Philip Willmott

An Introduction to Synchrotron Radiation

Techniques and Applications

WILEY

An Introduction to Synchrotron Radiation

An Introduction to Synchrotron Radiation

Techniques and Applications

PHILIP WILLMOTT

Swiss Light Source
Paul Scherrer Institut
Villigen, Switzerland

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

Willmott, Philip (Philip R.)
An introduction to synchrotron radiation : techniques and applications / Philip Willmott.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-74579-3 (cloth) – ISBN 978-0-470-74578-6 (pbk.)
1. Synchrotron radiation. 2. X-ray optics. I. Title.
QC793.5.E627W55 2011
539.7′35 – dc22

2011007847

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

Cloth ISBN: 978-0-470-74579-3
Paper ISBN: 978-0-470-74578-6
ePDF ISBN: 978-1-119-97096-5
oBook ISBN: 978-1-119-97095-8
ePub: 978-1-119-97286-0
eMobi: 978-1-119-97287-7

Typeset in 10/12pt Times by Laserwords Private Limited, Chennai, India
Printed in Singapore by Fabulous Printers Pte Ltd

To the memory of my father

Contents

Preface

In the 1980s, when the first dedicated synchrotron facilities came online, the large majority of users were condensed-matter physicists and physical chemists. As synchrotron technology matured and experimental techniques diversified and became more refined and easier to implement, new types of users began to recognize the unparalleled capability of synchrotron radiation to investigate the properties of matter. Structural biologists were among the first nonphysicists to take advantage of synchrotron radiation and they remain today the most ubiquitous of synchrotron users. At the end of the first decade in the 21st century, synchrotron facilities are inhabited by a hugely diverse menagerie of users, including (among others) archaeologists, environmental scientists, polymer chemists, biologists, art restoration experts, organic chemists and astrophysicists. Physicists and materials scientists have witnessed in less than two decades what was originally more or less exclusively their own dominion annexed by the whole gamut of natural scientists!

I am a physicist. However, when I began working with synchrotron light in April 2001 as a newly employed beamline scientist at the Swiss Light Source, my grasp of x-ray science was limited to undergraduate-level x-ray crystallography. My expertise until then being in condensed matter physics and surface science, I immediately undertook to improve my understanding of x-ray physics (indeed, if I were to remain employed as a beamline scientist, I had no other choice!) and began reading from a selection of four or five textbooks and classic and seminal research articles, while also interrogating my long-suffering and patient work colleagues.

I also, somewhat rashly, offered an elective, one-semester course at the University of Zürich, meant to guide the interested non-expert in basics of synchrotron physics and applications. In many ways, the course mirrored my own learning curve in those early days. As I began to collate the contents of the course it rapidly became clear to me, even with my rudimentary background (or maybe because of it), that there was a gap in the literature – most monographs on synchrotron science were pitched for graduate-level physicists upwards, or (less commonly) aimed at an audience including medics, biologists and those not necessarily comfortable or familiar with undergraduate-level mathematics and physics.

The central goal of this text, therefore, is to explain the physics of synchrotron light and synchrotron techniques in a manner that on the one hand is accessible to the non-physicist, while on the other is of value to traditional users of synchrotron light. Hence, although mathematical formulas feature prominently in most of the chapters, they are in general kept simple enough that a physical picture can be readily drawn from them. This aspect – a physical, *visual* insight is – I believe crucial in conveying new concepts, especially those which are not intrinsically difficult, but may be complex and detailed. As a result, graphics figure prominently in this text.

Another important aspect, often missing from monographs in general, is the day-to-day *manifestation* of a concept (be it an experimental technique, a piece of instrumentation or a physical phenomenon) – too often, it happens that an idea is explained in an abstract manner, perhaps supplemented by a single figure showing an example, but translating this information into useful guidelines as to how an experiment should be carried out or how a piece of equipment should be handled is often all but impossible. I have therefore attempted here to include detailed worked examples as much as possible, in order to demonstrate to the reader how an idea is realized on a practical level.

It is unrealistic to include a description of all synchrotron techniques, with the myriad subtle variations that have evolved. Instead, I have concentrated on the most commonly used 'classical' methods, but I have also included modern techniques derived from these, in particular those which have only become realistic propositions either because of the higher brilliance now available at synchrotron facilities, or thanks to other novel developments in instrumentation.

It is therefore my hope that this intermediate level of sophistication will translate into an introductory monograph useful to a broad spectrum of scientists.

Philip Willmott, Zürich

Acknowledgements

I am greatly indebted to many colleagues for their patience and tolerance in the face of my repeated questioning (some might say badgering), and their critical review of the manuscript and suggestions for improvements. In particular, I would like to thank Michael Böge, Oliver Bunk, Antonio Cervellino, Roy Clarke, Julie Fife, Uwe Flechsig, Fabia Gozzo, Daniel Grolimund, Gerhard Ingold, Juraj Krempasky, Federica Marone, Andreas Menzel, Frithjof Nolting, Bruce Patterson, Luc Patthey, Phil Pattison, Stephan Pauli, Christoph Quitmann, Thorsten Schmitt, Urs Staub, Marco Stampanoni, Paul Tafforeau, Jeroen van Bokhoven, Friso van der Veen, the editorial team at Wiley and the many students who helped to test the level of the contents of this book.

1

Introduction

In mid-1981, Hartmut Michel, from the Max-Planck Institute in Martinsried, West Germany, succeeded in growing the first single crystals of a protein from the purple bacterium *Rhodopseudomonas viridis*, containing the reaction centre responsible for photosynthesis [1]. Two months later, one of these crystals was x-rayed using a laboratory-based source and a high-quality set of diffraction patterns was recorded (see Figure 1.1(a)). The exposure times were typically several hours. Despite the daunting size of the structure, known to have a molecular mass of approximately 145 000 atomic mass units (a.m.u., or 'Daltons'), Michel, in collaboration with Robert Huber and Johann Deisenhofer, was able to determine the structure of this protein over the following three years, using the x-ray diffraction data and base-sequence determination. This tour de force was described in a milestone article in *Nature* in December, 1985 [2]. Michel, Deisenhofer, and Huber would receive the Nobel prize in chemistry for this work in 1988.

The importance of this work cannot be over-emphasized. Without hyperbole, it can be stated that this protein underpins all life on Earth. It catalyses the conversion of light energy into chemical energy (producing the molecule adenosine triphosphate, ATP, the 'molecular currency' of intracellular energy transfer). The structure of this molecule, and hence the underlying photochemical mechanisms, could only be elucidated thanks to x-ray techniques.

The year in which Michel recorded these diffraction patterns, 1981, also saw the start of experiments at the Synchrotron Radiation Source (SRS), the first facility to be constructed explicitly for the production of synchrotron x-ray light, in Daresbury, England. In the intervening quarter of a century, the performance of these dedicated synchrotron facilities has improved greatly, resulting in an increase in brilliance of several orders of magnitude. Think how much more quickly (and accurately) the data by Michel could have been recorded using synchrotron radiation!

Indeed, we need look no further than recent examples in the same field of research to see how it has advanced. Between 2006 and 2008, the group of Nenad Ban at the Federal Institute of Technology (ETH) in Zürich published articles on the overall architecture and

An Introduction to Synchrotron Radiation: Techniques and Applications, First Edition. Philip Willmott.

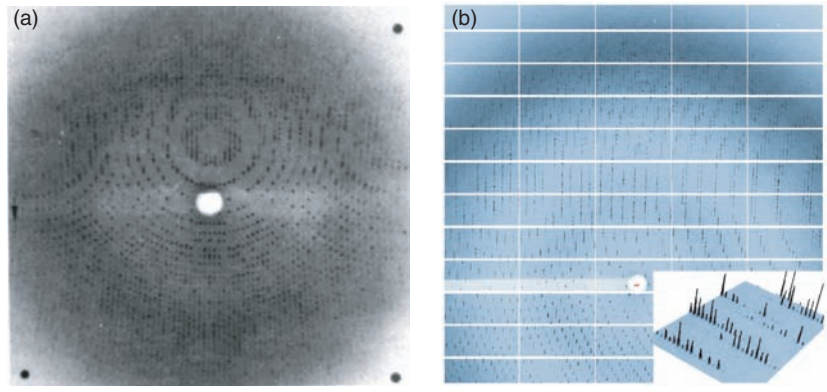

Figure 1.1 *Then and now: Recording x-ray diffraction data in protein crystallography. (a) X-ray diffraction pattern of a single crystal of photosynthetic reaction centre. The exposure time was 20 hours, using a laboratory-based x-ray source. Reprinted from [1] with permission of Elsevier. (b) A fine phi-slicing image of the diffraction pattern from a crystal of fungal fatty-acid synthase recorded using the Pilatus 6M detector. The inset shows a small section of the image rendered as three-dimensional intensity-spikes. The separation between adjacent spikes is approximately 0.03°. The entire data set was recorded in a few minutes. Courtesy Jasmin Lozza, ETH Zuerich.*

detailed structure of the enzymes responsible for fatty acid synthesis both in funghi and in mammals. Fatty acids are essential components in cells as energy-storage molecules, as chemical messengers, as well as being central building blocks of biological membranes.

Using high-brilliance synchrotron radiation, both enzyme types were mapped out to a resolution of approximately 3 Å (see Figure 1.1(b)). The larger, fungal, fatty-acid synthase (FAS) formed crystals made up from unit cells with a volume of approximately 40 million cubic Angstroms, containing a 2.6 MDa molecular assembly, nearly twenty times larger than the structure solved by Michel, Deisenhofer and Huber (see Figure 1.2) [3]. Note also that these diffraction data were recorded in a matter of minutes...

