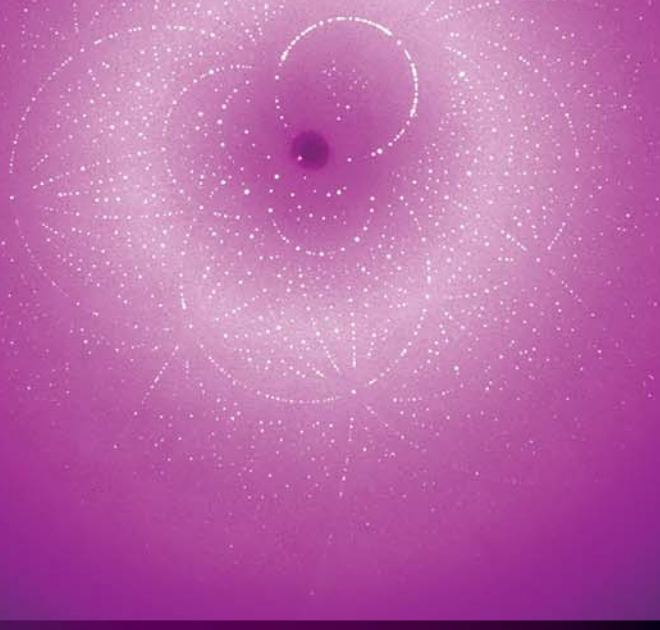
Jens Als-Nielsen • Des McMorrow

Elements of Modern X-ray Physics

Second Edition



WILEY

Elements of Modern X-ray Physics

Second Edition

Elements of Modern X-ray Physics

Second Edition

Jens Als-Nielsen

Ørsted Laboratory
Niels Bohr Institute
Copenhagen University

Des McMorrow

London Centre for Nanotechnology
University College London

WILEY

A John Wiley & Sons, Ltd Publication

This edition first published 2011

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Library of Congress Cataloging-in-Publication Data

Als-Nielsen, Jens
Elements of modern X-ray physics / Jens Als-Nielsen, Des McMorrow – 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-97395-0 (hardback) – ISBN 978-0-470-97394-3 (paper)
1. X-rays. I. Als-Nielsen, J. (Jens), 1937– II. Title.
QC481.A47 2011
539.7'222–dc22

2010051101

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

Print ISBN cloth: 978-0-470-97395-0
Print ISBN paper: 978-0-470-97394-3
ePDF ISBN: 978-1-119-99731-3
oBook ISBN: 978-1-119-99836-5
ePub ISBN: 978-1-119-97015-6
eMobi ISBN: 978-1-119-97016-3

Set by the authors using LaTeX.

Printed and bound in Singapore by Markono Print Media Pte Ltd.

Cover image by Michael Wulff.

Preface

In the decade or so since *Elements of Modern X-ray Physics* first appeared there has continued to be astonishing progress in the development of X-ray sources and the understanding of how to exploit them. This fact, taken together with the kind and generous comments we received in response to the first edition, has encouraged us to produce a second edition.

The second edition differs from the first in several key regards:

- An entirely new chapter on X-ray imaging has been included.
- The chapter dealing with kinematical diffraction has been divided into two separate chapters, which deal with non-crystalline and crystalline materials, respectively. This change has allowed us to include new material on the use of X-rays in the determination of the structure of liquids, glasses, and most importantly polymers and biomolecules.
- We have made many adjustments to various sections in the book with a view to improving the overall exposition.
- Small typographical errors have been corrected.
- Exercises have been included at the ends of all chapters except the first.

In preparing the second edition we have enjoyed the support and encouragement of many of our colleagues and friends without which our task would have been impossible. We would like to extend our deep gratitude to everyone who has contributed, most especially David Attwood, Martin Bech, Christian David, Martin Dierolf, Paul Emma, Kenneth Evans-Lutterodt, Per Hedegård, Mikael Häggström, John Hill, Moritz Hoesch, Torben Jensen, James Keeler, Ken Kelton, Carolyn Larabell, Bruno Lengeler, Anders Madsen, David Moncton, Theyencheri Narayan, Franz Pfeiffer, Harald Reichert, Ian Robinson, Jan Rogers, Joachim Stöhr, Joan Vila-Comamala, Simon Ward, and Tim Weitkamp.

Work on the second edition was initiated in Provence, France, during the summer of 2008 thanks to the generous support of the Ib Henriksen Foundation.

The front cover was designed by Marusa Design from an image kindly provided by Michael Wulff.

This book is dedicated to our respective families.

Jens Als-Nielsen and Des McMorrow

London, November 2010

Preface to the first edition

The construction of the first dedicated X-ray beamlines at synchrotron sources in the late 1970s heralded the start of a new era in X-ray science. In the intervening years tremendous progress has been made, both with respect to improvements to the sources, and with our knowledge of how to exploit them. Today's third-generation sources deliver extremely bright beams of radiation over the entire X-ray band (c. 1–500 keV), and with properties such as polarization, energy resolution, etc., that can be tailored to meet almost any requirement. These improvements have driven a surge of activity in X-ray science, and phenomena over a diverse range of disciplines can now be studied with X-rays that were undreamt of before the advent of synchrotron sources.

In light of these developments we believed that it was timely to produce a textbook at an introductory level. Our intention is to offer a coherent overview, which covers the basic physical principles underlying the production of X-rays, their interaction with matter, and also to explain how these properties are used in a range of applications. The main target audience for this book are final year undergraduates, and first year research students. Although the book has been written from the perspective of two physicists, we hope that it will be useful to the wider community of biologists, chemists, material scientists, etc., who work at synchrotron radiation facilities around the world. The main challenge in writing for a wider audience has been to convey the physical concepts without obscuring them in too much mathematical rigour. Therefore, many of the more difficult mathematical manipulations and theorems are explained in shaded boxes that may be studied separately. In addition appendices covering some of the required introductory physics have been included.

It is also our hope that this book will have appeal to more experienced research workers. Synchrotron radiation facilities are large laboratories where many different groups work on disparate areas of science. Cross fertilization of ideas is often the driving force of scientific progress. In order that these different groups, often working on neighbouring beamlines, can communicate their ideas, a common background is required. It is our intention that this book should provide at least some of this background knowledge. In addition, many X-ray techniques are becoming viewed as standard analytical tools, and it is no longer necessary to understand every aspect of the design of an instrument in order to be able to perform experiments. While this is undoubtedly a positive development, it can also be argued that a greater knowledge of the underlying principles not only adds to the overall feeling of satisfaction, but also allows better experiments to be designed.

This book has emerged from a lecture course that has been running for several years at the University of Copenhagen. The material covered in this book is taught in one semester, and is augmented by practical lessons both in an X-ray laboratory at the university, and also during a week long trip to the HASYLAB synchrotron facility. The list of subjects covered in this book inevitably reflects to some degree our own areas of specialization. There is, for example, very little on the vast and important subject of imaging. It was also decided at an early stage not to focus on subjects, such as classical crystallography, that we felt were well described in other texts. In spite of these shortcomings we hope that the reader, whatever his or her background, will learn something by studying this book, and be inspired to think of new ways to exploit the great opportunities that the development of synchrotron radiation offers.

Jens Als-Nielsen and Des McMorrow

Copenhagen, September 2000

Acknowledgements from the first edition

This book has grown out of our experiences of performing experiments at various synchrotron sources around the world. Our main thanks goes to our colleagues from these laboratories and elsewhere. In particular we would like to express our thanks to Henrik Bruus, Roger Cowley, Robert Feidenhans'l, Joseph Feldthaus, Francois Grey, Peter Gürtler, Wayne Hendrickson, Per Hedegård, John Hill, Mogens Lehmann, Les Leiserowitz, Gerd Materlik, David Moncton, Ian Robinson, Jochen Schneider, Horst Schulte-Schrepping, Sunil Sinha, and Larc Tröger for their detailed comments on different parts of the book. We are also indebted to the students at the Niels Bohr Institute who have attended the course on Experimental X-ray Physics. They have not only spotted untold numbers of typographical errors in early drafts, but also helped refine the material, and through their enthusiasm have ensured that teaching the course has been a rewarding, and even at times entertaining, experience. Birgitte Jacobsen deserves a special mention for her careful reading of the manuscript. Finally, we would like to thank Felix Beckmann, C.T. Chen, T.-C. Chiang, Trevor Forsyth, Watson Fuller, Malcolm McMahon, Benjamin Perman, and Michael Wulff for providing examples of their work, and Keld Theodor for help in preparing some of the figures.

This book has been typeset using LATEX, and we would like to express our thanks to everyone who has helped develop this system over the years, and in particular to Henrik Rønnow for helping with some of the trickier typesetting issues.

The image on the front cover was provided courtesy of Michael Wulff, ESRF, Grenoble, France.

Notes on the use of this book

The material in this book follows a more or less linear development. The scene is set in the first chapter, where the predominant mechanisms for the interaction of X-rays and matter are described. Many of the important concepts and results are introduced in this chapter, and forward references are made to the remaining chapters where these concepts are discussed more fully and the results derived. An attempt has been made to reduce to a minimum the level of mathematical skill required to follow the arguments. This has been done by placing most of the more taxing manipulations and theorems in shaded boxes, or in one of the appendices.

Computers are of course now an indispensable tool for helping to visualize mathematical and physical concepts. For this reason we have chosen to include a listing in the last appendix of some of the computer programmes that were used to generate the figures in this book. The hope is that this will ease the process of turning mathematical formulae into computer algorithms, and also aid the design of more complex programmes required for the analysis of data, etc. The programmes have been written using the MATLAB® programming environment, although the way that they are derived from the mathematics is transparent enough that they can easily be converted to other languages. Figures for which programme listings are given are indicated by a star, ⋆.

Contents

1

X-rays and their interaction with matter

X-rays were discovered by Wilhelm Conrad Röntgen in 1895. Since that time they have become established as an invaluable probe of the structure of matter. The range of materials for which X-rays have proved to be decisive in unravelling the structure is truly staggering. These include at one limit of complexity simple compounds, through to more complex and celebrated examples, such as DNA. In more recent times the structure of proteins, and even functional units of living organisms, can be solved on a regular basis. Progress in both our theoretical understanding of the interaction of X-rays with matter, and in our knowledge of how to exploit them experimentally, was steady from the period covering their discovery through to the mid 1970s. The main limitation in this period was the source, which had remained essentially unchanged from about 1912. In the 1970s it was realized that the synchrotron radiation emitted from charged particles circulating in storage rings constructed for high energy nuclear physics experiments was potentially a much more intense and versatile source of X-rays. Indeed synchrotrons have proven to be such vastly better sources that many storage rings have been constructed around the world dedicated solely to the production of X-rays.

This has culminated to date in the so-called third-generation synchrotron sources, which are more brilliant than the early lab-based sources by a factor of approximately 10^{12}, as indicated in Fig. 1.1. With the advent of synchrotron sources the pace of innovation in X-ray science increased markedly (though perhaps not a trillion fold!), and today shows no signs of slowing. The first X-ray free-electron lasers have recently come into service, and when they become fully operational further important breakthroughs will undoubtedly follow. In Chapter 2 we explain the basic physical principles of X-ray sources and outline their salient properties.

In Fig. 1.2 we show a schematic of the key components of a typical experimental beamline at a third-generation source. The details will of course vary considerably depending on the particular requirements, but many of the components shown will be found in one form or another on most beamlines. First there is the source itself. In this case the electrons do not follow a purely circular orbit in the storage ring, but traverse through straight sections where lattices of magnets, so-called undulator insertion devices, force them to execute small-amplitude oscillations. At each oscillation X-rays are emitted and, if the amplitude of the oscillations is small, then the different contributions from the passage of a single electron add coherently, and a very intense beam of X-rays results. The second key component is the monochromator, as in many applications it is required to work at a particular average wavelength. It may also be desirable to choose the wavelength bandwidth, and monochromators made

Elements of Modern X-ray Physics, Second Edition. Jens Als-Nielsen and Des McMorrow

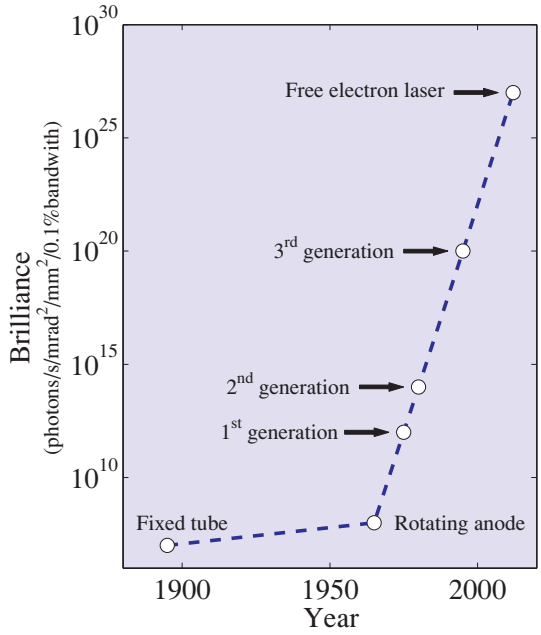


Fig. 1.1 The brilliance of X-ray sources as a function of time. Source brilliance is defined and discussed in Chapter 2, along with the principles underlying the production of X-rays from synchrotrons and free-electron lasers. For free-electron laser sources we plot the average brilliance. Due to the extremely short X-ray pulse length from a free-electron laser – of order 100 fs – the peak brilliance exceeds the average brilliance by a large factor.

from perfect crystals through to multilayers allow for a considerable variation in this parameter. Thirdly, if working with small samples it may be desirable to focus the monochromatic beam down to as small a size as achievable. This is accomplished by devices such as X-ray mirrors and refractive Fresnel lenses. Finally, X-rays are delivered to the sample itself on which the experiment is performed.

One of the main goals of this book is to explain the physical principles underlying the operation of the key components shown in Fig. 1.2. As a first step it is necessary to understand some of the basic aspects of the interaction of X-rays with matter.

1.1 X-rays: waves and photons

X-rays are electromagnetic waves with wavelengths in the region of an Ångström (10^{-10}m). In many cases one is interested in a monochromatic beam of X-rays as depicted in Fig. 1.3. The direction of the beam is taken to be along the z-axis, perpendicular to the electric, $\mathbf{E}$, and magnetic, $\mathbf{H}$, fields. For simplicity, we shall start by considering the electric field only and neglect the magnetic field. The top part of Fig. 1.3 shows the spatial dependence of the electromagnetic field at a given instance of time. It is characterized by the wavelength λ, or equivalently the wavenumber $k = 2\pi/\lambda$. Mathematically the electric field amplitude is expressed as a sine wave, either in its real form, $E_0 \sin(kz)$, or in its more compact complex form, $E_0 e^{ikz}$.

The lower part of Fig. 1.3 is an alternative illustration of the monochromatic plane wave. Only the wave crests are shown (full lines perpendicular to the z-axis), emphasizing that it is a plane wave with

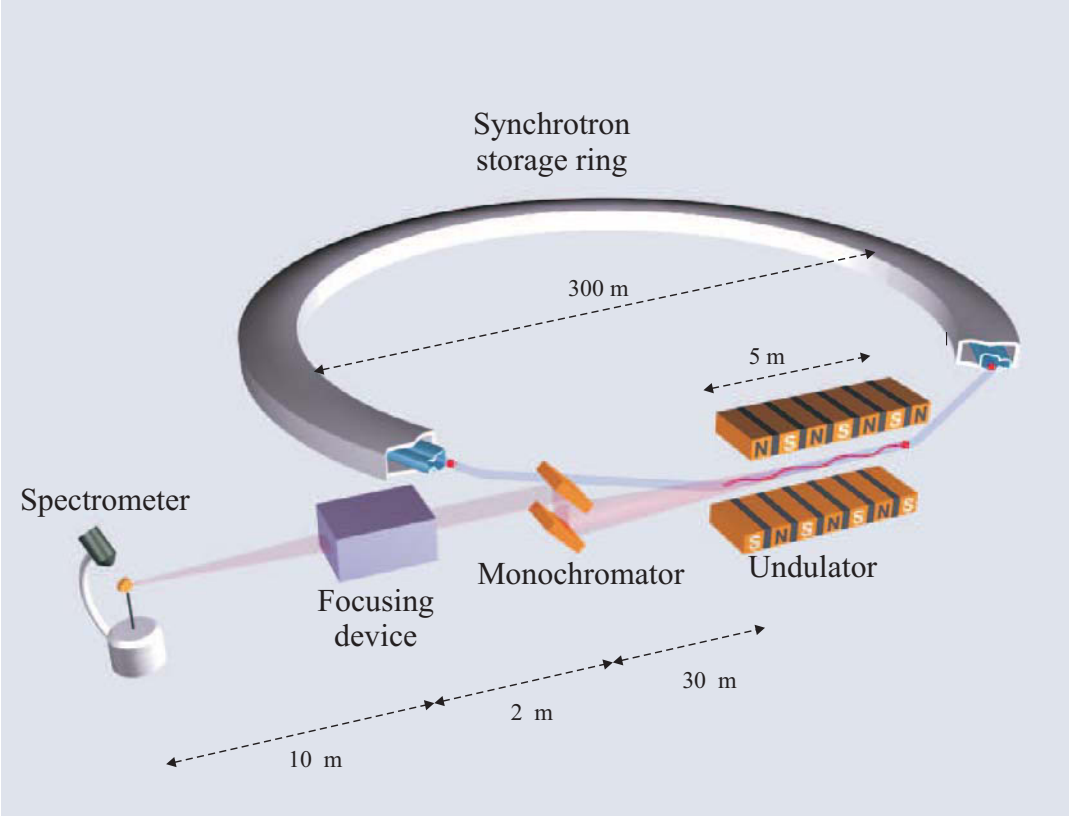


Fig. 1.2 A schematic of a typical X-ray beamline at a third generation X-ray source. Bunches of charged particles (electrons or positrons) circulate in a storage ring (typical diameter around 300 m). The ring is designed with straight sections, where an insertion device, such as undulator, is placed. The lattice of magnets in an insertion device forces the particles to execute small oscillations which produce intense beams of radiation. This radiation then passes through a number of optical elements, such as a monochromator, focusing device, etc., so that a beam of radiation with the desired properties is delivered to the sample. Typical distances are indicated.

an electric field that is constant anywhere in a plane perpendicular to the z-axis. Although a beam is never ideally collimated, the approximation of a plane wave is often valid. The spatial and temporal variation of a plane wave propagating along the z-axis can be encompassed in one simple expression, $\mathrm{E}_0\, \mathrm{e}^{i\,(\mathrm{k}z-\omega t)}$. More generally in three dimensions the polarization of the electric field is written as a unit vector $\hat{\boldsymbol{\varepsilon}}$, and the wavevector along the direction of propagation as $\mathbf{k}$, so that

$$\mathbf{E}(\mathbf{r}, t) = \hat{\boldsymbol{\varepsilon}}\, \mathrm{E}_0\, \mathrm{e}^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$$

Since electromagnetic waves are transverse we have $\hat{\boldsymbol{\varepsilon}} \cdot \mathbf{k}=0$, and $\mathbf{k} \cdot \mathbf{E}= \mathbf{k} \cdot \mathbf{H}=0$ as shown in Fig. 1.4.

This is the classical description of a linearly polarized, electromagnetic plane wave. From a quantum mechanical perspective, a monochromatic beam is viewed as being quantized into photons, each having an energy $\hbar\omega$ and momentum $\hbar\mathbf{k}$. The intensity of a beam is then given by the number of photons passing through a given area per unit time. As the intensity is also proportional to the square of the electric field, it follows that the magnitude of the field is quantized. Instead of quantizing the **E** and **H** fields separately, it turns out to be more convenient to work with the vector potential **A**, since both **E**

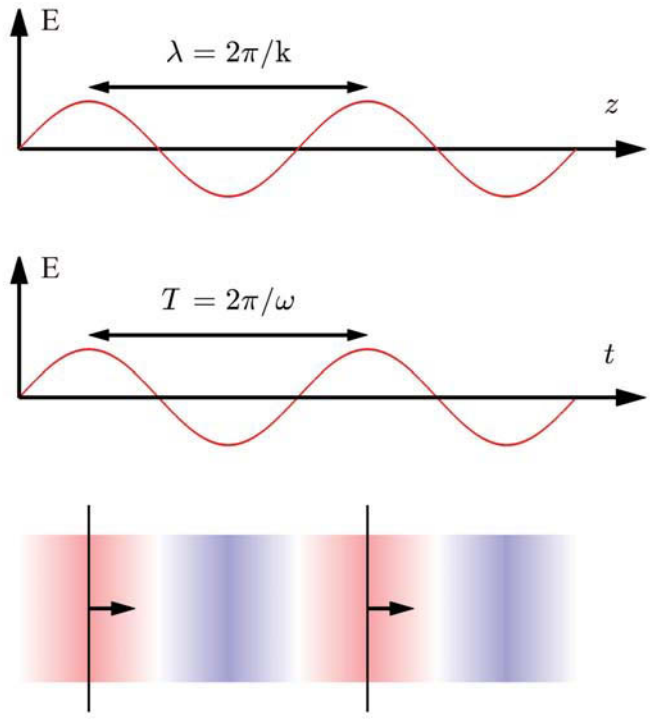


Fig. 1.3 Three representations of an electromagnetic plane wave. Only the electric field **E** is shown. Top: spatial variation, described by the wavelength λ or the wavenumber k, at a given instant in time. Middle: temporal variation, described by the period T or the cyclic frequency ω, at a given point in space. Bottom: Top view of a plane wave with the wave crests indicated by the heavy lines, and the direction of propagation by the arrows. The shading indicates the spatial variation of the amplitude of the field.

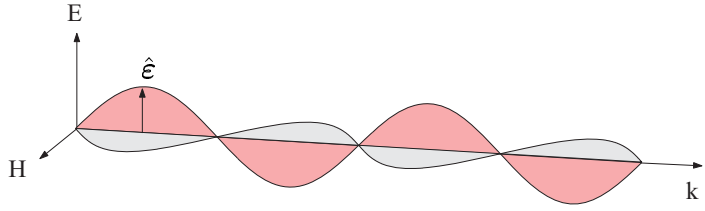


Fig. 1.4 An X-ray is a transverse electromagnetic wave, where the electric and magnetic fields, **E** and **H**, are perpendicular to each other and to the direction of propagation **k**. The direction of the electric field is given by the polarization unit vector $\hat{\boldsymbol{\epsilon}}$.

and **H** can be derived from **A**. In Appendix C it is explained how the vector potential is quantized, and the explicit form of the quantum mechanical Hamiltonian of the electromagnetic field is given. In this book we shall move freely between the classical and quantum descriptions, choosing whichever one leads us to the quickest and clearest understanding of the problem at hand.

The numerical relation[1] between wavelength λ in Å and photon energy $\mathcal{E}$ in keV is

$$\lambda\,[\text{Å}] = \frac{hc}{\mathcal{E}} = \frac{12.398}{\mathcal{E}\,[\text{keV}]} \tag{1.1}$$

An X-ray photon interacts with an atom in one of two ways: it can be scattered or it can be absorbed, and we shall discuss these processes in turn. When X-rays interact with a dense medium consisting of a very large number of atoms or molecules it is sometimes more convenient to treat the material as a continuum, with an interface to the surrounding vacuum (or air). At the interface the X-ray beam is refracted and reflected, and this is an alternative way in which the interaction may be discussed. The scattering and refraction descriptions are of course equivalent. In Chapter 3 we derive the X-ray reflectivity equations, and exploit this equivalence to relate the reflectivity to the microscopic properties of the medium of interest.

1.2 Scattering

To start with we shall consider the scattering of an X-ray by a single electron. In the classical description of the scattering event the electric field of the incident X-ray exerts a force on the electronic charge, which then accelerates and radiates the scattered wave. Classically, the wavelength of the scattered wave is the same as that of the incident one, and the scattering is necessarily *elastic*. This is not true in general in a quantum mechanical description, where the incident X-ray photon has a momentum of $\hbar\mathbf{k}$ and an energy of $\hbar\omega$. Energy may be transferred to the electron with the result that the scattered photon has a lower frequency relative to that of the incident one. This *inelastic* scattering process is known as the Compton effect, and is discussed at the end of this section. However, the elastic scattering of X-rays is the main process that is exploited in investigations of the structure of materials, and in this case it suffices to adopt what is essentially a classical approach.

One electron

The most elementary scattering object that we shall consider is a single, free electron. The ability of an electron to scatter an X-ray is expressed in terms of a *scattering length*, which we shall now derive.

Figure 1.5 shows a schematic of a generic scattering experiment. The fundamental quantity determined in such an experiment is the *differential scattering cross-section* $(d\sigma/d\Omega)$ which is defined by

$$\left(\frac{d\sigma}{d\Omega}\right) = \frac{I_{sc}}{\Phi_0\,\Delta\Omega} \tag{1.2}$$

The strength of the incident beam is given by the flux, Φ_0, which is simply the number of photons passing through unit area per second. The incident beam interacts with the scattering object and is scattered. The number of scattered photons recorded per second in a detector is I_{sc}, where the detector is positioned a distance R away from the object and subtends a solid angle $\Delta\Omega$. The differential cross-section is thus a measure of the efficiency of the scattering process where details of the experiment,

[1] In this book we shall mostly limit the wavelength band to 0.1 – 2 Å corresponding to the energy band 120 – 6 keV. The first limit, 0.1 Å or 120 keV, ensures that relativistic effects are negligible since the X-ray energy is considerably lower than the rest mass of the electron, mc^2= 511 keV. The second limit, 2 Å or 6 keV, is a practical limit ensuring that the X-rays have a high penetration power through light materials, such as beryllium. In many X-ray tubes, and in synchrotron radiation beam lines, the X-rays must be transmitted through a Be window, and above 6 keV the transmission of a 0.5 mm Be window exceeds 90%. Lower energy X-rays are called soft X-rays and will not be dealt with in this book.