These examples beautifully illustrate that, since their discovery in late 1895 by Wilhelm Röntgen, x-rays have played a pivotal role in society, particularly in medicine, pharmacy, physics and chemistry. Whereas research using x-rays was originally the dominion of physicists, x-rays are now a ubiquitous tool for research in almost all branches of scientific endeavour, from determining the internal architecture of cells and other biological structures, to the chemical composition, fabrication techniques and provenance of archaeological artefacts, to insights into the hidden earlier artistic efforts of one of the foremost influential post-impressionist painters.

This broad range of applications of x-rays has in the last two decades expressed itself in the diverse disciplines served and the broad palette of techniques now available at synchrotron facilities, which represent one of the principal examples of multidisciplinary research. Today, there are more than 70 facilities worldwide in operation or under construction, providing services for approximately 100 000 users from virtually every discipline of the natural sciences.

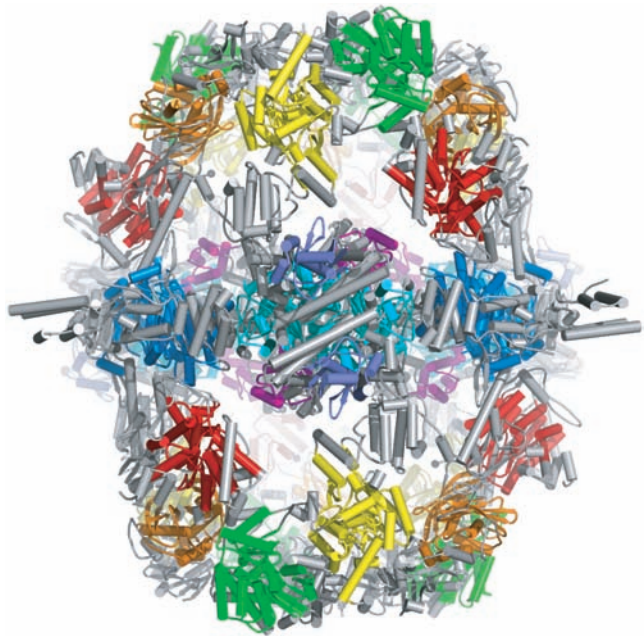

Figure 1.2 *The architecture of the fatty-acid synthase in funghi, solved using synchrotron-based x-ray diffraction. The individual enzymes, highlighted in colour, are embedded in the structural matrix, shown in grey. Courtesy Simon Jenni and Marc Leibundgut, ETH Zuerich.*

Why are x-rays (and ultraviolet radiation, which bridges the gap between visible light and x-rays) especially important for investigating the properties of materials? There are several reasons. Firstly, x-rays are able to probe deeply and nondestructively in solid materials, their degree of penetration depending on their energy and on the electron density (and hence the atomic weights and spatial distribution) of the constituent elements that make up the sample under investigation. In contrast, if the x-ray beam is made to impinge at very shallow angles to the sample surface, the beam will be totally reflected and only a so-called evanescent wave will penetrate the surface to a depth of a few nanometres. In this manner, x-rays can also be used as a highly surface-sensitive probe.

Secondly, x-ray photons with energies of several kiloelectronvolts (keV) have wavelengths comparable to those of typical atomic spacings in solid materials, measured in Angstroms ($1\,\text{Å} = 10^{-10}\,\text{m}$). Under certain conditions, crystalline arrays of atoms can therefore act as interference gratings (i.e. ordered arrays of scattering centres) for x-ray light, which can therefore be diffracted. On the other hand, electronic transitions involving valence electrons have energies that are in the range of ultraviolet (UV) photons, which are hence useful as probes for the chemical nature and electronic properties of materials.

Synchrotron storage rings are very powerful sources of x-rays. Synchrotrons research facilities are nowadays designed and dedicated to generate tunable beams of electromagnetic radiation from the far infrared to the hard x-ray regime, with intensities (defined by

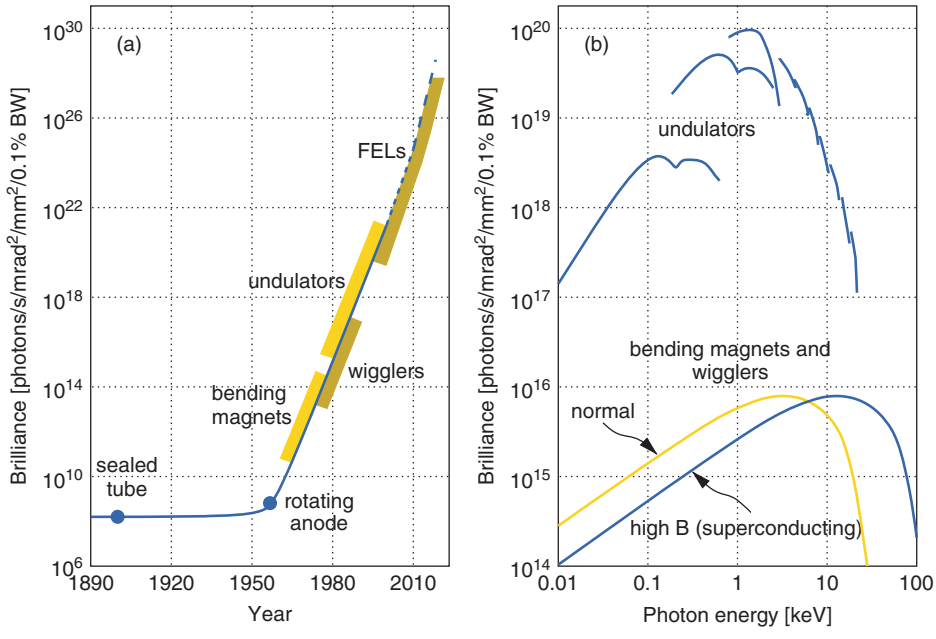

Figure 1.3 *The brilliance of x-ray sources. (a) A historical graph showing the enormous increase in brilliance starting in the second half of the twentieth century. (b) Typical spectral brilliance curves of devices commonly used in third-generation synchrotrons.*

their 'flux' or 'brilliance', explained precisely in Chapter 3), many orders of magnitude greater than those produced by laboratory-based sources (see Figure 1.3). Indeed, in the last four decades, there has been an increase in the brilliance available to researchers of a factor of 10^{14}. Up to another nine orders of magnitude of peak brilliance are promised with the development of x-ray free-electron lasers (XFELs).

Synchrotrons consist of an evacuated storage ring in which high-energy electrons circulate at highly relativistic velocities, and so-called beamlines that utilize synchrotron light emitted by the electrons tangentially to their orbital path at positions defined by components known as bending magnets and insertion devices (see Figure 1.4).

This text will cover the use of synchrotron radiation for the determination of materials structures and properties in physics, chemistry, biology and related disciplines such as archaeology and environmental science. In general, a biologist has a very different training in science compared, for example, to that of a physicist. I have therefore tried to pitch the contents of this text in such a manner that it should be accessible to all newcomers to synchrotron radiation. Certain chapters and sections will undoubtedly leave some readers cold – for example, it is unlikely that an archaeologist will see much relevance to his or her research in the arcane joys of surface x-ray diffraction. Accordingly, the tone of the different subjects and the depth to which they are explained is adapted as much as possible to the demands and scientific profile of the probable users.

After the Introduction, fundamentals of the interaction with high-energy electromagnetic radiation, especially x-rays, will be discussed in Chapter 2. Next, the basic physics

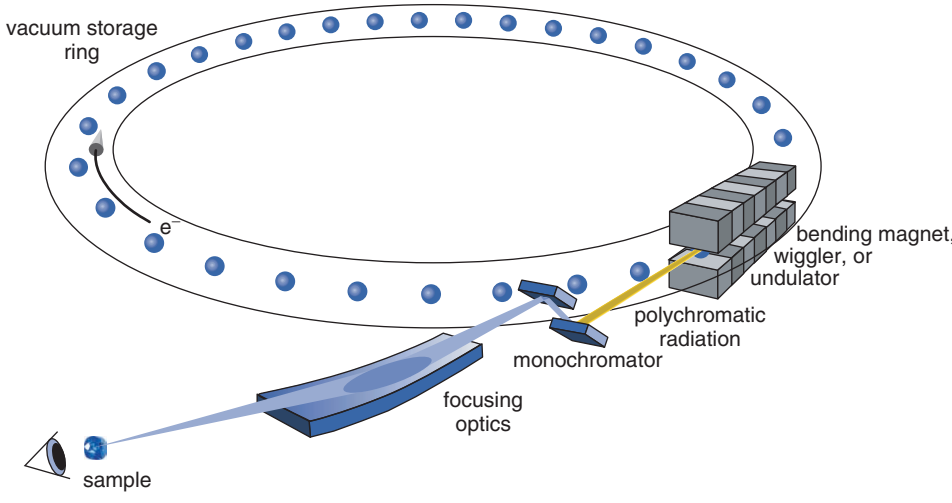

Figure 1.4 *Schematic of a third-generation synchrotron. Electrons moving at highly relativistic velocities in an evacuated storage ring emit electromagnetic (synchrotron) radiation as their direction is changed by bending magnets, used to steer them in a closed path, or by wigglers or undulators placed in straight sections of the storage ring, which 'shake' the electrons to and fro but keep an average straight trajectory. At the beamline, tangential to the storage ring, the radiation is normally (but not always) monochromated and focused using x-ray optics onto a sample.*

and the most important technical aspects of synchrotron facilities will be covered in Chapter 3, while technical details of the optics and instrumentation found at beamlines are summarized in Chapter 4. Although many users do not need to understand exactly how the x-rays they are using have been generated and manipulated, it is hoped that the information provided in these chapters will help guide them as to what any particular facility or beamline can realistically provide for potential experiments.

The second part of the book covers synchrotron techniques and applications. The methods described in Chapters 5 to 7 are, as far as possible, accompanied by representative and illustrative worked examples. The reader is encouraged to refer also to the original sources of these selected examples to obtain still deeper insights and become aware of subtleties peculiar to individual experiments – beamtime is expensive, and a user with an understanding of potential pitfalls and experimental obstacles is a valuable member of any team.