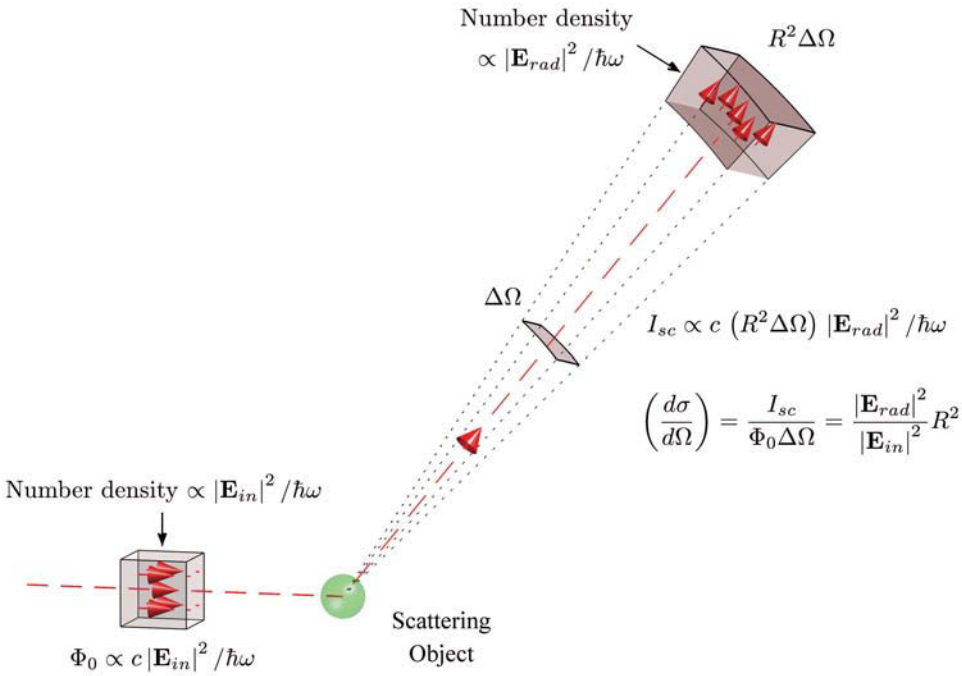


Fig. 1.5 Schematic layout of a generic scattering experiment used to determine the differential cross-section ($d\sigma/d\Omega$): see Eq. (1.2). The incident beam flux Φ_0 is the number of particles per second per unit area. For an electromagnetic wave this is proportional to $|\mathbf{E}_{in}|^2$ times the velocity of light, c. The incident beam interacts with the target object to produce the scattered beam. A detector records the scattered intensity, I_{sc}, defined as the number of counts recorded per second, which is proportional to $|\mathbf{E}_{rad}|^2$ times the area of the detector and the velocity of light. The detector is located a distance R from the target object, and subtends a solid angle of $\Delta\Omega$.

specifically the flux of the incident beam and size of the detector, have been normalized away. (See Appendix A for a more complete discussion.)

For the particular case of the scattering of an electromagnetic wave indicated in Fig. 1.5 an expression for Φ_0 can be obtained in terms of the electric field $\mathbf{E}_{in}$ of the incident beam. Since the energy density is proportional to $|\mathbf{E}_{in}|^2$, the number density of photons is proportional to $|\mathbf{E}_{in}|^2 / \hbar\omega$, while the flux Φ_0 is the number density multiplied by speed of light, c. (This follows from the realisation that in one second a beam of area A sweeps out a volume equal to $A\,c$.) A similar argument applies to the intensity I_{sc} of the scattered beam. In this case the number density is proportional to the modulus squared of the radiated electric field, $|\mathbf{E}_{rad}|^2$. This quantity then must be multiplied by the area of the detector, $R^2\Delta\Omega$, and c to yield and expression for I_{sc}. With these considerations, the differential cross-section is given by

$$\left(\frac{d\sigma}{d\Omega}\right) = \frac{|\mathbf{E}_{\mathrm{rad}}|^2 R^2}{|\mathbf{E}_{\mathrm{in}}|^2} \tag{1.3}$$

In a classical description of the scattering process an electron will be forced to vibrate when placed in the electric field of an incident X-ray beam, as illustrated in Fig. 1.6(a). A vibrating electron acts as a source, and radiates a spherical wave $\mathbf{E}_{rad} \propto \hat{\boldsymbol{\varepsilon}}'\, e^{ikR}/R$. The problem then is to evaluate the radiated field at an observation point X. This calculation is performed in Appendix B starting from

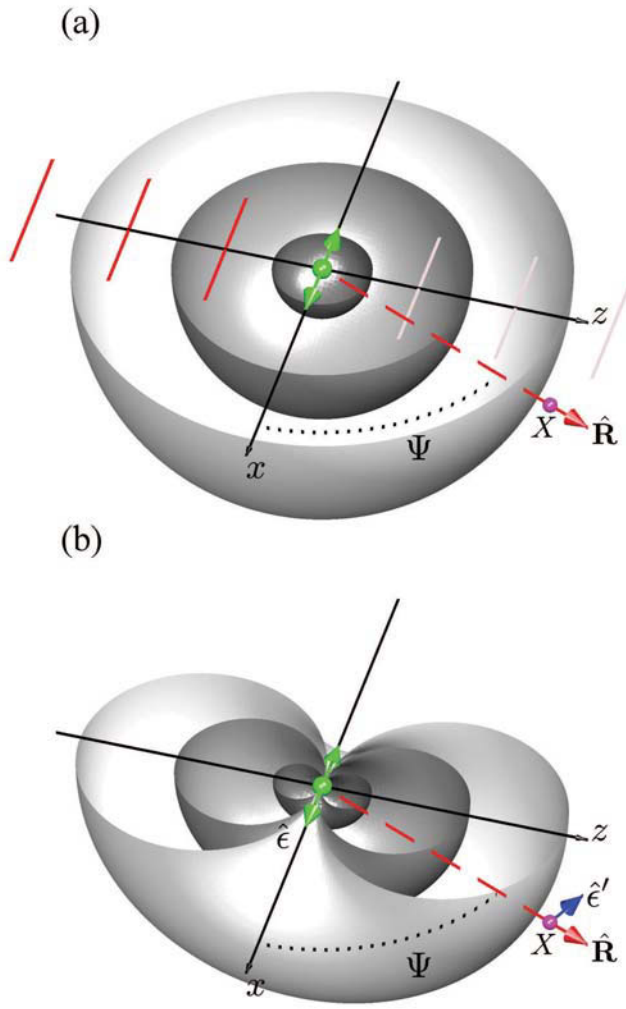


Fig. 1.6 The classical description of the scattering of an X-ray by an electron. (a) The electric field of an incident plane wave sets an electron in oscillation which then radiates a spherical wave. (For clarity the radiated wave is shown for positive values of y only, and for the simplest case of an isotropic spherical wave the phase and amplitude are constant on spherical surfaces.) The incident wave propagates along the z axis and has its electric field polarized along x. The wave crests of the incident wave lie in between those of the scattered spherical wave because of the 180° phase shift in Thomson scattering. In the text the radiated field at an observation point X is calculated. Point X lies in the plane spanned by the polarization vector and the propagation direction of the incident wave, and the observed acceleration has to be multiplied by a factor of $\sin\Psi$. (b) From geometry, $\sin\Psi = -\hat{\boldsymbol{\varepsilon}} \cdot \hat{\boldsymbol{\varepsilon}}'$ where $\hat{\boldsymbol{\varepsilon}}$ ($\hat{\boldsymbol{\varepsilon}}'$) represents the polarization of the incident (scattered) beam. The effect of this factor on the radiated wave is illustrated by plotting surfaces of constant amplitude.

Maxwell's equations. Here, an heuristic argument is outlined. Initially we consider the situation where the observation point X lies in the plane spanned by the polarization vector and the propagation direction of the incident wave, and at an angle $90° - \Psi$ with respect to the direction of propagation of the incident beam (Fig. 1.6(a)).

The radiated field is proportional to the charge of the electron, $-e$, and to the acceleration, $a_X(t')$, evaluated at a time t' earlier than the observation time t due to finite speed c at which the radiation propagates. The radiated field is thus expected to be of the form

$$\mathrm{E}_{\mathrm{rad}}(R,t) \propto \frac{-e}{R}\, a_X(t') \sin\Psi \tag{1.4}$$

where $t' = t - R/c$. The total energy flow through a spherical shell of radius R is the energy density, proportional to $\left|\mathbf{E}_{\mathrm{rad}}\right|^2$, multiplied by the surface area, proportional to R^2, so with $\left|\mathbf{E}_{\mathrm{rad}}\right| \propto R^{-1}$ the total energy flow becomes independent of R, as it must[2]. A further factor of $\sin\Psi$ has been included to allow for the variation of the acceleration with observation angle. For an observer at point X in the $x-z$ plane, the acceleration observed is zero for $\Psi = 0°$, and a maximum for $\Psi = 90°$. Therefore, the acceleration observed is the full acceleration multiplied by $\sin\Psi$.

To proceed we evaluate the full acceleration from the force on the electron divided by its mass, which yields

$$a_X(t') = \frac{-e\,\mathrm{E}_0\,\mathrm{e}^{-i\omega t'}}{m} = \frac{-e}{m}\,\mathrm{E}_{\mathrm{in}}\,\mathrm{e}^{i\omega(R/c)} = \frac{-e}{m}\,\mathrm{E}_{\mathrm{in}}\,\mathrm{e}^{ikR}$$

where $\mathrm{E}_{\mathrm{in}} = \mathrm{E}_0\mathrm{e}^{-i\,\omega t}$ is the electric field of the incident wave. Hence Eq. (1.4) can be rearranged to read

$$\frac{\mathrm{E}_{\mathrm{rad}}(R,t)}{\mathrm{E}_{\mathrm{in}}} \propto \left(\frac{e^2}{m}\right)\frac{\mathrm{e}^{ikR}}{R}\sin\Psi \tag{1.5}$$

For an observation point at an arbitrary angle with respect to the polarization of the incident beam, the factor of $\sin\Psi$ must be reevaluated. If $\hat{\boldsymbol{\varepsilon}}$ is the polarization of the incident field, and $\hat{\boldsymbol{\varepsilon}}'$ that of the radiated field, then from Fig. 1.5(b), $\hat{\boldsymbol{\varepsilon}} \cdot \hat{\boldsymbol{\varepsilon}}' = \cos(90° + \Psi) = -\sin(\Psi)$. The advantage of writing the trigonometric factor for the apparent acceleration in this way is that it is valid for all possible angles of observation. This is ensured by the azimuthal symmetry evident in Fig. 1.6(b) of the radiated field around the x axis.

To complete the derivation of the differential cross-section it is necessary to check whether we have the correct units. Clearly, the ratio of electric fields given in Eq. (1.5) is dimensionless. This requires that whatever the factor is multiplying the spherical wave form e^{ikR}/R, it must have units of length. The appropriate length can be found by noting that in SI units the Coulomb energy at distance r from a point charge $-e$ is $e^2/(4\pi\epsilon_0 r)$, while dimensionally the energy is also of the form mc^2. Thus equating these two expressions for energy and rearranging provides an expression for the fundamental length scale in the problem, namely

$$r_0 = \left(\frac{e^2}{4\pi\epsilon_0 mc^2}\right) = 2.82\times10^{-5}\ \text{Å} \tag{1.6}$$

This is referred to as the Thomson scattering length, or classical radius, of the electron. While these arguments fix the magnitude of the scattering length, they do not fix its phase. As shown in Appendix

[2]Equation (1.4) represents the electric field radiated by an oscillating dipole in the far-field limit.

B, the scattering amplitude from a single electron is in fact equal to $-r_0 \left|\hat{\boldsymbol{\varepsilon}} \cdot \hat{\boldsymbol{\varepsilon}}'\right|$. Physically, the factor of -1 represents the fact that there is a 180° phase shift between the incident and scattered waves. This phase shift also has consequences for the refractive index n which in the X-ray region is less than unity, as discussed in Section 1.4 and Chapter 3.

The ratio of radiated to incident electric fields is therefore

$$\frac{\mathrm{E}_{\mathrm{rad}}(R,t)}{\mathrm{E}_{\mathrm{in}}} = -r_0 \frac{\mathrm{e}^{ikR}}{R} \left|\hat{\boldsymbol{\varepsilon}} \cdot \hat{\boldsymbol{\varepsilon}}'\right| \tag{1.7}$$

and from Eq. (1.3) the differential cross-section becomes

$$\boxed{\left(\frac{d\sigma}{d\Omega}\right) = r_0^2 \left|\hat{\boldsymbol{\varepsilon}} \cdot \hat{\boldsymbol{\varepsilon}}'\right|^2} \tag{1.8}$$

This equation describes the Thomson differential scattering cross-section of an electromagnetic wave by a free electron.

The factor of $\left|\hat{\boldsymbol{\varepsilon}} \cdot \hat{\boldsymbol{\varepsilon}}'\right|^2$ has important implications for the choice of optimal geometry for different types of X-ray experiments. For example, synchrotron sources naturally produce X-rays which are linearly polarized in the horizontal plane of the synchrotron. It follows that scattering experiments are best performed in a vertical scattering plane as then $\left|\hat{\boldsymbol{\varepsilon}} \cdot \hat{\boldsymbol{\varepsilon}}'\right|^2 = 1$ independent of the scattering angle, $\psi = 90° - \Psi$. Conversely, if one wants to study fluorescence from a sample, then it is possible to suppress the scattering by working in the horizontal plane at $\psi = 90°$ since $\left|\hat{\boldsymbol{\varepsilon}} \cdot \hat{\boldsymbol{\varepsilon}}'\right|^2 = 0$. These considerations lead us to define P, the polarization factor for scattering, which depends on the X-ray source:

$$P = \left|\hat{\boldsymbol{\varepsilon}} \cdot \hat{\boldsymbol{\varepsilon}}'\right|^2 = \begin{cases} 1 & \text{synchrotron: vertical scattering plane} \\ \cos^2\psi & \text{synchrotron: horizontal scattering plane} \\ \frac{1}{2}\left(1 + \cos^2\psi\right) & \text{unpolarized source} \end{cases} \tag{1.9}$$

The *total* cross-section for Thomson scattering is found by integrating the differential cross-section over all possible scattering angles. Exploiting the rotational symmetry of the radiated field around $\hat{\boldsymbol{\varepsilon}}$, it can be shown that the average value of $< \left(\hat{\boldsymbol{\varepsilon}} \cdot \hat{\boldsymbol{\varepsilon}}'\right)^2 >$ over the unit sphere is (2/3). Thus the total cross-section σ_{T} is equal to $4\pi r_0^2 \times (2/3) = 8\pi r_0^2/3 = 0.665 \times 10^{-24}$ cm^2 =0.665 barn. It is evident that the classical cross-section, both the differential and total, for the scattering of an electromagnetic wave by a free electron is a constant, independent of energy. This result is particularly relevant to the X-ray part of the electromagnetic spectrum, as here a photon is energetic enough that even atomic electrons respond to a good approximation as if they are free. Where it breaks down entirely is at low energies in the optical part of the spectrum, or when the energy of photon passes a threshold for resonantly exciting electrons from deeply bound atomic states, as outlined later in this section. In Chapter 8 we discuss the origin and consequences of such resonant scattering processes.