1.1 A Potted History of X-rays

On the evening of 8 November 1895, Wilhelm Röntgen first detected x-rays. He found that when running a high-voltage discharge tube enclosed in thick black cardboard, which excluded all visible light, in his darkened room, a paper plate covered on one side with barium platinocyanide would nevertheless fluoresce, even when it was as far as two metres from the discharge tube. Röntgen concluded that the discharge must be emitting a

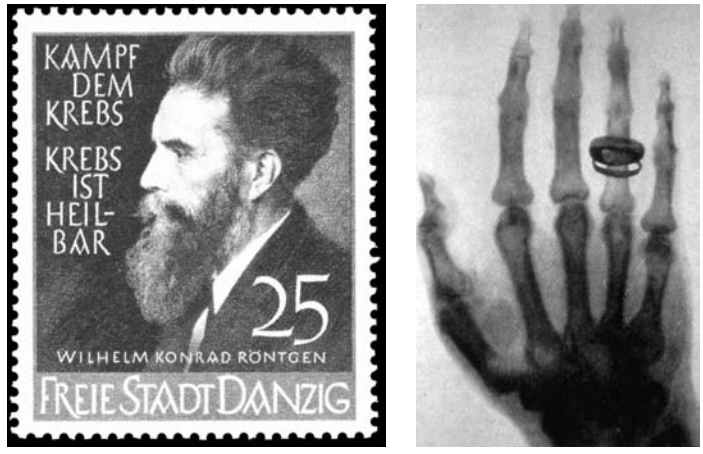

Figure 1.5 *Left: this postage stamp, depicting Wilhelm C. Röntgen, was first issued in April 1939. The message translates to 'Fight cancer. Cancer is curable.' – the use of x-rays in oncology was established soon after their discovery. Reprinted from http://commons.wikimedia.org/wiki/File:Danzig_Wilhelm_Konrad_R%C3%B6ntgen_25_Pf_1939.jpg. Right: one of the first radiographic images, recorded by Röntgen. Reprinted from Lino Piotto, Toger Gent and Giovanni Bibbo, Myth busting-in the world of x-rays, The Radiographer, 54, Copyright 2007, with permission of Lino Piotto.*

new form of invisible radiation that could escape the confines of the glass discharge tube and the surrounding covering – their unknown nature led him to name them x-rays [4].

Röntgen immediately discovered that these x-rays also stained photographic plates. Using these, subsequent experiments demonstrated that objects of different thicknesses placed in the path of these rays showed differing degrees of transparency to them. When Röntgen placed the hand of his colleague at Würzburg, the Swiss anatomist Albert von Kölliker, in the path of the rays over a photographic plate, he observed after development of the plate an image showing the shadows thrown by the finger bones and a ring, surrounded by the penumbra of the flesh, which absorbed less of the rays and therefore threw a fainter shadow (Figure 1.5).[1] In further experiments, Röntgen showed that the x-rays were produced by the impact of high-energy electrons on a material object. For his discovery of x-rays, Röntgen would receive the first ever Nobel prize in physics in 1901.

In 1909, characteristic x-ray radiation was discovered by the British physicists Barkla and Sadler (see Figure 1.6) – Barkla would also receive the Nobel prize in 1917. The term 'characteristic' refers to the narrow, intense lines found when a material emits x-radiation – each element has a unique set of lines, which, like a fingerprint, characterize that element alone, and is a manifestation of the discrete energy levels of electronic states in atoms.

In 1912, the year he became Professor of Physics at Zürich University, Max von Laue conjectured that periodic, crystalline, structures might diffract waves if their wavelength

[1] There seems to be some dispute as to whether the first x-ray photograph was of von Kölliker's or Röntgen's wife's hand. The image shown in Figure 1.5 was recorded at a demonstration at the Physical Institute on 23 January 1896 and is Kölliker's hand, while another, very similar image is undated and was only later attributed to Röntgen's wife.

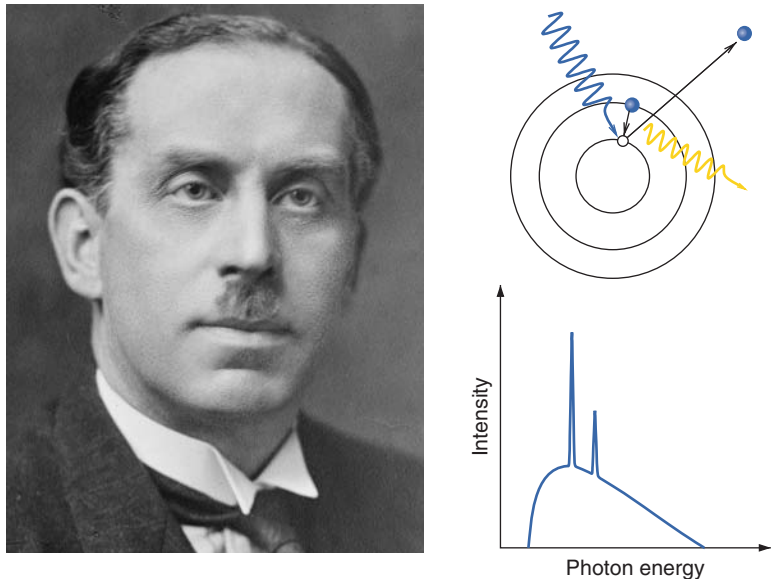

Figure 1.6 *Charles G. Barkla, and the characteristic x-radiation he discovered.*

were of the same order of magnitude as the crystalline periodicity and that x-rays might just fit the bill. He discussed these ideas with his colleagues Arnold Sommerfeld and Wilhelm Wien during a skiing expedition. Despite objections raised to the idea (exactly what these were seems now unclear), Walter Friedrich, one of Sommerfeld's assistants, and Paul Knipping (a diploma student of Röntgen's) succeeded in recording the first diffraction pattern of a crystal of copper sulfate (Figure 1.7). Von Laue was awarded the Nobel prize for physics in 1914 for this discovery.

This was indeed a hectic and exceedingly productive period. In 1913, the year between the recording of the first diffraction pattern and von Laue receiving his Nobel prize, the father-and-son team of Sir William Henry Bragg and William Lawrence Bragg were working on the theory and experimental techniques required to obtain detailed atomic structural information about crystals from their diffraction spot intensities and positions (Figure 1.8). The importance of their work is reflected by the fact that they would be awarded the Nobel prize for physics only two years later in 1915 'for their services in the analysis of crystal structure by means of x-rays'. Among other things, they formulated the famous Bragg law, which describes the mechanism by which x-ray diffraction occurs. Crucially, the son, W. L. Bragg, also explained why diffracted spots had the intensities they did, and was able, using these principles, to determine the internal atomic structure of some simple minerals. The foundations of x-ray crystallography had been laid down [5].

This culminated some 40 years later in one of the most important scientific discoveries of the twentieth century, namely the determination in 1953 of the structure of deoxyribose nucleic acid (DNA) and the mechanism by which it replicates [6], for which James Watson, Francis Crick and Maurice Wilkins would be awarded the Nobel Prize in medicine in 1962. Their discovery, which revolutionized biology and provided an

Figure 1.7 *Max von Laue, and the first x-ray diffraction pattern of copper sulfate.*

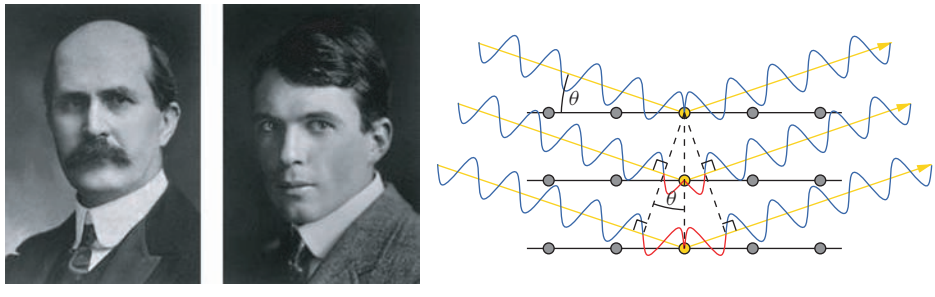

Figure 1.8 *W. H. Bragg and his son W. L. Bragg and the diffraction of x-rays by crystals.*

atomistic view of the stuff of life, was in no small measure assisted by x-ray diffractograms recorded by Rosalind Franklin at University College, London [7]. The detailed information which was gleaned by these patterns (in particular the central 'X' motif, so characteristic of diffraction from double-helical structures) was only made possible by the high-quality of both the crystals of salts of DNA prepared by Franklin, and of the diffractometer that was available (see Figure 1.9).

Although the diffraction patterns recorded by Franklin provided one of the last and most important clues for Watson and Crick in their model-building approach to the structural solution of DNA, the molecular details were primarily deduced using chemical and geometrical arguments. In the same year that Watson, Crick and Wilkins were awarded the Nobel prize in medicine, the prize in chemistry was won by their Cavendish Laboratory colleagues, John Kendrew and Max Perutz, for their determination by x-ray

Figure 1.9 *Watson and Crick and their model of DNA. The determination of its structure was much facilitated by the beautiful x-ray diffraction images of crystals of DNA salts recorded by Rosalind Franklin. Middle image reprinted with permission of Image State.*

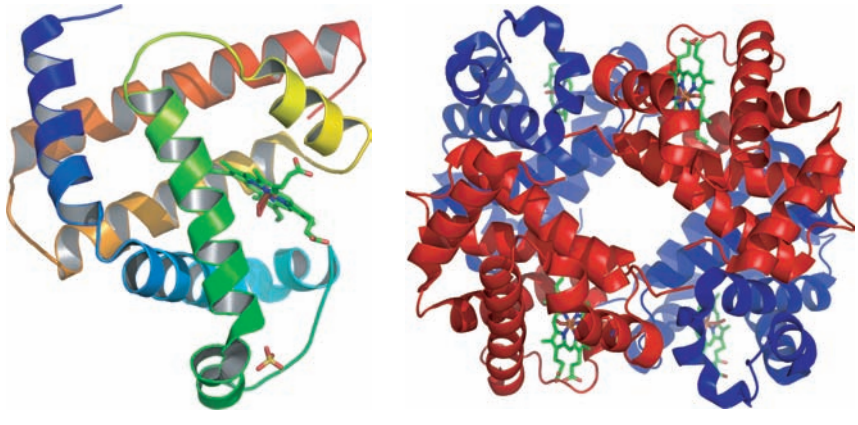

Figure 1.10 *The structures of myoglobin (left) and haemoglobin (right), determined by Kendrew and Perutz using x-ray diffraction. Right image reprinted from Wikipedia, Copyright 2007 Free Software Foundation, Inc.*

diffraction of the atomic structures of myoglobin (Kendrew [8]) and haemoglobin (Perutz [9]) (Figure 1.10). It was Perutz's invention of the technique of multiple isomorphous replacement in x-ray crystallography that provided the breakthrough in what had until then seemed the insurmountable problem of solving the structure of a crystal containing a molecule of a mass of approximately 17 000 Da.

Most recently, the Nobel prizes in chemistry in 2003, 2006 and 2009 have all been awarded to biochemistry studies exploiting x-ray crystallography[2] – McKinnon and Agre (2003) described the three-dimensional structure of potassium ion channels across cell membranes [11]; Kornberg was awarded the prize in 2006 for his studies of the process

[2] Incidentally, such tour de forces of structural biology only became possible with technological advances, not least in x-ray detectors. The 2009 Nobel prize in physics was partly awarded to Willard S. Boyle and George E. Smith for their pioneering work on charged-couple device (CCD) imaging technology [10]. CCDs are nowadays ubiquitous as x-ray detectors.

Table 1.1 *Nobel prizes awarded in the field of x-ray research.*

Year	Recipient(s)	Research discipline
1901	W. C. Röntgen	Physics; discovery of x-rays
1914	M. von Laue	Physics; x-ray diffraction from crystals
1915	W. H. Bragg and W. L. Bragg	Physics; crystal structure derived from x-ray diffraction
1917	C. G. Barkla	Physics; characteristic radiation of elements
1924	K. M. G. Siegbahn	Physics; x-ray spectroscopy
1927	A. H. Compton	Physics; for scattering of x-rays by electrons
1936	P. Debye	Chemistry; diffraction of x-rays and electrons in gases
1962	M. Perutz and J. Kendrew	Chemistry; structures of myoglobin and haemoglobin
1962	J. Watson, M. Wilkins, and F. Crick	Medicine; structure of DNA
1979	A. McLeod Cormack and G. Newbold Hounsfield	Medicine; computed axial tomography
1981	K. M. Siegbahn	Physics; high-resolution electron spectroscopy
1985	H. Hauptman and J. Karle	Chemistry; direct methods to determine x-ray structures
1988	J. Deisenhofer, R. Huber, and H. Michel	Chemistry; determining the structure of proteins crucial to photosynthesis
2003	R. MacKinnon and P. Agre	Chemistry; structure and operation of ion channels
2006	R. D. Kornberg	Chemistry; atomic description of DNA transcription
2009	V. Ramakrishnan, T. A. Steitz, and A. E. Yonath	Chemistry; structure and function of the ribosome

by which genetic information from DNA is copied to RNA (this process is common to all life forms containing cell nuclei, that is, all life except the prokaryotae) [12], while Ramakrishnan, Steitz and Yonath were able to solve the structure of ribosomes and thereby identify their core function in translating DNA code into living matter [13].

It can be argued, judging by the disproportionate number of Nobel prizes given in the field (see Table 1.1), that x-ray physics has been the one of the most influential areas of pure science in the twentieth century, while its impact has reached many other disciplines. Its importance in physics, chemistry, biology and medicine can hardly be overstated. It seems certain that x-rays will continue to play a hugely important role in modern science and technology in the twenty-first century, as ever more materials, compounds and structures in all their varied guises are discovered.

1.2 Synchrotron Sources Over the Last 50 Years

Short-wavelength synchrotron radiation generated by relativistic electrons in man-made circular accelerators dates back to shortly after the Second World War. However, the

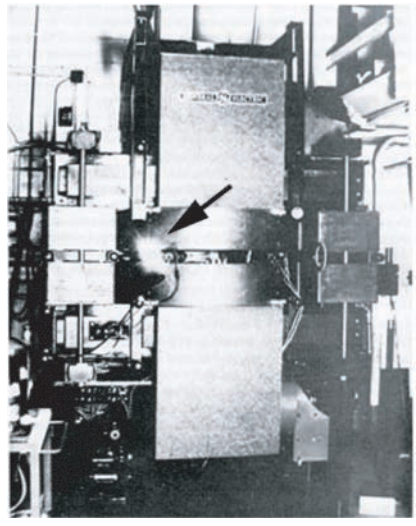

Figure 1.11 *Synchrotron light from the 70 MeV electron synchrotron at General Electric.*

theoretical basis for synchrotron radiation traces back to the time of Thomson's discovery of the electron in 1897. In the same year, Joseph Larmor derived an expression from classical electrodynamics for the instantaneous total power radiated by an accelerated charged particle, and the following year, the French physicist Alfred Liénard showed the radiated power emitted by electrons moving on a circular path to be proportional to $(\mathcal{E}/mc^2)^4/R^2$, where $\mathcal{E}$ is the electrons' kinetic energy, m is the electron rest mass and R is the radius of the trajectory.

The first observation of synchrotron radiation came on 24 April 1947, at the General Electric Research Laboratory in Schenectady, New York, (see Figure 1.11), although it was not being sought and was not immediately recognized for what it was. Here is an excerpt from Herb C. Pollock's description of this historical event over two decades later in a letter to Dmitri Ivanenko, dated 25 September 1970:

> On April 24, Langmuir and I were running the machine and as usual were trying to push the electron gun and its associated pulse transformer to the limit. Some intermittent sparking had occurred and we asked the technician to observe with a mirror around the protective concrete wall. He immediately signaled to turn off the synchrotron as 'he saw an arc in the tube.' The vacuum was still excellent, so Langmuir and I came to the end of the wall and observed. At first we thought it might be due to Cerenkov radiation, but it soon became clearer that we were seeing Ivanenko and Pomeranchuk radiation.

A more formal description of their findings can be found in [14]. 'Ivanenko and Pomeranchuk radiation' alludes to the two Russian physicists who, in 1944, published their calculations predicting the rate of energy loss due to radiating electrons [15].

Initially, synchrotron radiation was seen as an unwanted but unavoidable loss of energy in accelerators designed (ironically) to produce intense beams of x-rays by directing accelerated electrons onto a suitable target. The potential advantages of synchrotron

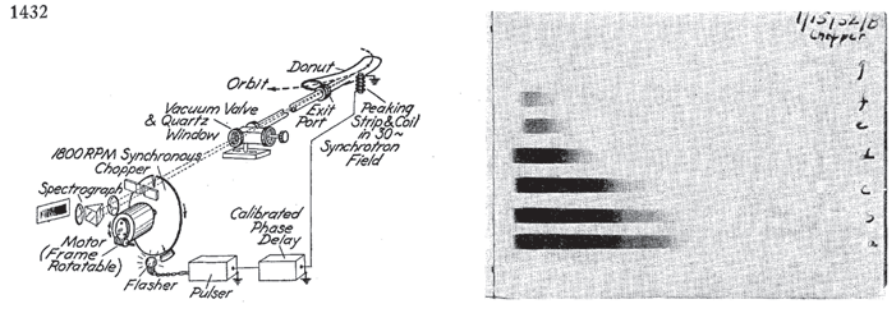

FIG. 11. Diagram showing the arrangement to be used for recording the radiation from essentially monoenergetic electrons. For work in the vacuum ultraviolet the rotating disk is enclosed, the drive shaft coming through a vacuum seal. So far the disk has not been used in the investigation of the far-ultraviolet spectrum. However, the plate of Fig. 12 was obtained with this arrangement and the quartz optical system indicated here.

FIG. 12. Reproduction of a plate obtained with the arrangement of Fig. 11, showing spectra of the continuous radiation emitted by essentially monoenergetic electrons. The various exposures correspond to electron energies ranging from 60 Mev at the top to 110 Mev at the bottom. An exposure at 50 Mev is not visible in the reproduction. The exposures were adjusted so that, in each case, approximately the same total number of radiating electrons was involved.

Figure 1.12 *Reproduction of Figures 11 and 12 in the* Physical Review *paper by Tomboulian and Hartmann [16], showing the experimental setup and photographic plate of the spectra produced by synchrotron radiation from monoenergetic electrons. Reprinted from [16] with permission of the American Physical Society.*

radiation for its own end were detailed by Diran Tomboulian and Paul Hartman from Cornell University in their seminal paper 'Spectral and Angular Distribution of Ultraviolet Radiation from the 300 MeV Cornell Synchrotron' in the *Physical Review* [16], describing their findings after being granted two-week's access there to investigate the possible applications of synchrotron light. The experimental setup and photographic-plate results reported in this paper, showing the broadband synchrotron radiation emitted by monoenergetic electrons, are reproduced in Figure 1.12.

Five years later, in 1961, a pilot experimental programme exploiting synchrotron radiation began when the National Bureau of Standards modified its 180-MeV electron synchrotron in Washington, D. C. to allow access to the radiation via a tangent section into the machine's vacuum system. Thus was born the Synchrotron Ultraviolet Radiation Facility (SURF I), the first facility catering for regular users of synchrotron radiation.