Finally in this section we note that the classical derivation of the scattering of a photon by a free electron given here yields the same result as the full quantum mechanical derivation given in Appendix C.

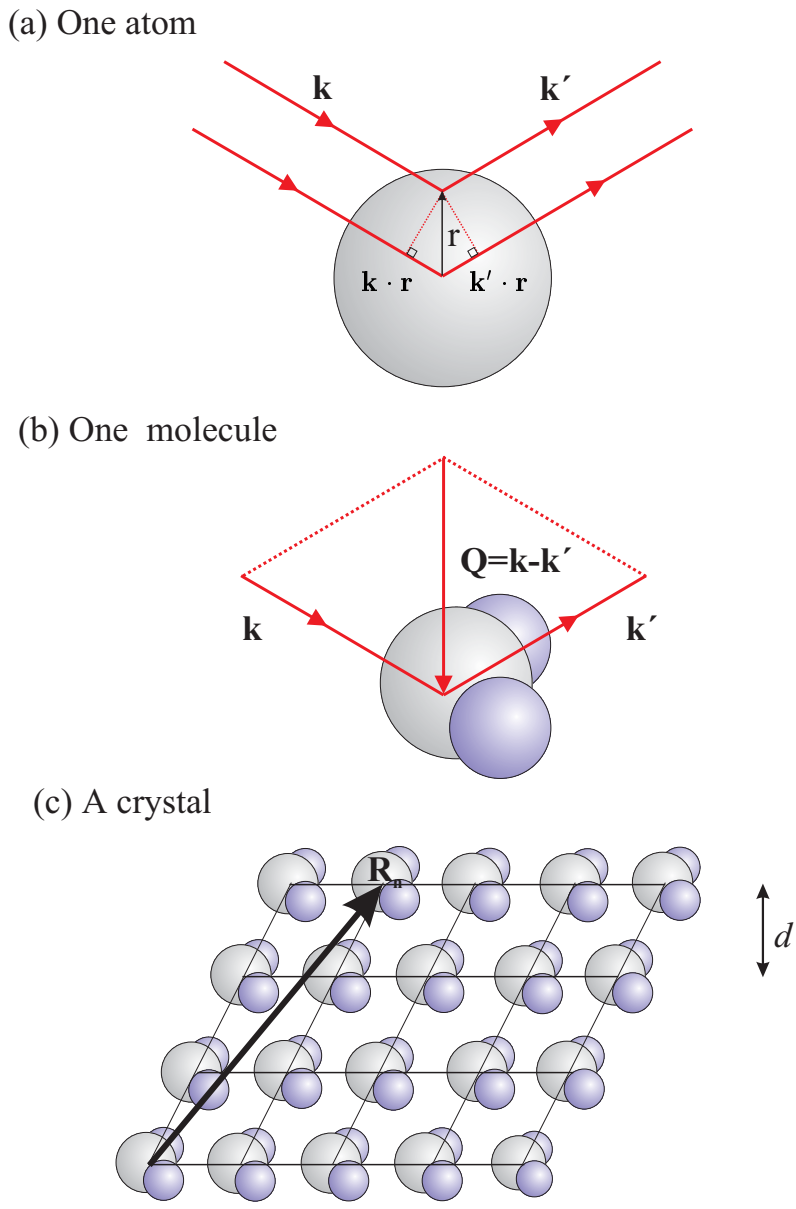


Fig. 1.7 (a) Scattering from an atom. An X-ray with a wavevector $\mathbf{k}$ scatters from an atom to the direction specified by $\mathbf{k}'$. The scattering is assumed to be elastic, i.e. $|\mathbf{k}| = |\mathbf{k}'| = 2\pi/\lambda$. The difference in phase between a wave scattered at the origin and one at a position $\mathbf{r}$ is $(\mathbf{k} - \mathbf{k}') \cdot \mathbf{r} = \mathbf{Q} \cdot \mathbf{r}$. This defines the wavevector transfer $\mathbf{Q}$. (b) The scattering from a molecule. Here the scattering triangle is shown which relates $\mathbf{k}$, $\mathbf{k}'$ and $\mathbf{Q}$. (c) Scattering from a molecular crystal. The molecules are organized on a lattice with position vectors $\mathbf{R_n}$, and a lattice plane spacing of d.

One atom

Let us now proceed from the scattering by a single electron to consider the elastic scattering from an atom with Z electrons.

To start with a purely classical description will be used, so that the electron distribution is specified by a number density, $\rho(\mathbf{r})$. The scattered radiation field is a superposition of contributions from different volume elements of this charge distribution. In order to evaluate this superposition one must keep track of the phase of the incident wave as it interacts with the volume element at the origin and the one at position $\mathbf{r}$, as shown in Fig. 1.7(a). The phase difference between two successive crests is 2π. The phase difference between the two volume elements is 2π multiplied by the ratio of $\mathbf{r}$, projected onto the incident direction, and the wavelength. This is nothing other than the scalar product of the two vectors $\mathbf{k}$ and $\mathbf{r}$. The simplicity of this expression is one of the reasons why it is so convenient to use the wavevector $\mathbf{k}$ to describe the incident wave. In the vicinity of the observation point X in Fig. 1.6, the scattered wave is locally like a plane wave with wavevector $\mathbf{k}'$. The phase difference, between the scattered wave from a volume element around the origin and one around $\mathbf{r}$ is $-\mathbf{k}' \cdot \mathbf{r}$. The resulting phase difference is thus

$$\Delta\phi(\mathbf{r}) = (\mathbf{k} - \mathbf{k}') \cdot \mathbf{r} = \mathbf{Q} \cdot \mathbf{r}$$

where

$$\boxed{\mathbf{Q} = \mathbf{k} - \mathbf{k}'} \tag{1.10}$$

$\mathbf{Q} = \mathbf{k} - \mathbf{k}'$ is known as the *wavevector transfer* or *scattering vector*. The scattering events depicted in Fig. 1.7 are elastic, with $|\mathbf{k}| = |\mathbf{k}'|$, so that from the scattering triangle we have $|\mathbf{Q}| = 2|\mathbf{k}| \sin\theta = (4\pi/\lambda)\sin\theta$. As we shall see, $\mathbf{Q}$ is the natural variable to describe elastic scattering processes and is usually expressed in units of Å^{-1}.

Thus a volume element $d\mathbf{r}$ at $\mathbf{r}$ will contribute an amount $-r_0\rho(\mathbf{r})d\mathbf{r}$ to the scattered field with a phase factor of $e^{i\mathbf{Q}\cdot\mathbf{r}}$. The total scattering length of the atom is

$$\boxed{-r_0 f^0(\mathbf{Q}) = -r_0 \int \rho(\mathbf{r})\, e^{i\mathbf{Q}\cdot\mathbf{r}}\, d\mathbf{r}} \tag{1.11}$$

where $f^0(\mathbf{Q})$ is known as the *atomic form factor*. In the limit that $Q \to 0$ all of the different volume elements scatter in phase so that $f^0(\mathbf{Q} = 0) = Z$, the number of electrons in the atom. As Q increases from zero the different volume elements start to scatter out of phase and consequently $f^0(Q \to \infty) = 0$. The right hand side of Eq. (1.11) is recognizable as a Fourier transform. Indeed one of the recurrent themes of this book is that the scattering length may be calculated from the Fourier transform of the distribution of electrons in the sample[3]. It should be clear that to calculate the scattered intensity we have to evaluate Eq. (1.11) and multiply by its complex conjugate (see Eq. (1.2) and accompanying discussion).

Atomic electrons are of course governed by quantum mechanics, and have discrete energy levels. The most tightly bound electrons are those in the K shell, which have energies comparable to those of a typical X-ray photon. If the X-ray photon has an energy much less than the binding energy of the K shell, the response of these electrons to an external driving field is reduced by virtue of the fact that they are bound. Electrons in shells that are less tightly bound (L, M, etc.) will be able to respond to the

[3]The reader is reminded of the definition and properties of Fourier transforms in Appendix E.

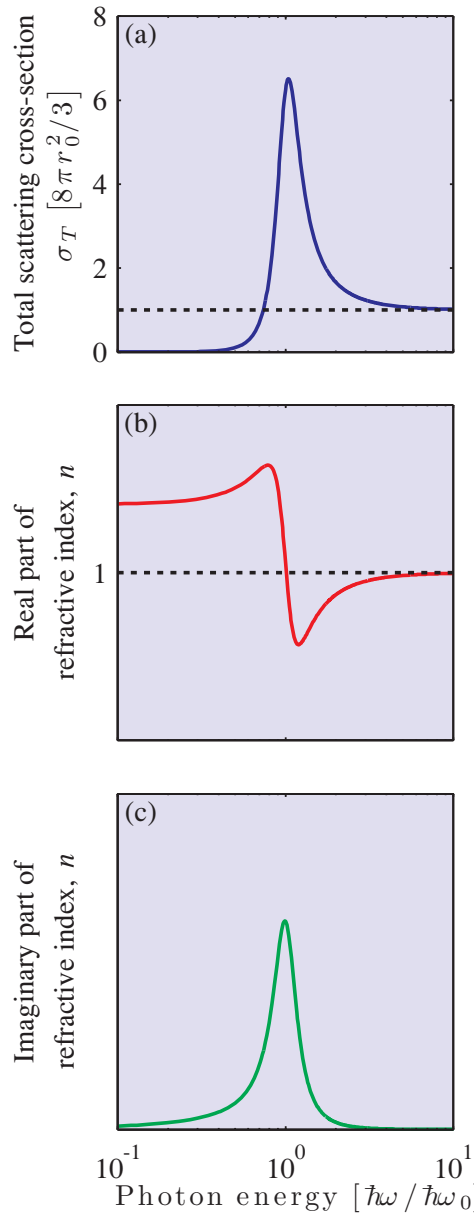


Fig. 1.8 The calculated frequency dependence of (a) the total Thomson scattering cross-section, and the real (b) and imaginary (c) parts of the refractive index, n, when including the dispersion corrections to the Thomson scattering (see Section 8.1). In general, the X-ray part of the electromagnetic spectrum corresponds to the high-frequency (or energy) limit. In this limit, the total scattering cross-section approaches that from a free electron, $\sigma_T = 8\pi r_0^2/3$, and the real part of the refractive index is less than one. It should be noted that important resonances, associated with the K, L and M absorption edges, occur in the X-ray part of the spectrum as discussed in Chapters 7 and 8. For clarity the width of the resonance at $\hbar\omega_0$ has been exaggerated.

driving field more closely, but overall we expect that the scattering length of an atom to be reduced by some amount, which is by convention denoted f'. At energies much greater than the binding energy the electrons can be treated as if they are free and f' is zero. For energies in between these limits f' displays resonant behaviour at energies corresponding to atomic absorption edges, which are discussed in Section 1.3. In addition to altering the real part of the scattering length, we also expect that, by analogy with a forced harmonic oscillator, the response of the electron to have a phase lag with respect to the driving field. This is allowed for by including a term if'', which represents the dissipation in the system, and, as we shall see in Chapters 3 and 8, it is related to the absorption. Collecting these results together means that the atomic form factor is

$$\boxed{f(\mathbf{Q},\hbar\omega) = f^0(\mathbf{Q}) + f'(\hbar\omega) + if''(\hbar\omega)} \tag{1.12}$$

where f' and f'' are known as the dispersion corrections[4] to f^0. We have written f' and f'' as functions of the X-ray energy $\hbar\omega$ to emphasize that their behaviour is dominated by tightly bound inner-shell electrons, and as a consequence cannot have any appreciable dependence on $\mathbf{Q}$. As might be expected from these introductory remarks, f' and f'' assume their extremal values when the X-ray energy is equal to one of the absorption edge energies of the atom. This resonant behaviour is manifestly element specific, and in Chapter 8 it is explained how it may be exploited to solve the structure of complex materials.

In Fig. 1.8(a) we illustrate the effects of including the dispersion corrections in a calculation of the total scattering cross-section (see Section 8.1). For frequencies much less than $\hbar\omega_0$ the binding of the electron drastically reduces the cross-section. When $\omega \approx \omega_0$ the cross section is significantly enhanced. It is only at high frequencies that the electrons behave as if they are free and the value of the total cross-section calculated for Thomson scattering is realised. Figure 1.8(b) shows the calculated variation in the real part of the refractive index n as a function of photon energy. For $\hbar\omega \ll \hbar\omega_0$ the real part of the refractive index tends to a constant greater than unity, whereas above $\hbar\omega_0$ it is less than unity as is found to be true for X-rays. The consequences of the real part of n being less than unity for X-rays are introduced in Section 1.4 and discussed further in Chapter 3. Figure 1.8 also serves to further underline the fact that scattering and refraction of electromagnetic waves are essentially different views of the same physical phenomenon.