The first generation of synchrotron radiation sources were sometimes referred to as parasitic facilities (reflecting, perhaps, the perception of the primary users of these facilities towards these interlopers), as the synchrotrons were primarily designed for high-energy or nuclear physics experiments. In addition to SURF I, facilities in Frascati and in Japan soon after began to attract a regular stream of physicists keen to explore the possibilities of synchrotron radiation. More would soon follow suit. Because most of these early facilities had storage-ring energies around or below 1 GeV, experiments were concentrated in the ultraviolet and soft x-ray regimes. 1964 saw the first users of synchrotron radiation from the 6 GeV Deutsches Elektronen-Synchrotron (DESY) in Hamburg – suddenly the range of synchrotron radiation was extended to the hard x-ray region down to 0.1 Å (about 125 keV). Many further facilities began to accommodate synchrotron-radiation users. One of the first of these was the 240 MeV machine 'Tantalus' in Wisconsin. Although not originally designed to provide synchrotron radiation, it became the first

facility to be exclusively used for synchrotron experiments, providing radiation in the ultraviolet up to a few tens of eV.

After the development of efficient electron storage rings for long-term operation, the time was ripe to develop the first dedicated facilities designed specifically for synchrotron radiation. The 2 GeV Synchrotron Radiation Source (SRS) at Daresbury, England, was the first of these so-called 'second-generation' synchrotron sources. Experiments began in 1981. Several more new facilities were soon built and commissioned, while some first-generation sources were upgraded to second-generation status.

At approximately the same time it was becoming clear that increased x-ray beam brilliance could be achieved by optimizing the property of the electron beam in the storage ring called the 'emittance', described in detail in Chapter 3, by careful design of the array of magnets (the 'magnet lattice') used to manipulate the electrons, and by employing so-called 'insertion devices'. These include 'wigglers' and 'undulators' and are placed in straight sections in between the curved arcs of large storage rings. They operate by perturbing the path of the electrons in an oscillatory manner, so that although their *average* direction remains unchanged, synchrotron radiation is produced. Details of undulators and wigglers are covered in Chapter 3.

The inclusion of insertion devices in storage rings defines the third generation of synchrotron sources, which are designed for optimum brilliance, that is, the amount of power per unit frequency, surface area and solid angle. The first third-generation facility to be completed was the European Synchrotron Radiation Facility (ESRF, 6 GeV storage ring) in Grenoble, France, which began experiments in 1994. At the time of writing, the ESRF is undergoing a facility-wide upgrade which promises an enhancement of performance by approximately an order of magnitude.

The most recent third-generation facilities are the National Synchrotron Light Source II (NSLS II) at the Brookhaven National Laboratories in Upton, New York and MAX IV in Lund, Sweden. Construction of NSLS II began in 2009, with MAX IV about one year behind. The projected specifications for these medium-energy storage rings will redefine the benchmark for synchrotrons – it is planned to provide spatial and energy resolutions of the order of 1 nm and 0.1 meV, respectively, while the brilliance and emittance, at 10^{21} photons/mm^2/mrad^2/0.1% bandwidth and 0.5 nm rad, improve on previous performances by an order of magnitude. These significant advances are largely made possible by recent advances in the design of storage-ring magnets, but also (in the case of NSLS II) by the fact that although the storage-ring energy is, at 3 GeV, fairly moderate, its circumference of 800 m is more typical for high-energy facilities.

Beyond this, a fourth generation of synchrotron radiation facilities is coming of age in the first decade of the twenty-first century, which will be defined by a greatly improved performance, especially with regards to the coherence and brilliance of the x-rays, using so-called energy recovery linacs (ERLs), and free electron lasers (FELs). There is, respectively, a further 3 and 10 orders of magnitude in peak brilliance to be gained using ERLs and FELs, and facilities such as the Linac Coherent Light Source (LCLS) at Stanford, and the FLASH and European XFEL facilities at Hamburg, are presently striving to achieve this goal. The first hard x-rays from an FEL were produced at the LCLS in April 2009.

References

[1] H. Michel, "Three-dimensional crystals of a membrane-protein complex – the photosynthetic reaction center from rhodopseudomonas-viridis," *J. Molec. Biol*., 158, pp. 567–572, 1982.

[2] J. Deisenhofer, O. Epp, K. Miki, R. Huber, and H. Michel, "Structure of the protein subunits in the photosynthetic reaction centre of rhodopseudomonas viridis at 3 Å resolution," *Nature*, 318, pp. 618–624, 1985.

[3] S. Jenni, M. Leibundgut, D. Boehringer, C. Frick, B. Mikolasek, and N. Ban, "Structure of fungal fatty acid synthase and implications for iterative substrate shuttling," *Science*, 316, pp. 254–261, 2007.

[4] J. Maddox, "The Sensational Discovery of X-rays," *Nature*, 375, p. 183, 1995.

[5] M. F. Perutz, "How W. L. Bragg invented x-ray analysis," *Acta Crystallogr. A*, 46, pp. 633–643, 1990.

[6] J. D. Watson and F. H. C. Crick, "Molecular structure of nucleic acids – A structure for deoxyribose nucleic acid," *Nature*, 171, pp. 737–738, 1953.

[7] J. D. Watson, *The double helix: a personal account of the discovery of the structure of DNA*. Atheneum, 1968.

[8] J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, and D. C. Phillips, "A three-dimensional model of the myoglobin molecule obtained by x-ray analysis," *Nature*, 181, pp. 662–666, 1958.

[9] M. F. Perutz, M. G. Rossmann, A. F. Cullis, H. Muirhead, G. Will, and A. C. T. North, "Structure of haemoglobin: A three-dimensional Fourier synthesis at 5.5 Å resolution, obtained by x-ray analysis," *Nature*, 185, pp. 416–422, 1960.

[10] G. E. Smith, "Nobel Lecture: The invention and early history of the CCD," *Rev. Mod. Phys*., 82, pp. 2307–2312, 2010.

[11] D. A. Doyle, J. M. Cabral, R. A. Pfuetzner, A. L. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait, and R. MacKinnon, "The Structure of the Potassium Channel: Molecular Basis of K^+ Conduction and Selectivity," *Science*, 280, pp. 69–77, 1998.

[12] P. Cramer, D. A. Bushnell, and R. D. Kornberg, "Structural basis of transcription: RNA polymerase II at 2.8 Å resolution," *Science*, 292, pp. 1863–1876, 2001.

[13] E. Pennisi, "The race to the ribosome structure," *Science*, 285, pp. 2048–2051, 1999.

[14] F. R. Elder, R. V. Langmuir, and H. C. Pollock, "Radiation from electrons accelerated in a synchrotron," *Phys. Rev*., 74, pp. 52–56, 1948.

[15] D. Iwanenko and I. Pomeranchuk, "On the maximal energy attainable in a betatron", *Phys. Rev*., 65, p. 343, 1944.

[16] D. H. Tomboulian and P. L. Hartman, "Spectral and angular distribution of ultraviolet radiation from the 300-MeV Cornell Synchrotron," *Phys. Rev*., 102, pp. 1423–1447, 1956. http://prola.aps.org/abstract/PR/v102/i6/p1423_1.

2

The Interaction of X-rays with Matter

2.1 Introduction

How does electromagnetic radiation interact with matter? We are all familiar at some level with phenomena in the visible region. Refraction describes the dependence of the velocity of light on the medium in which it is travelling – the higher the refractive index, the slower the light travels. This causes light to be bent when it crosses the boundary between one medium and another. Below a certain incidence angle, a beam of light impinging on the surface of a material with a lower refractive index will be completely reflected and only penetrate the boundary in the form of a decaying, or 'evanescent" standing wave to a depth comparable to the light's wavelength.

Conducting materials which contain free or nearly-free electrons absorb and re-emit (that is, they reflect) light incident at any angle efficiently – the reflectivity of aluminium over the entire visible range and also well into the ultraviolet region is near unity, while gold reflects generally less well and has a yellow (or golden!) appearance, because of an absorption band in the blue.

Visible light, however, spans only a very narrow energy range of about 1.75 to 3.2 eV (or, more familiarly, 400 to 700 nm in wavelength), and hence virtually all interactions in this region only involve free electrons or those occupying the outermost (valence) shell,[1] as all other electrons bound to atomic nuclei are much more strongly bound and less affected by such relatively weak perturbations. But how does material respond to electromagnetic radiation of higher photon energy?

We will see that the above-mentioned phenomena of refraction, reflection and absorption are all also found when studying the interaction of ultraviolet rays and x-rays with matter, whereby the main differences compared to the familiar properties of visible light

[1] The electron shell can be thought of classically as the 'orbit' of the electron around the atomic nucleus. It is defined by the principal quantum number, n. For a given atom, the maximum value of n, which is normally associated with the most weakly bound 'valence' electrons, is given by the period in which the atom resides in the periodic table.

An Introduction to Synchrotron Radiation: Techniques and Applications, First Edition. Philip Willmott.

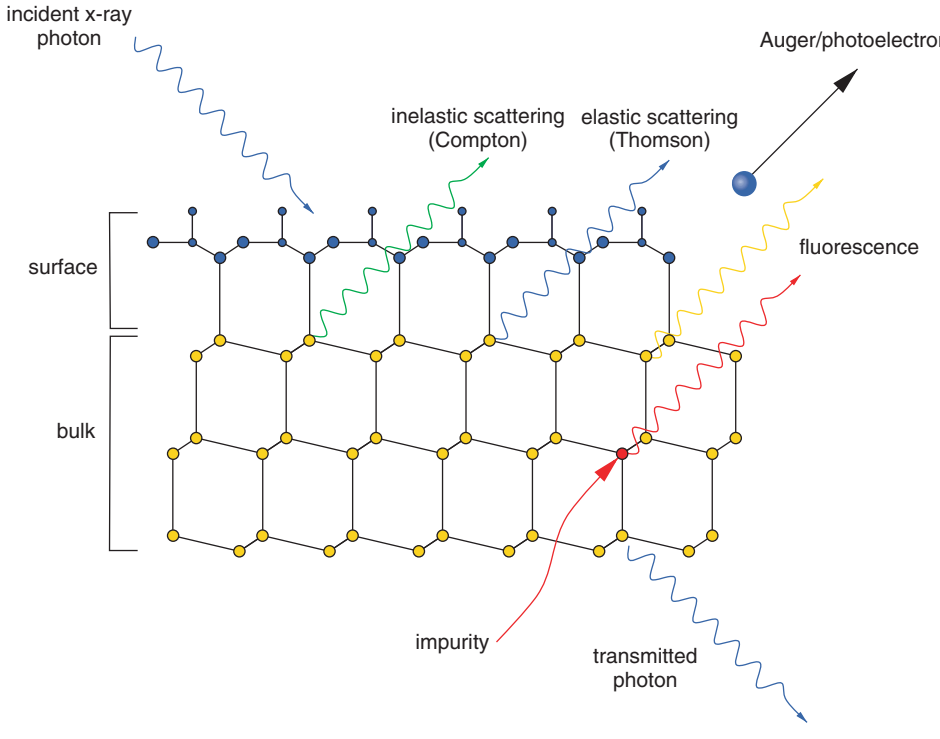

Figure 2.1 *The interaction of x-rays with matter. Surface (and interface) regions of a solid or liquid material are characterized by physical properties and structures that may differ significantly from those of the bulk structure. The x-rays may be elastically or inelastically scattered, or absorbed, in which case electrons or lower-energy photons can be emitted. If none of the above occur, the photon is transmitted through the sample.*

with matter can be explained by the much higher photon energies involved. The most important phenomena are sketched in Figure 2.1.

The binding energy of the only electron in a hydrogen atom is 13.6 eV. To a first approximation, the innermost electrons in an atom of atomic number Z have binding energies which are proportional to Z^2, as shown in Figure 2.2 for the $1s$-electrons. X-rays therefore have energies similar to or larger than those of tightly-bound core electrons, and hence the interaction strengths or 'cross-sections' between x-rays and core electrons are stronger than those between x-rays and valence electrons.

In this chapter, we will discuss how x-rays interact with matter and how this manifests itself in the phenomena of refraction, reflection, absorption and fluorescence. Diffraction phenomena are treated separately in Chapter 5.

2.2 The Electromagnetic Spectrum

The use of light to investigate the properties and structures of materials goes back to the dawn of science. Before the end of the nineteenth century, mankind only had visible light

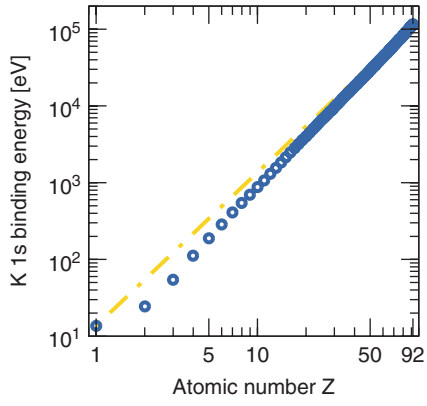

Figure 2.2 *The binding energies of the 1s-electrons in the elements up to uranium. Also shown is the zeroth-order approximation $E_B(1s) = 13.6Z^2$ (which ignores any Coulomb interaction between the electrons in the atom) as a comparison. The actual binding energies deviate somewhat from this simple expression, due to the screening effect of the other electrons in the atom, particularly evident for low-Z elements, for which the average distances and hence binding energies of the other electrons are more similar to that of the 1s-electron.*

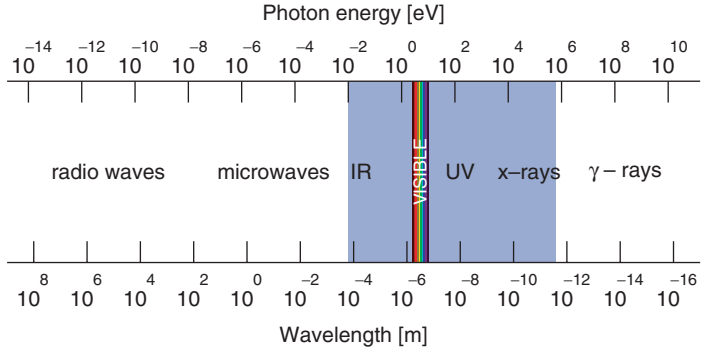

Figure 2.3 *The electromagnetic spectrum. Worldwide, synchrotrons cover photon energies from the infrared at around 0.01 eV to very hard x-rays at 5×10^5 eV, as indicated by the blue region.*

at its disposal. Visible light is just a small section of the entire electromagnetic (EM) spectrum, which spans longwave radio waves, through microwaves, infrared radiation, visible light, ultraviolet light, x-rays and gamma rays (see Figure 2.3).

The difference in energies in the spectrum shown in Figure 2.3 spans some 23 orders of magnitude (for example the energy of a gamma ray photon produced by the annihilation of a proton with its antiparticle is about 1 GeV, some 10^{18} times greater than that of a photon emitted by a longwave radio antenna). In general, the energy of a photon is given by

$$E = h\nu = hc/\lambda, \tag{2.1}$$

where $h = 6.626 \times 10^{-34}$ J s is Planck's constant, ν is the frequency of the radiation in Hz, $c = 2.9979 \times 10^{8}$ m s^{-1} is the speed of light in vacuum, and λ is the wavelength of light in vacuum, given in metres. A convenient way of re-expressing Equation (2.1) to determine the wavelength of x-ray photons for a given photon energy is

$$\lambda\,[\text{Å}] = \frac{12.3984}{E\,[\text{keV}]}. \tag{2.2}$$

Hence, a 1 Å x-ray photon has an energy of 12 398.4 eV (or about 2 fJ).

In this chapter, we first consider the interaction of x-rays with individual electrons (their primary scatterers in our region of interest), followed by scattering from an ensemble of bound electrons in an atom. From this, we discuss the phenomena of refraction, absorption and reflection. The phenomena of diffraction from a regular array of atoms, and of small-angle scattering from mesoscopic objects, although logical progressions in complexity, are deferred to Chapter 5.

The cross-sections for the processes involving the interaction of x-rays with matter are shown in Figure 2.4 for the example of elemental barium. For applications using synchrotron radiation, only photoelectric absorption, elastic Thomson scattering and (to a lesser extent) inelastic Compton scattering are important [1]. These processes are now described in more detail.

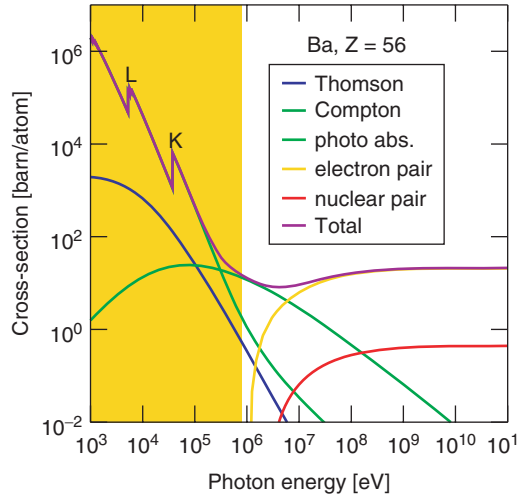

Figure 2.4 *The cross-sections for various processes involving the interaction of x-rays with matter, for the example of the element barium (Ba). The unit 'barn' is equal to 10^{-24} cm^2. The yellow region indicates the upper energy range covered by synchrotron sources (which also continues down to a few meV, five orders of magnitude lower still). In this range, two processes dominate, namely photoelectric absorption (here, the K- and L-edges of Ba are marked, see later) and elastic (Thomson) scattering, although inelastic (Compton) scattering also becomes significant above approximately 30 keV.*

2.3 Thomson Scattering

For all but the most energetic gamma rays, the primary scattering particle for x-rays in a medium is the electron. Classically, elastic scattering occurs when an electron is sinusoidally accelerated by the incoming electromagnetic field and reradiates light at exactly the same frequency – the process is therefore elastic.

If the charged particles (here, electrons) are not moving relativistically, they are accelerated mainly by the electric field component of the incident wave. As shown in Figure 2.5, the electron will move in the same direction as the oscillating electric field, resulting in emission of electromagnetic radiation. The strength of the re-emitted radiation seen at some observation point therefore depends on the magnitude of the component of the electric field seen from that point. It should therefore be clear from Figure 2.5 that no scattering occurs along any line parallel to the electric field vector ($\chi = \pi/2$) as, in such directions, the projected oscillation amplitude of the electron is zero. In contrast, an observer anywhere in the plane perpendicular to the electric field vector sees a maximal scattering oscillation amplitude of the electron. The scattered *intensity*, which is proportional to the square of the electric field amplitude, is therefore proportional to $\cos^2 \chi$, where χ is the angle between the plane perpendicular to the electric field and the direction of observation. The form of this intensity distribution lends it the name electromagnetic *dipole* radiation.