One molecule

So far we have introduced the scattering length for an electron and subsequently for an atom composed of electrons. The next step in complexity is naturally molecules composed of atoms (Fig. 1.7(b)). It is obvious that just as the scattering length of an atom has a form factor, so will the scattering length of a molecule. Labelling the different atoms in the molecule by index j we may write

$$F^{\mathrm{mol}}(\mathbf{Q}) = \sum_j f_j(\mathbf{Q})\, \mathrm{e}^{i\mathbf{Q}\cdot\mathbf{r}_j}$$

where as before $f_j(\mathbf{Q})$ is the atomic form factor of the j'th atom in the molecule, and it must be remembered to include the multiplicative factor of $-r_0$ if the intensity is required in absolute units. If one can determine $\left|F^{\mathrm{mol}}(\mathbf{Q})\right|^2$ experimentally for sufficiently many values of scattering vector $\mathbf{Q}$ then

[4] These are also sometimes referred to as the anomalous dispersion corrections, but it is generally agreed that there is in fact nothing anomalous about them. It should be noted that with our sign convention f'' is negative.

one can (at least by trial and error) determine the positions $\mathbf{r}_j$ of the atoms in the molecule. However, the scattering length of a single molecule is not sufficient to produce a measurable signal, even in the very intense X-ray beams produced by today's synchrotron sources. For that bulk samples containing many molecules are required, assembled either as non-crystalline or crystalline forms of matter. The scattering from these distinct phases of matter are dealt with in Chapters 4 and 5, respectively. However, it is expected that in the future, the spectacular increase in peak brilliance offered by free-electron sources will permit the imaging of single molecules.

A crystal

The defining property of a crystalline material is that it is periodic in space[5], as shown for a molecular crystal in Fig. 1.7(c). In elementary treatments of the scattering of X-rays from a crystal lattice, Bragg's law

$$m\lambda = 2d\sin\theta$$

is derived, where m is an integer. This is the condition for the constructive interference of waves which have an angle of incidence θ to a set of lattice planes a distance d apart. While this is a useful construction, it does have its limitations, principal among which is that it does not enable us to calculate the intensity of the scattering for which constructive interference occurs.

For that we need to build on what we have already developed and write down the scattering amplitude of the crystal. To do so we note that a crystal structure may be specified in the following way. First, a lattice of points is defined in space, which must reflect the symmetry of the crystal, and then a choice of unit cell is made, in other words a choice is made over which atoms to associate with each lattice site. If $\mathbf{R_n}$ are the lattice vectors that define the lattice, and $\mathbf{r}_j$ the position of the atoms with respect to any one particular lattice site, then the position of any atom in the crystal is given by $\mathbf{R_n} + \mathbf{r}_j$. It follows that the scattering amplitude for the crystal factorizes into the product of two terms, which we write as

$$F^{\text{crystal}}(\mathbf{Q}) = \overbrace{\sum_j f_j(\mathbf{Q})\mathrm{e}^{i\,\mathbf{Q}\cdot\mathbf{r}_j}}^{\text{Unit cell structure factor}} \overbrace{\sum_{\mathrm{n}} \mathrm{e}^{i\,\mathbf{Q}\cdot\mathbf{R_n}}}^{\text{Lattice sum}} \tag{1.13}$$

where the first term is the *unit cell structure factor*, the second term is a sum over lattice sites, and where again we have neglected a leading factor of $-r_0$. In applications, such as solid state physics it is the structure of the material that is of interest in its own right. For many other applications, such as in molecular and protein crystallography, the lattice is of no interest whatsoever, and assembling the molecules on a lattice merely serves to amplify the signal.

All the terms in the lattice sum given in Eq. (1.13) are phase factors located on the unit circle in the complex plane. The sum will therefore be of order unity unless the scattering vector happens to fulfill

$$\mathbf{Q}\cdot\mathbf{R_n} = 2\pi \times \text{integer} \tag{1.14}$$

in which case it becomes of order N, the number of unit cells. The lattice vectors $\mathbf{R_n}$ are of the form

$$\mathbf{R_n} = n_1\,\mathbf{a}_1 \;+\; n_2\,\mathbf{a}_2 \;+\; n_3\,\mathbf{a}_3$$

where $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ are the basis vectors of the lattice and (n_1, n_2, n_3) are integers. A unique solution to Eq. (1.14) can be found by introducing the important concept of the reciprocal lattice. This new lattice

[5] See, however, Section 5.2 on quasicrystals.

is spanned by the *reciprocal lattice basis* vectors which are defined by

$$\mathbf{a}_1^* = 2\pi\frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}, \quad \mathbf{a}_2^* = 2\pi\frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}, \quad \mathbf{a}_3^* = 2\pi\frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}$$

so that any lattice site in the reciprocal lattice is given by

$$\mathbf{G} = h\,\mathbf{a}_1^* + k\,\mathbf{a}_2^* + l\,\mathbf{a}_3^*$$

where (h, k, l) are all integers. We can see that the product of a lattice vector in the reciprocal ($\mathbf{G}$) and direct ($\mathbf{R_n}$) spaces is

$$\mathbf{G} \cdot \mathbf{R_n} = 2\pi(hn_1 + kn_2 + ln_3) = 2\pi \times \text{ integer}$$

and hence the solution to Eq. (1.14) that we are seeking is to require that

$$\boxed{\mathbf{Q} = \mathbf{G}}$$

This proves that $F^{\text{crystal}}(\mathbf{Q})$ is non-vanishing if and only if $\mathbf{Q}$ coincides with a reciprocal lattice vector. This is the Laue condition for the observation of diffraction from a crystalline lattice which may be shown to be completely equivalent to Bragg's law (Chapter 5, page 155).

Scattering from a crystal is therefore confined to distinct points in reciprocal space. The intensity in each point is modulated by the absolute square of the unit cell structure factor. From a (large) set of intensities from a given crystal it is possible to deduce the positions of the atoms in the unit cell. These considerations may of course be generalized to crystals containing molecules. Indeed these methods have had an enormous impact on our knowledge of molecular structure. More than 95% of all molecular structures come from X-ray diffraction studies. Data sets from crystals of large molecules such as proteins or even viruses encompass tens of thousands of reflections and sophisticated methods have been developed to get from the measured intensities to the atomic positions in the molecule. In Chapter 5 these concepts will be further developed, and the principles behind these methods will be explained.

In this section it has been tacitly assumed that the interaction between the X-ray and crystal is weak, since we have not allowed for the possibility that the scattered beam may be scattered a second or third time before leaving the crystal. This assumption leads to considerable simplicity and is known as the *kinematical approximation*. In Chapter 6 it is explained how this assumption breaks down when dealing with macroscopic perfect crystals, where multiple scattering effects become important, and we are then in what is known as the *dynamical* scattering limit.

Compton scattering by a free electron

The alternative to the classical description used so far in this section, is to view the incident X-ray as a beam of photons. For simplicity assume that the electron is initially at rest and is free. In a collision energy will be transferred from the photon to the electron, with the result that the scattered photon has a lower energy than that of the incident one. This is the Compton effect. Historically this was of considerable importance as it could not be explained using classical concepts, and thus gave further support to the then emerging quantum theory. The energy loss of the photon is readily calculated by considering the conservation of energy and momentum during the collision. The collision process is sketched in Fig. 1.9, while the kinematics of the collision are worked through in the box on page 17.

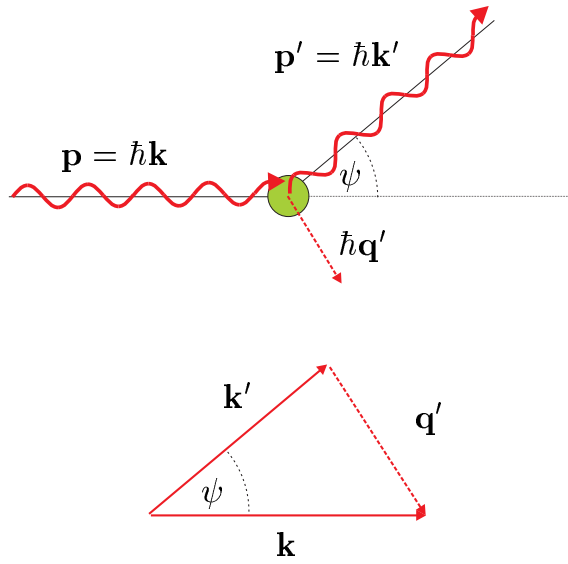


Fig. 1.9 Compton scattering. A photon with energy $\mathcal{E} = \hbar ck$ and momentum $\hbar\mathbf{k}$ scatters from an electron at rest with energy mc^2. The electron recoils with a momentum $\hbar\mathbf{q}' = \hbar(\mathbf{k} - \mathbf{k}')$ as indicated in the scattering triangle in the bottom half of the figure.

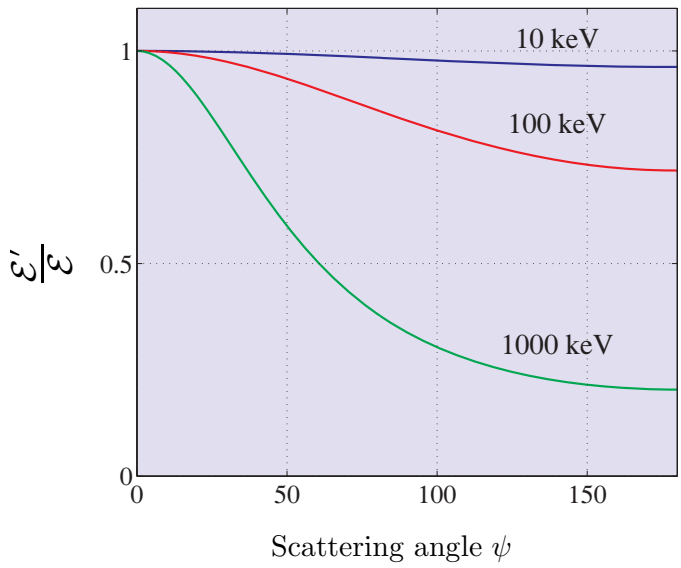


Fig. 1.10 The ratio of the energy $\mathcal{E}'$ of the scattered photon to the energy $\mathcal{E}$ of the incident one as function of scattering angle. The curves have been calculated from Eq. (1.15) with $\lambda\!\!\!^{-}_{C}k = \mathcal{E}/mc^2 = \mathcal{E}[\text{keV}]/511$.

Kinematics of Compton scattering

Conservation of energy for the scattering of a photon by an electron shown in Fig. 1.9 leads to

$$mc^2 + \hbar c\mathrm{k} = \sqrt{(mc^2)^2 + (\hbar c\mathrm{q}')^2} + \hbar c\mathrm{k}'$$

Dividing both sides by mc^2, and using the definition of the Compton wavelength, $\lambda\!\!\!^{-}_C = \hbar c/(mc^2)$, leads to

$$1 + \lambda\!\!\!^{-}_{\mathrm{C}}(\mathrm{k} - \mathrm{k}') = \sqrt{1^2 + \left(\lambda\!\!\!^{-}_{\mathrm{C}}\mathrm{q}'\right)^2}$$

This can be rewritten to obtain an expression for q'^2 by squaring both sides and collecting terms to give

$$\mathrm{q}'^2 = (\mathrm{k} - \mathrm{k}')^2 + 2\frac{(\mathrm{k} - \mathrm{k}')}{\lambda\!\!\!^{-}_{\mathrm{C}}}$$

Conservation of momentum (or equivalently wavevector) reads

$$\mathbf{q}' = \mathbf{k} - \mathbf{k}'$$

Taking the scalar product of $\mathbf{q}'$ with itself gives

$$\begin{aligned}\mathbf{q}' \cdot \mathbf{q}' = \mathrm{q}'^2 &= (\mathbf{k} - \mathbf{k}') \cdot (\mathbf{k} - \mathbf{k}') \\ &= \mathrm{k}^2 + \mathrm{k}'^2 - 2\mathrm{k}\mathrm{k}' \cos\psi\end{aligned}$$

Equating this with the expression for q'^2 derived from energy conservation yields

$$\mathrm{k}^2 + \mathrm{k}'^2 - 2\mathrm{k}\mathrm{k}' \cos\psi = \mathrm{k}^2 + \mathrm{k}'^2 - 2\mathrm{k}\mathrm{k}' + 2\frac{(\mathrm{k} - \mathrm{k}')}{\lambda\!\!\!^{-}_{\mathrm{C}}}$$

or

$$\mathrm{k}\mathrm{k}'(1 - \cos\psi) = \frac{(\mathrm{k} - \mathrm{k}')}{\lambda\!\!\!^{-}_{\mathrm{C}}}$$

which may be recast in the form

$$\frac{\mathrm{k}}{\mathrm{k}'} = 1 + \lambda\!\!\!^{-}_{\mathrm{C}}\mathrm{k}(1 - \cos\psi) = \frac{\mathcal{E}}{\mathcal{E}'} = \frac{\lambda'}{\lambda} \tag{1.15}$$

The result of the calculation is that the change in wavelength is proportional to the Compton scattering length defined by

$$\lambda\!\!\!^{-}_C = \frac{\hbar}{mc} = 3.86 \times 10^{-3}\ \text{Å} \tag{1.16}$$

There are thus two fundamental scattering lengths for the X-ray, the Thomson scattering length, r_0, and the Compton scattering length, $\lambda\!\!\!^{-}_C$. The ratio of these two is the fine structure constant

$$\alpha = \frac{r_0}{\lambda\!\!\!^{-}_C} \approx \frac{1}{137}$$

The ratio of the final to initial energy of the photon is given in Eq. (1.15) and is plotted in Fig. 1.10. For a given scattering angle, the scattering becomes progressively more inelastic as the energy $\mathcal{E}$ of the incident X-ray is increased. The energy scale is set by the rest mass energy of the electron, mc^2 = 511 keV.