The ability of an electron to scatter an x-ray is quantified by a scattering length. For an isolated electron, a convenient unit is the *Thomson scattering length*, which can be thought of as the 'classical' radius of a free electron (this can be physically understood

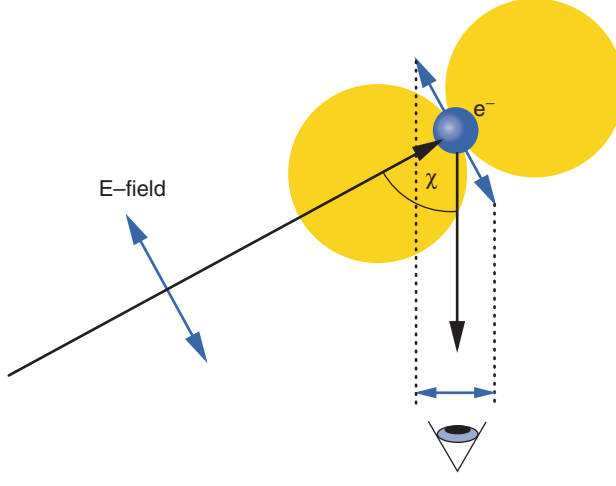

Figure 2.5 *Dipole radiation caused by Thomson scattering of electromagnetic radiation by a charged particle.*

by equating the rest-mass energy of the electron mc^2 with the electrostatic energy at radius r_0). It is given by

$$r_0 = \left(\frac{e^2}{4\pi\varepsilon_0 mc^2}\right) = 2.82 \times 10^{-5}\text{Å}. \tag{2.3}$$

The elastic scattering amplitude of an atom is proportional to the Thomson scattering length and Z, the number of electrons in the (neutral) atom. The elastic scattering *intensity* therefore increases with Z^2.

2.4 Compton Scattering

Photoabsorption, described in Section 2.6, and elastic (Thomson) scattering are the most important processes in the majority of techniques used to investigate the structure of materials. In the quantum-mechanical description of the excitation of an electron by electromagnetic radiation, however, kinetic energy may be transferred to the electron, resulting in the scattered photon having a lower energy [2]. This inelastic process is called the Compton effect (see Figure 2.6). One should not confuse photoabsorption with Compton scattering. In the former, a photon is entirely removed from the system. This may or may not result in the ejection of a lower-energy photon via fluorescent decay, but these discrete-energy photons do not have the same origin as Compton-scattered photons, which exhibit a continuous spectrum. The energy loss in Compton scattering is simple to determine by applying conservation of total energy and momentum, and results in

$$\frac{h\nu_0}{h\nu} = \frac{k_0}{k} = \frac{\lambda}{\lambda_0} = 1 + \lambda_C k_0(1 - \cos\psi), \tag{2.4}$$

where the incoming photon has the frequency $\nu_0 = c/\lambda_0$, its wavevector is $k_0 = 2\pi/\lambda_0$, ψ is the angle by which the photon is scattered and λ_C is the Compton scattering length, given by

$$\lambda_C = \frac{h}{mc} = 2.43 \times 10^{-2}\ \text{Å}. \tag{2.5}$$

Therefore, the fractional loss of the photon energy increases with increasing scattering angle and incident photon energy. For most applications using synchrotron radiation, the

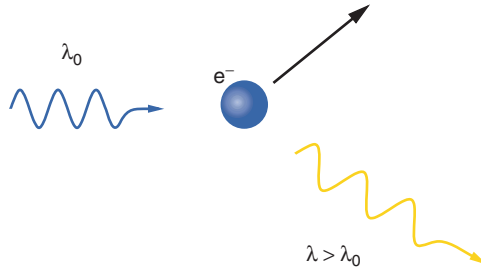

Figure 2.6 *The Compton effect, whereby a fraction of the energy of a photon is transferred to the kinetic energy of a photon, resulting the scattered photon having a lower energy.*

cross-section for Compton scattering and its effect are small and can be neglected (see Figure 2.4).

2.5 Atomic Scattering Factors

2.5.1 Scattering From a Cloud of Free Electrons

The optical properties of materials in the energy range above about 30 eV can be described by the atomic scattering factors (also called the atomic form factors). We begin by assuming the electrons in an atom can move freely and are unaffected by being in bound states. It has already been stated that the electron is the most important scattering particle for the photon energies of interest. An atom will not act as a point scattering centre, but will scatter x-rays throughout the volume occupied by the electron cloud surrounding the nucleus.[2] In the forward-scattering direction, the scattering strength is simply Z times that from a single electron. In this case, all of the elements of the electron cloud scatter in phase with each other.

However, the scattering amplitude at some angle 2θ to the incident beam will be less than Z, as the scattered waves from different parts of the electron cloud will have a nonzero phase relative to one another and the scattering amplitude at 2θ will therefore be equal to the vector sum of the scattering amplitudes in this direction from each point of the cloud, taking into account the phases between them, as shown schematically in Figure 2.7(a).

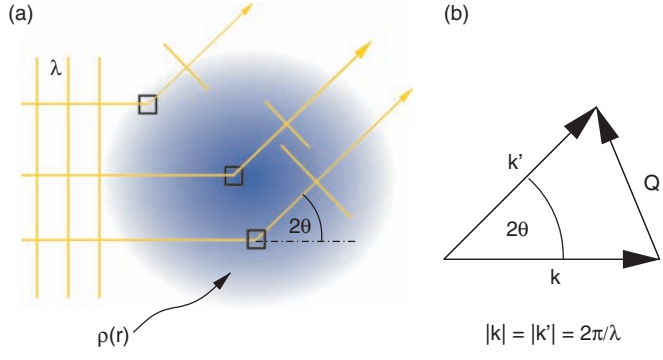

Figure 2.7 *Elastic scattering of x-rays by an electron cloud around an atom with an electron density $\rho(\mathbf{r})$. (a) Scattered waves from volume elements of the electron cloud have different scattering amplitudes, proportional to $\rho(\mathbf{r})$, and a phase ϕ, relative to some arbitrary origin, which depends on the position* **r** *relative to that origin. (b) The scattering vector* **Q** *is the vectorial difference between the incident and elastically scattered wavevectors* **k** *and* **k**′*, respectively.*

[2] One might expect the positively charged nucleus to also scatter x-rays. However, as we have seen in Equation (2.3), the Thomson scattering length is inversely proportional to the mass of the particle, while the scattering cross-section is proportional to the square of the Thomson scattering length, and hence the nucleus contributes to less than one part in a million to the scattering amplitude and can be completely ignored.

It is a trivial exercise to show that the vectorial difference between an incoming x-ray beam with wavevector $\mathbf{k}$ and wave $\mathbf{k}'$ elastically scattered through an angle 2θ, given by

$$\mathbf{Q} = \mathbf{k} - \mathbf{k}', \tag{2.6}$$

has a magnitude

$$Q = \frac{4\pi}{\lambda} \sin\theta. \tag{2.7}$$

$\mathbf{Q}$ is the 'scattering vector', shown in Figure 2.7(b), and $\hbar\mathbf{Q}$ is the momentum transfer imparted on the scattered photon. An important aspect of Thomson scattering is that the scattered photon has a phase shifted by π relative to the incoming photon. This is a consequence of the refractive index of materials in the x-ray regime being less than unity.

The atomic scattering factor (also called the atomic form factor) f describes the total scattering amplitude of an atom as a function of $\sin\theta/\lambda$ (which is proportional to the scattering vector $\mathbf{Q}$) and is expressed in units of the scattering amplitude produced by one electron. It falls off quasimonotonically with increasing Q.

As already stated, in the forward scattering direction, $f = f(0)$ is nothing more than the electron density, integrated over the atom's electron cloud, and is therefore equal to Z, the atomic number. Note, however, that in the case of an ionic species, $f(0)$ will differ from Z by an amount equal to the amount of electron transfer to or from neighbouring ions. Hence cations have $f(0)$-values less than Z and anions values greater than Z.

It is conventionally assumed that the electron density of an atom is roughly spherically symmetric and hence f depends only on the magnitude of $\mathbf{Q}$ and not on its orientation relative to the scattering atom. Values for $f(\sin\theta/\lambda)$ can be either directly found in the *International Tables for Crystallography*, or can be calculated using nine tabulated coefficients and the expression

$$f(\sin\theta/\lambda) = \sum_{i=1}^{4} a_i \exp(-b_i \sin^2\theta/\lambda^2) + c. \tag{2.8}$$

Examples of the dependence of f on $\sin\theta/\lambda$ are given in Figure 2.8. Note that, for both La and B, $f(0)$ does indeed equal Z.

2.5.2 Correction Terms for the Atomic Scattering Factor

The above description assumes that the electrons are essentially free, that is, their movement is not damped by the fact that they are bound in an atom. In reality, the response of bound electrons in an atom, molecule or bulk material, to the electric field of the incoming x-rays is such that additional energy-dependent terms must be added to more completely describe the structure factor. We consider these now.

Electrons bound in material have discrete energy levels. If the incident x-ray photon has an energy much less than the binding energy of the tightly-bound core electrons, the response (i.e. oscillation amplitude) of these electrons to the driving EM-field will be damped because they are bound. The real part of the elastic scattering factor, which we

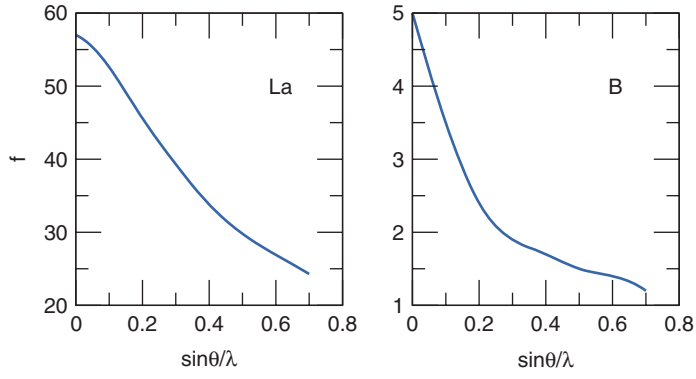

Figure 2.8 *The real parts of the atomic scattering factors for La and B as a function of* $\sin\theta/\lambda$.