One important difference between Thomson and Compton scattering is that the latter is *incoherent*. It has already been shown how X-rays that are elastically scattered from a crystal add up coherently when Bragg's law (or equivalently the Laue condition) is fulfilled. The scattering is then restricted to lie at points on the reciprocal lattice. The same is not true for Compton scattering, as it is the interaction between a single photon and electron, and the variation of the Compton cross-section[6] varies only slowly with scattering angle. As far as diffraction experiments are concerned, Compton scattering gives rise to a smoothly varying background which sometimes needs to be subtracted from the data.

Compton scattering may be used to obtain unique information on the electronic structure of materials. So far we have assumed that the electron in the Compton scattering process is initially at rest. This assumption breaks down for electrons in a solid, which instead have a finite momentum. When the kinematics are worked through for this case, it turns out that the Compton cross-section gives a measure of the electronic momentum distribution.

1.3 Absorption

Now let us turn to the absorption process. It is depicted in Fig. 1.11(a). An X-ray photon is absorbed by an atom, and the excess energy is transferred to an electron, which is expelled from the atom, leaving it ionized.

The process is known as *photoelectric absorption*. Quantitatively, the absorption is given by the linear absorption coefficient μ. By definition μdz is the attenuation of the beam through an infinitesimal sheet of thickness dz at a depth z from the surface (Fig. 1.12). The intensity $I(z)$ through the sample must therefore fulfill the condition

$$-dI = I(z)\,\mu\,dz \tag{1.17}$$

which leads to the differential equation

$$\frac{dI}{I(z)} = -\mu\,dz$$

[6] The calculation of the Compton cross-section is beyond the scope of this book. It is discussed by Lovesey and Collins [1996].

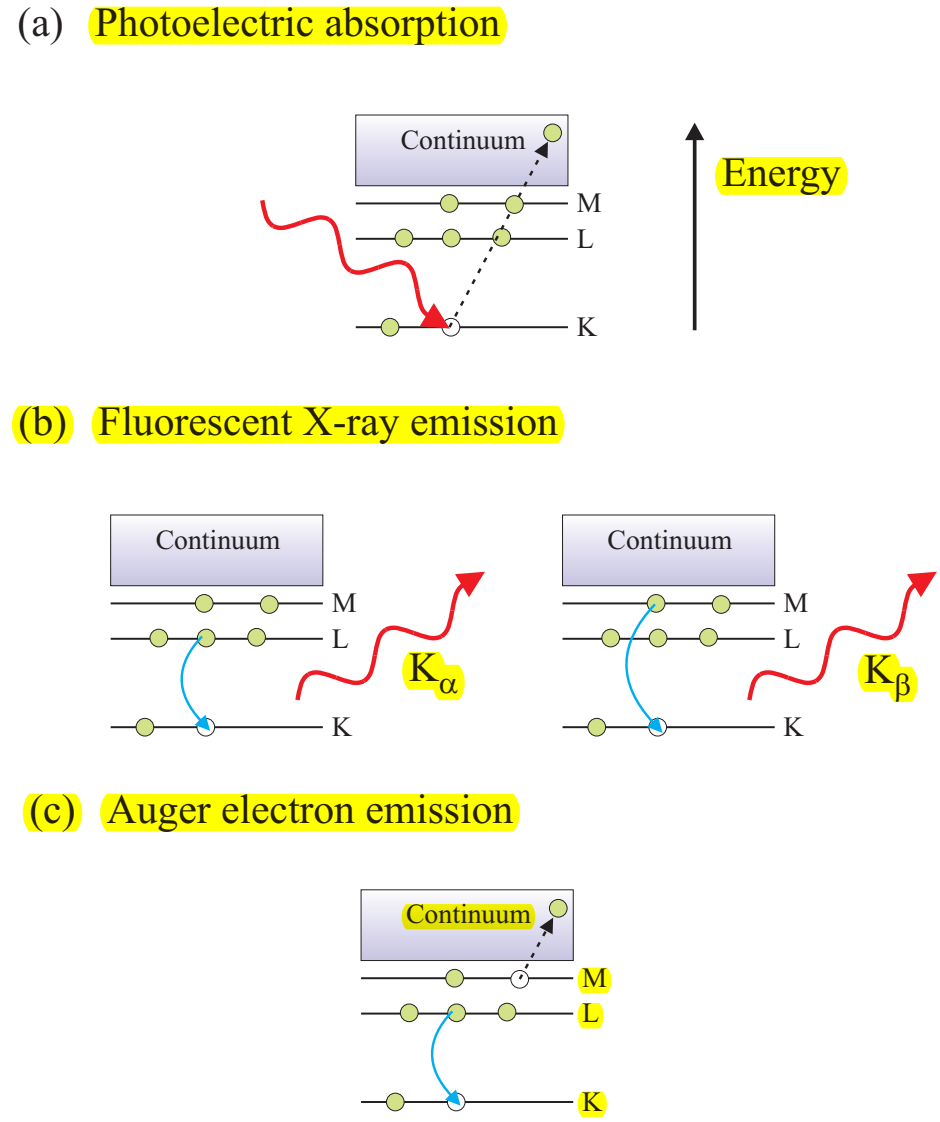


Fig. 1.11 Schematic energy level diagram of an atom. For clarity we have indicated only the energy of the three lowest shells; the rest are merged into the continuum. (a) The photoelectric absorption process. An X-ray photon is absorbed and an electron ejected from the atom. The hole created in the inner shell can be filled by one of two distinct processes: (b) Fluorescent X-ray emission. One of the electrons in an outer shell fills the hole, creating a photon. In this example the outer electron comes either from the L or M shell. In the former case the fluorescent radiation is referred to as the K_α line, and in the latter as K_β. (c) Auger electron emission. The atom may also relax to its ground state energy by liberating an electron.

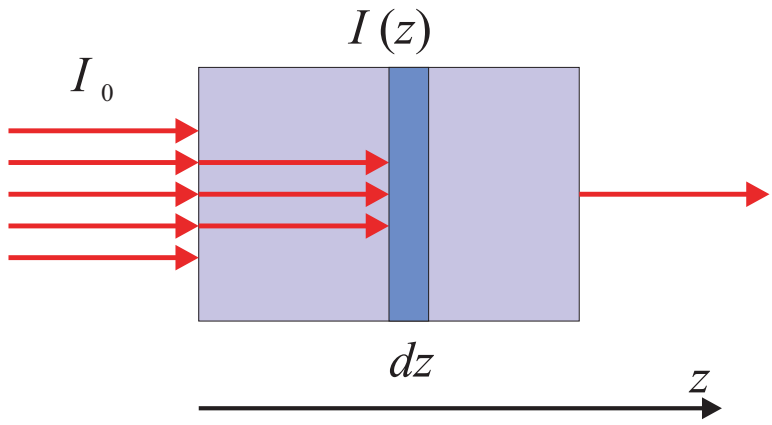


Fig. 1.12 The attenuation of an X-ray beam through a sample due to absorption. The attenuation follows an exponential decay with a characteristic linear attenuation length $1/\mu$, where μ is the absorption coefficient.

The solution is found by requiring that $I(z = 0) = I_0$, the incident beam intensity at $z = 0$, and we have

$$I(z) = I_0 \, \mathrm{e}^{-\mu z}$$

One can therefore readily determine μ experimentally as the ratio of beam intensities with and without the sample. The number of absorption events, W, in the thin sheet is proportional to I, and to the number of atoms per unit area, $\rho_{at}\, dz$, where ρ_{at} is the atomic number density. The proportionality factor is by definition the absorption cross-section, σ_a, so that

$$W = I(z)\, \rho_{at}\, dz\, \sigma_a = I(z)\, \mu\, dz$$

where in the last step we have used Eq. (1.17). The absorption coefficient is therefore related to σ_a by

$$\mu = \rho_{at} \sigma_a = \left(\frac{\rho_m N_\mathrm{A}}{M} \right) \sigma_a \tag{1.18}$$

where N_A, ρ_m and M are Avogadro's number, the mass density, and molar mass, respectively. In a composite material with several kinds of atoms, each with a number density $\rho_{at,j}$ and absorption cross section $\sigma_{a,j}$, the total probability for absorption in a layer dz is obtained by summing over $\rho_{at,j}\sigma_{a,j}dz$, the total probability of absorption for an atom of type j. Thus the absorption coefficient for a composite material is

$$\mu = \sum_j \rho_{at,j} \sigma_{a,j} \tag{1.19}$$

When an X-ray photon expels an electron from an inner atomic shell it creates a hole in that shell. In Fig. 1.11(a) we illustrate this for the case of an electron excited from a K shell. The hole is subsequently filled by an electron from an outer shell, L say, with the simultaneous emission of a photon with an energy equal to the difference in the binding energies of the K and L electrons (Fig. 1.11(b)). The emitted radiation is known as fluorescence. Alternatively, the energy released by an electron hopping from the L shell to the hole in the K shell can be used to expel yet another electron from one of the

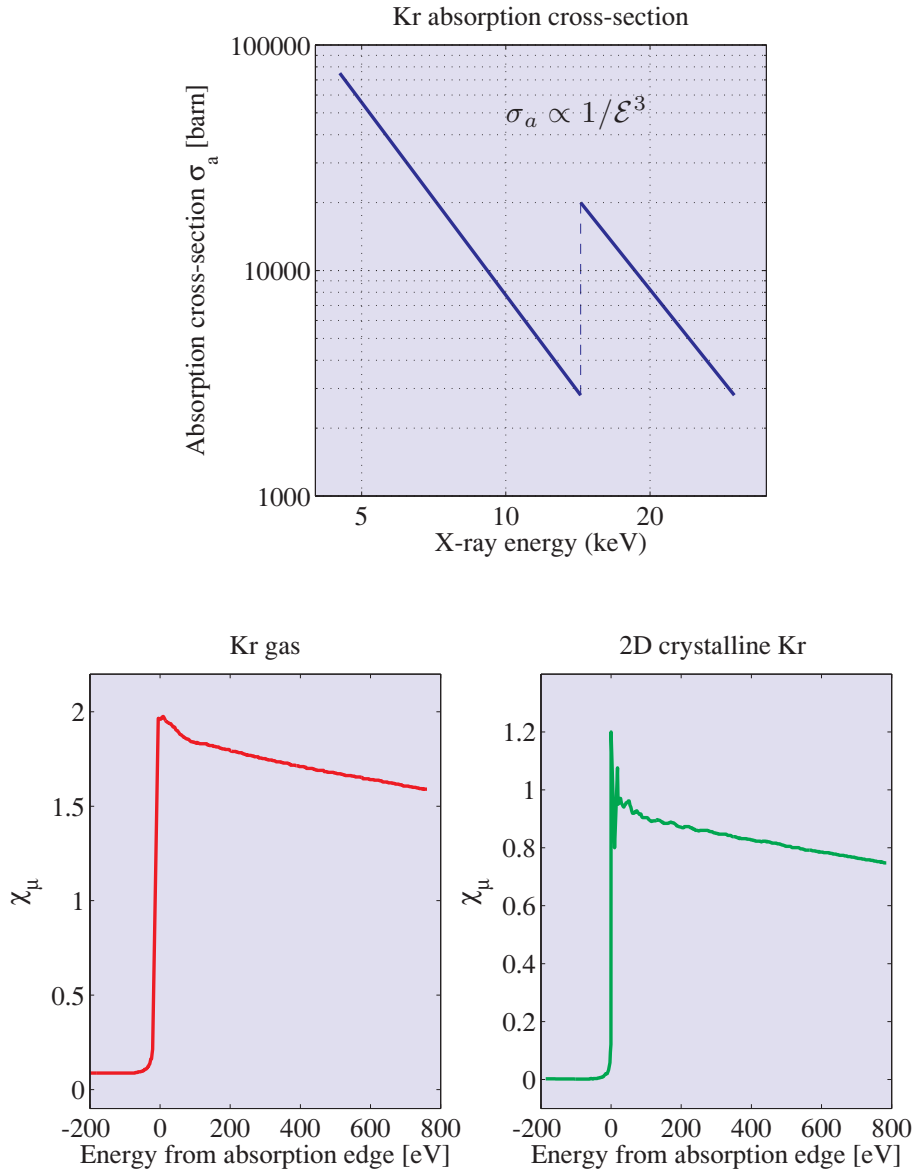


Fig. 1.13 Top: The absorption cross-section of gaseous krypton. Above a photon energy of 14.325 keV a K shell electron can be expelled from the atom and a new absorption 'channel' opens. The double logarithmic plot illustrates that the cross-section varies as $1/\mathcal{E}^3$. Bottom: A comparison of the absorption spectra of krypton in its gaseous form and physisorbed on graphite where the krypton atoms form a two-dimensional lattice. In the latter case fine structure, or wiggles, are evident which are known as EXAFS. The quantity χ_μ is proportional to the absorption cross-section σ_a.

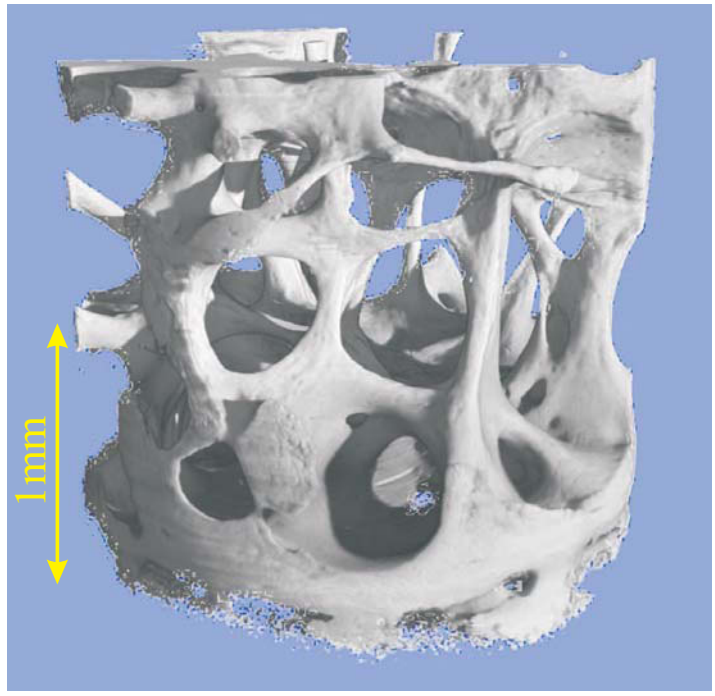


Fig. 1.14 Three-dimensional micro-CAT reconstruction of a cylindrical human vertebral bone specimen scanned with 3.6 μm spatial resolution. Note the difference between the cortical end-plate and the underlying trabecular bone. (Image courtesy of a collaboration between Aarhus University, Denmark, and HASYLAB at DESY, Germany.)

outer shells, as sketched in Fig. 1.11(c). This secondary emitted electron is called an Auger electron, named after the French physicist who first discovered the process.

The monochromatic nature of fluorescent X-rays is a unique fingerprint of the kind of atom that produces the fluorescence. It was Moseley who first discovered the empirical law

$$\mathcal{E}_{K_\alpha}\,[\mathrm{keV}] \approx 1.017 \times 10^{-2}\,(Z-1)^2 \tag{1.20}$$

where $\mathcal{E}_{K_\alpha}$ is the energy of the K_α line of a given element and Z is its atomic number[7]. The analysis of fluorescent radiation can be utilized for non-destructive chemical analysis of samples, and has the advantage that it is very sensitive. The radiation that creates the hole in the first place does not have to be an X-ray: it could also be from a beam of particles, such as of protons or electrons. For example, the latter is a standard option on electron microscopes, enabling the chemical composition of samples to be determined with a very fine spatial resolution.

The absorption cross-section has a distinct dependence on photon energy. An example is shown in the top panel of Fig. 1.13 for the rare gas krypton. Below a photon energy of 14.32 keV the X-ray photon

[7] Moseley's original work, published in 1913 against the backdrop of the emergence of quantum mechanics, played a key role in establishing the Bohr model of the atom. Moseley's law also allowed the position of elements to be understood in terms of Z, and was used to predict the existence of elements that up until that point had not been discovered.

can only expel electrons from the L and M shells. The cross-section is approximately proportional to $1/\mathcal{E}^3$. At a characteristic energy, the so-called K-edge energy, the X-ray photon has enough energy to also expel a K electron, with a concomitant discontinuous rise in the cross-section of about one decade. From then on the cross-section continues to fall off as $1/\mathcal{E}^3$.

If we examine the fine structure of the absorption just around the edge it is apparent that it depends on the structure of the material. This is again illustrated for Krypton in Fig. 1.13 [taken from Stern and Heald, 1983]. The wiggles in the spectrum from two-dimensional crystalline Krypton on graphite demonstrate the phenomenon of Extended X-ray Absorption Fine Structure (EXAFS) in condensed matter systems. We shall return to the interpretation of EXAFS data in Chapter 7.

The photoelectric absorption cross-section varies with the atomic number Z of the absorber, approximately as Z^4. It is this variation, and thus the contrast, between different elements that make X-rays so useful for imaging, as we describe in Chapter 9. Tissue is mainly water and hydrocarbons and thus has a 1/e thickness of many centimetres for hard X-rays, whereas bone contains a lot of Ca and a correspondingly smaller X-ray transmission. It was this, by now well-known, ability to look through the body that produced a sensation, when Wilhelm Conrad Röntgen discovered X-rays over a 100 years ago. When coupled with the computer power available today one can obtain the internal structure of parts of the body with remarkable precision. The technique is called CAT scanning, an acronym for Computer Axial Tomography (or Computer Aided Tomography). The idea is to take two-dimensional 'shadow' pictures from many angles, and then reconstruct the three-dimensional object using a computer program. An example of the type of exquisite images that can be obtained with modern CAT scanning is given in Fig. 1.14. Another tomography application where computer power is essential, utilizes subtraction of pictures taken above and below the K edge of the element one is particularly interested in. In this way the element-sensitivity is enhanced dramatically.

While photoelectric absorption arises from a physical process that is distinct from the scattering of a photon, it should always be bourne in mind that the two are nonetheless related (see Fig. 1.8). In Section 3.3 the relationship between the absorption cross section and the imaginary part of the scattering amplitude is established, while this interrelationship is more fully explored in Chapter 8.

1.4 Refraction and reflection

The interaction of X-ray photons with matter has so far been discussed mostly at the atomic level. However, since X-rays are electromagnetic waves, one should also expect some kind of refraction phenomena at interfaces between different media. To describe such refractive phenomena, the media of interest are taken to be homogeneous with sharp boundaries between them, each having its own refractive index n. By definition the refractive index of vacuum is one. It is well known that for visible light in glass n is large and can vary considerably, ranging from 1.5 to 1.8 depending on the type of glass. This of course enables lenses to be designed for focusing light and thereby obtaining magnified images. For X-rays the difference from unity of n is very small, and as we shall see in Chapter 3 is of order 10^{-5} or so. In general for X-rays, the refractive index can be expressed as

$$n = 1 - \delta + i\beta \tag{1.21}$$

where δ is of order 10^{-5} in solids and only around 10^{-8} in air. The imaginary part β is usually much smaller than δ. That the real part of n is less than unity is due to the fact that the X-ray spectrum generally lies to the high-frequency side of various resonances associated with the binding of electrons,

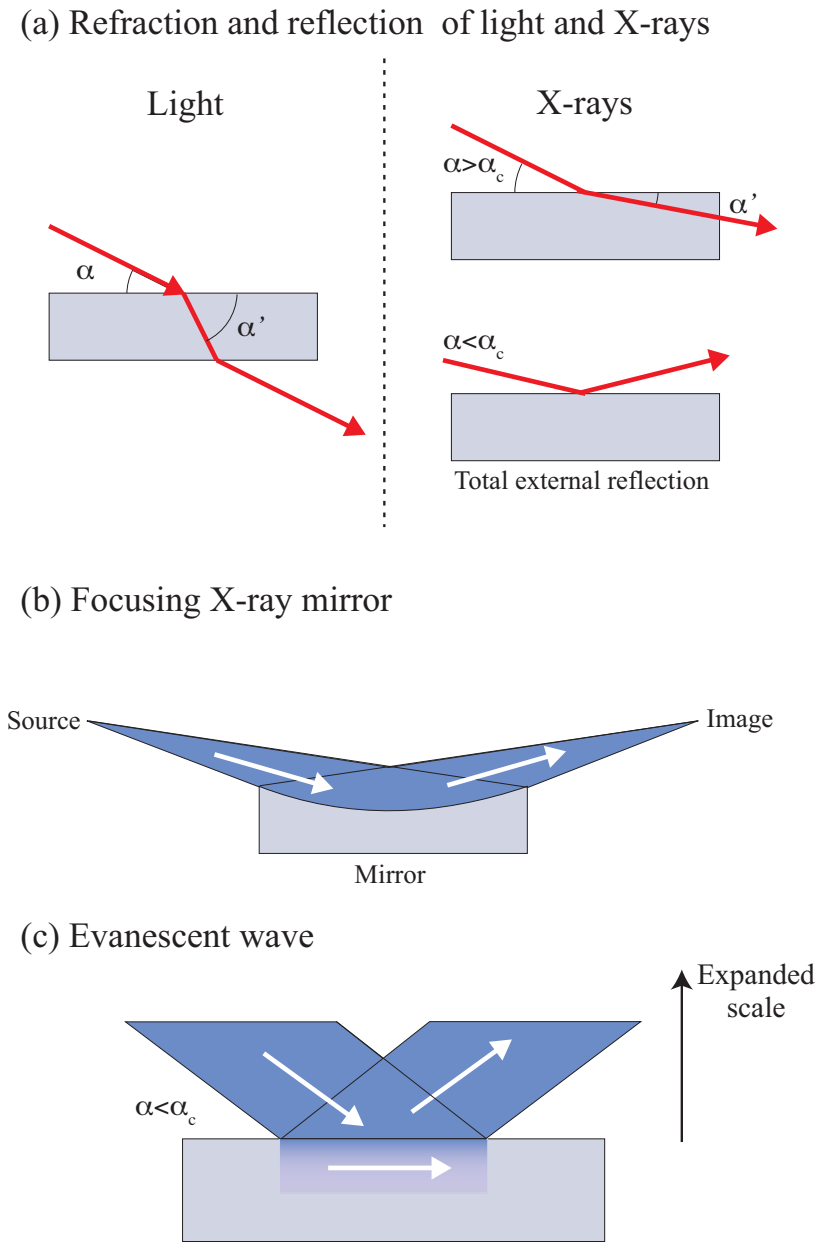


Fig. 1.15 (a) The refraction of light shows that in the visible part of the spectrum the refractive index of glass is considerably greater than one. In contrast, the index of refraction for X-rays is slightly less than one, implying total external reflection at glancing angles below the critical angle α_c. (b) A focusing X-ray mirror can be constructed by arranging that the incident angle is below the critical angle for total external reflection. (c) At glancing angles below the critical angle the reflectivity is almost 100%, and the X-ray only penetrates into the material as an evanescent wave with a typical penetration depth of ≈10 Å. In this way X-rays can be made to be surface sensitive.

as illustrated in Fig. 1.8. One consequence of the real part of n being less than unity is that it implies that the phase velocity inside the material, c/n, is larger than the velocity of light, c. This does not, however, violate the law of relativity, which requires that only signals carrying 'information' do not travel faster than c. Such signals move with the group velocity, not the phase velocity, and it can be shown that the group velocity is in fact less than c.

Snell's law relates the incident grazing angle α to the refracted grazing angle α' (see Fig. 1.15(a))

$$\cos\alpha = n\cos\alpha' \tag{1.22}$$

An index of refraction less than unity, implies that below a certain incident grazing angle called the critical angle, α_c, X-rays undergo total external reflection. Expansion of the cosine in Eq. (1.22) with $\alpha = \alpha_c$, $\alpha' = 0$ and using Eq. (1.21) allows us to relate δ to the critical angle α_c:

$$\alpha_c = \sqrt{2\delta}$$

where for simplicity we have taken $\beta = 0$. With δ being typically around 10^{-5}, α_c is of the order of a milli-radian. We shall see in Chapter 3, that the refractive constants δ and β can be derived from the scattering and absorption properties of the medium, respectively.

Total external reflection has several important implications for X-ray physics. First, total reflection from a curved surface enables focusing optics to be constructed as shown in the Fig. 1.15(b). A small source size is thus desirable, since from geometrical optics, a small source will be focused to a small image. A second consequence of total external reflection is that for $\alpha < \alpha_c$ there is a so-called evanescent wave within the refracting medium, see Fig. 1.15(c). It propagates parallel to the flat interface, and its amplitude decays rapidly in the material: typically with a penetration depth of only a few nanometers. This should be compared with a penetration depth of several micrometers at a glancing angle of several times α_c.

The much-reduced penetration of X-rays for angles less than α_c increases their surface sensitivity. This allows the scattering from the surface and near surface region to be studied, often in great detail, and indeed X-rays have become a valuable tool for the investigation of surfaces and interfaces.

1.5 Coherence

Throughout this introductory survey we have assumed that we are dealing with an X-ray beam in a perfect plane-wave state. This is obviously an idealization, and in this section we shall briefly discuss its limitation by recalling the concept of a *coherence length* of a real beam, and its relation to the source and monochromator. A real beam deviates from an ideal plane wave in two ways: it is not perfectly monochromatic, and it does not propagate in a perfectly well defined direction. Let us discuss these limitations in turn.

The top part of Fig. 1.16 shows two plane waves A and B with slightly different wavelengths, λ and $\lambda - \Delta\lambda$ say, but both propagating in exactly the same direction. The two waves are exactly in phase at the wavefront P. The question is how far do we have to go away from P before the two waves are out of phase? This defines the *longitudinal* coherence length L_L. If the two waves are out of phase after travelling L_L, then they will be in phase again after travelling $2L_L$. Let that distance be N wavelengths λ, or equivalently $(N+1)(\lambda - \Delta\lambda)$, i.e.

$$2L_L = N\lambda = (N+1)(\lambda - \Delta\lambda)$$

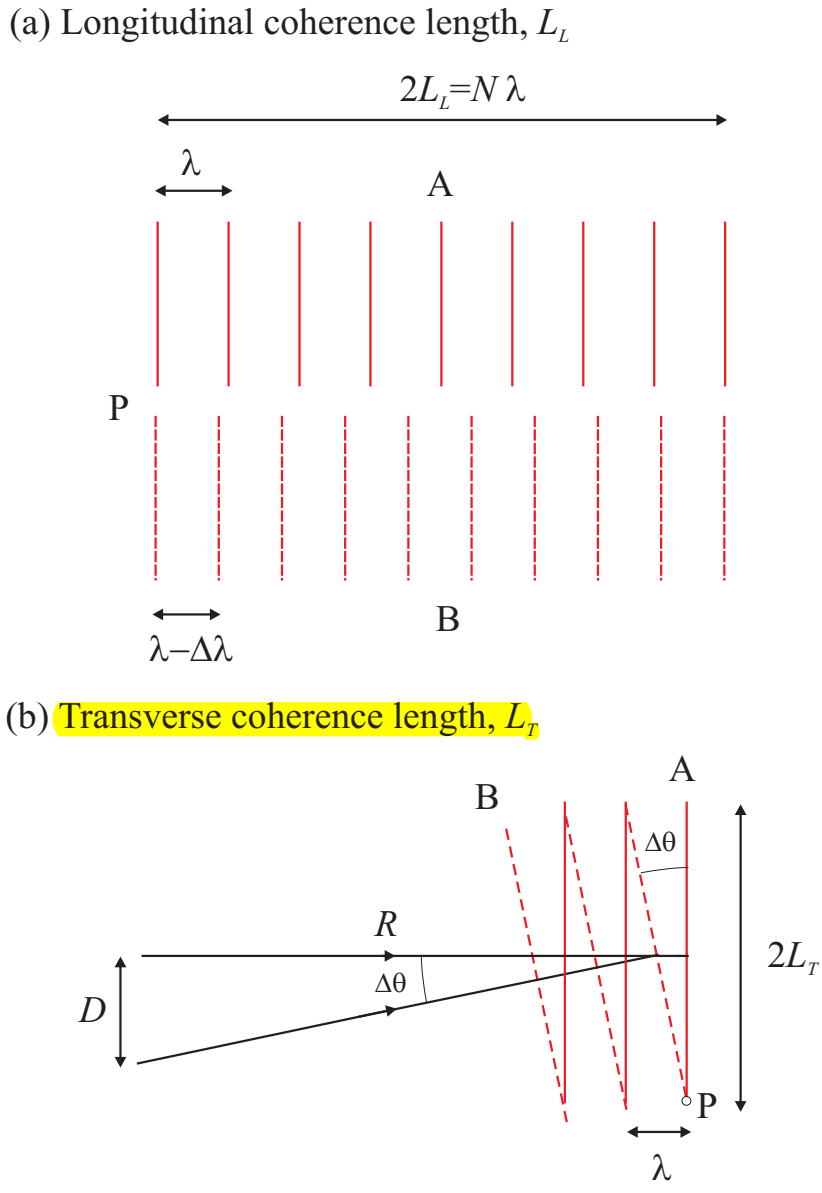


Fig. 1.16 Longitudinal and transverse coherence lengths. (a) Two plane waves with different wavelengths are emitted in the same direction. For clarity we have shown the waves displaced from each other in the vertical direction. After a distance L_L, the longitudinal coherence length, the two are out of phase by a factor of π. (b) Two waves with the same wavelength are emitted from the ends of a finite sized source of height D.

The second equation implies that $(N + 1)\Delta\lambda = \lambda$, or $N \approx \lambda/\Delta\lambda$, and using this result the first equation can be rearranged to read

$$L_L = \frac{1}{2}\frac{\lambda^2}{\Delta\lambda} \tag{1.23}$$

The bottom panel of Fig. 1.16 shows the other case: two waves A and B of the same wavelength, but with slightly different directions of propagation, say by an angle of $\Delta\theta$. Their wavefronts coincide at point P, and the question is now how far do we have to go from P along the wavefront of wave A before it is out of phase with wave B? By definition that distance is the *transverse* coherence length L_T. Clearly, if proceeding to a distance of $2L_T$, the two waves will be in phase again, and it is obvious from the figure that $2L_T\Delta\theta = \lambda$, i.e. $L_T{=}\lambda/(2\Delta\theta)$. Suppose that the different directions of propagation arise because the two waves originate from two different points on the source, let us say a distance D apart. If the distance from the observation point P to the source is R, then $\Delta\theta = D/R$ and we have

$$L_T = \frac{1}{2}\frac{\lambda}{(D/R)} = \frac{\lambda}{2}\left(\frac{R}{D}\right) \tag{1.24}$$

It is instructive to consider typical values for the coherence lengths, but to do so we need to make some assumptions about the source. At a third generation synchrotron the vertical source size is around 100 μm, and the experiment may be performed some 20 m away, so that for 1 Å X-rays L_T is approximately 10 μm in the vertical plane. To calculate the longitudinal coherence length we need to make some additional assumption about the device used to monochromate the beam. If a perfect crystal is used, $\Delta\lambda/\lambda \approx 10^{-5}$ (see Chapter 6), and then, according to Eq. (1.23), L_L is ≈ 5 μm for 1 Å X-rays, similar in order of magnitude to L_T. The consequence of a finite coherence length is that it places an upper limit on the separation of two objects if they are to give rise to interference effects. To take a simple example, consider the scattering from two electrons. If the projection of their separation on the wavevector transfer $\mathbf{Q}$ is much greater than the coherence length, then the total scattered intensity is the sum of scattered intensities from the individual electrons, and not the modulus squared of the sum of amplitudes as has been described thus far.

In Chapter 9 we describe how coherent beams of X-rays are utilized in modern imaging methods.

1.6 Magnetic interactions

The discussion so far has centred on the interaction between the electric field of the X-ray and the charge of the electron. What has been neglected is the magnetic field of the X-ray and the spin of the electron. When these are included in a full treatment of the interaction, terms emerge in the scattering cross-section that are sensitive to the spin and orbital magnetic moments of the electron. In this way it is possible to use X-rays to investigate magnetic structures. The history of X-ray magnetic scattering is much more recent than that of classical X-ray diffraction. In fact the first observation of X-ray magnetic scattering had to wait until 1972 and the pioneering experiments of de Bergevin and Brunel on antiferromagnetic NiO [de Bergevin and Brunel, 1972].

The reason for this is simply that magnetic scattering is much weaker than charge scattering. The

amplitude ratio of magnetic to charge scattering for a *single electron* is

$$\frac{A_{\text{magnetic}}}{A_{\text{charge}}} = \left(\frac{\hbar\omega}{mc^2}\right)$$

[Blume, 1985]. For 5.11 keV X-rays this ratio is 0.01, so the intensity of Bragg peaks that are purely magnetic in origin are weaker than the charge peaks by a factor of approximately 10^{-4}. In fact, as only relatively few atomic electrons contribute to the magnetic scattering (namely those with unpaired angular momenta in open shells), while all of them contribute to the charge scattering, the intensity ratio is typically depressed by an additional factor of 10^{-2} or so. Progress in the field of X-ray magnetic scattering was at first slow, but the routine availability of synchrotron radiation has given a tremendous boost to this subject, to the extent that it has now flourished into a field in its own right.

Sensitivity to magnetism is not restricted to scattering experiments, however, but also occurs in absorption processes. For example, the difference in absorption of left- and right-hand circularly polarized light by a solid (known generally as circular dichroism, or more specifically in the case of magnetic systems as X-ray magnetic circular dichroism (XMCD)) can be directly related to the ferromagnetic magnetization density, as described in Chapter 7. It has also been found that magnetic scattering itself is a much richer phenomena than early expectations, with the discovery that resonant magnetic scattering processes occur when the energy of the incident X-ray is tuned close to certain atomic absorption edges [Namikawa et al., 1985, Gibbs et al., 1988]. These subjects take us beyond the scope of this volume, but it is important to realize that the study of the interaction of X-rays with matter is still an active field of research some 100 years or so after the discovery of the X-ray [see, for example, Lovesey and Collins, 1996].

1.7 Further reading

Röntgen Centennial – X-rays in Natural and Life Sciences, Eds. A Haase, G. Landwehr, and E. Umbach (World Scientific, Singapore, 1997).

Fifty Years of X-ray Diffraction, P. P. Ewald, (International Union of Crystallographers, N. V. A. Oosthoek's uitgeversmaatchappj, Utrecht, 1962).

X-rays 100 Years Later, Physics Today (special issue) **48**, (1995).

2

Sources

2.1 Early history and the X-ray tube

Röntgen discovered X-rays in November 1895 in his laboratory at the University of Würzburg, Germany. He was examining the light and other radiation associated with the discharge from electrodes in an evacuated glass tube. He had covered the tube, a so-called Geisler discharge tube, so that no visible light could escape. The laboratory was also darkened. All that could be seen was a faint yellow-green light from a fluorescent screen placed close to the tube. The fluorescent light was flickering, since the high voltage was supplied by the *ac* output of an induction coil, and could be seen even when the screen was several metres away from the tube. To his amazement the radiation from the tube passed through paper and wood, whereas metal pieces of equipment cast a shadow on the screen. The most stunning phenomenon occurred when he placed his hand into the space between the tube and the screen and saw the bones inside. Röntgen was a keen amateur photographer and he quickly had the idea to photograph the X-ray beam instead of using the fluorescent screen. The photographs were convenient scientific documentation of his discovery, which was first published in the annals of the local Würzburg Scientific Society in late December of 1895. The paper is entitled 'Uber eine neue Art von Strahlen – vorläufige Mitteilung'[1]. The fact that one could now 'see' inside the human body was a sensation that spread worldwide within a few weeks, with implications for medical science that can hardly be overstated.

It became clear from Röntgen's subsequent investigations that the imaging of bones in the body is based on the fact that X-ray absorption is strongly dependent on the atomic number of the elements; it varies approximately as Z^4. The other important application of X-rays, based on diffraction phenomena, showing how crystalline matter is built up by atoms forming a periodic lattice, had to wait until 1912 when von Laue and his coworkers obtained the first diffraction pattern from a crystal of copper sulfate. In the following year W.H. Bragg and W.L. Bragg (father and son) examined the diffraction of X-rays from a number of crystals and laid the foundations of the field of crystallography, which subsequently allowed one to determine the structure of molecules.

The younger Bragg also found a particular simple way to interpret the diffraction patterns which

Elements of Modern X-ray Physics, Second Edition. Jens Als-Nielsen and Des McMorrow

[1] 'On a new kind of radiation – preliminary communication'

proved unambiguously that X-rays are nothing other than electromagnetic radiation of short wavelength: Röntgen had also played with the same idea, and tried to prove it experimentally, but without success. His influence in German physics at the time was so great that even von Laue and coworkers were not tempted to reach the same conclusion from their diffraction experiments as Bragg.

The standard X-ray tube and the rotating anode

The X-ray tube Röntgen used was a tricky business to run reliably. It was therefore a tremendous practical step forward when in 1912 W.D. Coolidge from General Electric Research Laboratories in New York developed a new tube, where electrons were produced by a glowing filament and subsequently accelerated towards a water-cooled metal anode (see Fig. 2.1). Now one could vary the high voltage and the current independently, and the limitation of intensity was set only by the cooling efficiency. It turns out that the maximal power for such a device is around 1 kW. The Coolidge tube served as the standard X-ray tube for many decades with only marginal technical improvements.

Although it was appreciated early on that by spinning the anode the heat could be dissipated over a much larger volume than in a standard tube, allowing the total power to be correspondingly increased, it was not until the 1960s that so-called rotating anode generators became available on a commercial basis. One of the technical difficulties to overcome had been the problem of how to make a high-vacuum seal on the rotating shaft, inside which the cooling water must flow in and out.

The spectrum of X-rays generated from electrons impinging on a metal anode has two distinct components. There is a continuous part due to the electrons being decelerated, and eventually stopped in the metal. This is consequently known as *bremsstrahlung* radiation (after the German *bremsen* for brake), and has a maximum energy that corresponds to the high voltage applied to the tube. Superimposed on this broad spectrum is a sharp line spectrum. In a collision with an atom the incident electron may also cause an atomic electron to be removed from one of the inner shells, creating a vacancy. The subsequent relaxation of an electron from an outer shell into the vacancy may produce an X-ray with a characteristic energy equal to the difference in energy between the two shells. This is the fluorescent radiation. For experiments requiring a monochromatic beam one often utilizes the K_α line which is several orders of magnitude more intense than the bremsstrahlung spectrum. However, only a very small fraction of the photons emitted into the solid angle of 2π can be utilised in a beam requiring an angular divergence of a few squared milli-radian. In addition, the line source is not continuously tuneable so the optimal wavelength for the experiment cannot be chosen, or scanned, at will. As we shall see in the following sections, X-rays generated from synchrotron sources do not have these drawbacks, and have a brilliance which is enormously higher than that of standard laboratory sources.

2.2 Introduction to synchrotron radiation

Synchrotron radiation takes its name from a specific type of particle accelerator. However, synchrotron radiation has become a generic term to describe radiation from charged particles travelling at relativistic speeds in applied magnetic fields which force them to travel along curved paths. Besides synchrotrons themselves, synchrotron radiation is produced in storage rings where electrons or positrons are kept circulating at constant energy. In a storage ring the synchrotron radiation is produced either in the bending magnets needed to keep the electrons in a closed orbit, or in insertion devices such as wigglers or undulators situated in the straight sections of the storage ring. In these devices an alternating magnetic field forces the electrons to follow oscillating paths rather than moving in a straight line. In a wiggler the amplitude of the oscillations is rather large, and the radiation from different wiggles