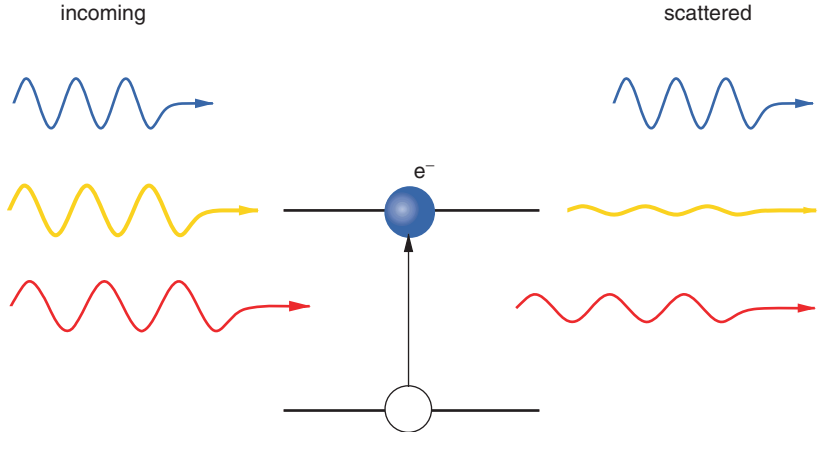

Figure 2.9 *Elastic scattering of an x-ray by a bound electron. The amplitude of the scattered wave decreases as one approaches a discrete energy level of an electron bound in an atom. Far from resonant edges, the electrons are essentially free and the damping term f′ is insignificant.*

henceforth refer to as f_1 (in order to distinguish it from our 'start' definition f for a free electron) will therefore be reduced by an additional term f', that is

$$f_1 = f - f'. \tag{2.9}$$

In contrast, at photon energies far above absorption resonances, the electrons are essentially free and $f' = 0$ (see Figure 2.9).

Close to an absorption edge, the x-rays will be partially absorbed. After a delay, some x-radiation is re-emitted, which interferes with the directly (but also elastically) scattered part, thereby altering both the phase and the amplitude. This is expressed in the description of the atomic form factor by introducing a second additional term accounting for the increasing phase lag relative to the driving electromagnetic field of the x-rays as the photon energy approaches an absorption edge. Note that the interference between

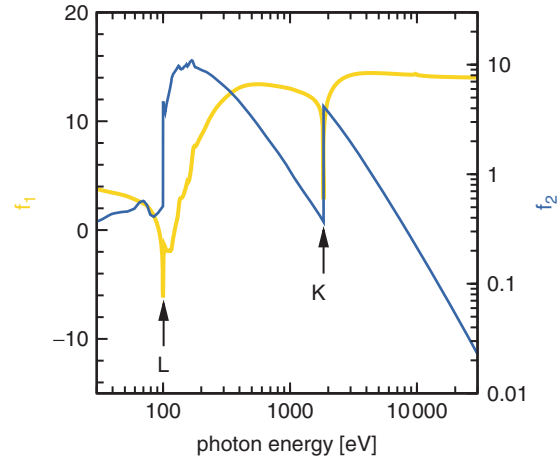

Figure 2.10 *The real ($f_1 = f - f'$) and imaginary ($f_2 = f''$) components of the atomic form factor for Si as a function of photon energy. Note that f_1 converges to a constant value at photon energies far above absorption energies equal to the number of electrons in the atom, $Z = 14$. f_2, on the other hand, falls away as E^{-2}.*

the incoming and emitted radiation causes energy loss (i.e. absorption), and the new second additional term if'' is imaginary. The mathematical justification for this is given in Section 2.6.3. This term shows sharp increases at absorption edges, but otherwise falls off as E^{-2} (note the gradient of -2 in the double logarithmic plot of Figure 2.10).

More precisely, f'' is related to the photoabsorption cross-section σ_a at that photon energy by

$$f''(0) = \frac{\sigma_a}{2r_0\lambda}. \tag{2.10}$$

From Figure 2.4, we can see that, for example, at 12.4 keV photon energy (1 Å wavelength), the photoabsorption cross-section of Ba is of the order of 10^{-24} m^2 and hence $f'' \sim 1$.

We refer to the imaginary component of the total (i.e. complex) atomic scattering factor as $f_2 = f''$, and hence

$$f_{\text{tot}} = f_1 + if_2. \tag{2.11}$$

The atomic form factors for forward scattering for Si are shown in Figure 2.10. We can qualitatively understand the curves as follows. At the binding energies of the K- and L-shells, there are sharp decreases in f_1 and corresponding sharp increases in f_2. Mid between these resonances, coupling of the x-rays to the material is inefficient, and f_1 increases again. At high photon energies, there is only a weak interaction of the x-ray photons with the atoms, and $f_1 \approx Z = 14$ and f_2 (which, remember, represents absorption), falls off to very small values.

2.6 The Refractive Index, Reflection and Absorption

2.6.1 The Refractive Index

The index of refraction n at a wavelength λ describes the response of electrons in matter to electromagnetic radiation, and is a complex quantity which we express as

$$n = n_R + in_I, \tag{2.12}$$

In the x-ray regime, n is related to the atomic scattering factors of the individual atoms in a material by

$$n = 1 - \frac{r_0}{2\pi}\lambda^2 \sum_i N_i f_i(0), \tag{2.13}$$

where N_i is the number of atoms of type i per unit volume and $f_i(0)$ is the complex atomic scattering factor in the forward direction of the ith atom.

Because the phenomena of refraction and absorption both depend on the response of the electrons in a material to the oscillating electric field of an electromagnetic wave propagating through that material, the real and imaginary components of $f(0)$ are coupled,[3] as can be clearly seen in Figure 2.10.

2.6.2 Refraction and Reflection

We have argued that when electromagnetic radiation passes through a solid, it interacts with the electrons within the material. For visible light, this normally means a reduction of the group velocity by a factor equal to the real part of the refractive index n of that material. This can, for example, be used to bend or focus light with optical lenses. But n is a function of the wavelength of the radiation; in other words, the material disperses the light. This is why rainbows occur – they are manifestations of refraction and reflection (and are *not*, as commonly misconceived, a diffraction phenomenon). Normally, as an absorption maximum of a material is approached from lower frequencies, the refractive index increases. Therefore blue light is refracted more by water or quartz than is red light.

Past the absorption maximum, however, the refractive index drops to values less than unity. In such cases, it is the *phase velocity* that has the value c/n. This is greater than the velocity of light, but it is the *group velocity* that carries the energy, and this remains below c – the laws of relativity remain inviolate!

This is the region of so-called *anomalous dispersion*. This means that (a) as radiation enters a material with a refractive index less than unity, it is refracted to *shallower* angles to the interface, and (b) the scattered wave is phase-shifted by π radians with respect to the incident wave. We express the complex refractive index as

$$n = 1 - \delta + i\beta, \tag{2.14}$$

[3] The precise relationship between these properties is described by the Kramers–Kronig relation, though it lies outside the scope of this book to detail this.

where the parameters δ and β are called the refractive index decrement and absorption index, respectively. Comparing this equation with Equation (2.12), we immediately recognize that $n_R = 1 - \delta$, while the absorption index $\beta = n_I$.

The refractive index decrement δ is equal to the real part of the summed term in Equation (2.13). Far from an absorption edge, we can ignore the correction term f' in Equation (2.9), and the sum $\sum_i N_i f_{1,i}(0) = \rho$ is merely the average density of electrons, and hence

$$\delta = \frac{2\pi\rho r_0}{k^2}. \tag{2.15}$$

δ is therefore inversely proportional to the square of the photon energy. If we assume a typical value of one electron per cubic Angstrom (for example, Si has an average electron density of 0.7 e Å^{-3}), we obtain a value of $\delta \approx 5 \times 10^{-6}$. Hence, for x-rays, the real part of n is really only *very* slightly less than unity.

2.6.2.1 Snell's Law and Total External Reflection

Snell's law relates the incident angle α with the refracted angle α' (see Figure 2.11)

$$\frac{\cos\alpha}{\cos\alpha'} = n_R. \tag{2.16}$$

This means that for a value of n_R of less than unity there will be a minimum incident angle for which the x-rays are refracted into the bulk of the material. This is known as the critical angle, α_c, such that

$$\cos\alpha_c = n_R. \tag{2.17}$$

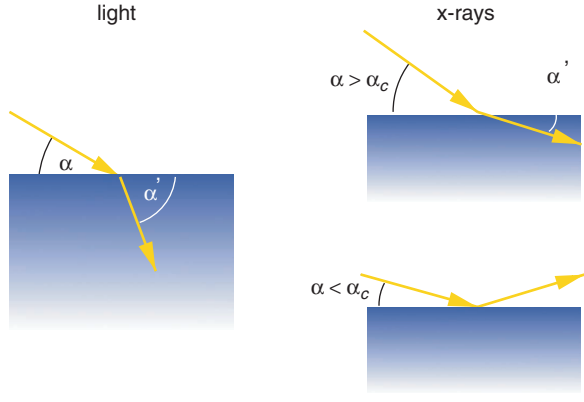

Figure 2.11 *The refraction of visible light by a transparent medium is greater than unity (water has a refractive index in the visible of approximately 1.33). In contrast, the refractive index for x-rays is marginally less than one. At grazing angles less than the critical angle α_c, total external reflection will occur.*

For x-rays impinging on a surface at an angle below α_c, they will undergo *total external reflection*.

The first two terms of the Taylor expansion of the cosine function are $1 - \theta^2/2$, which is a very accurate approximation for small angles. For example, the fractional difference between the precise value for $\cos(1^\circ)$ and this approximation is less than four parts in one billion. So, combining Equations (2.14) and (2.17), we obtain

$$\alpha_c \approx \sqrt{2\delta}. \tag{2.18}$$

For typical values of δ of around 10^{-5}, α_c is therefore of the order of a few milliradians, or a few tenths of a degree. From Equation (2.15), it can also be deduced that α_c is inversely proportional to the photon energy.

Hence, at low incident angles, x-rays can be made to reflect from surfaces. Indeed, if the surface of an x-ray mirror is made to be concave (ideally, parabolic), the reflected x-rays can be focused, as long as curvature is sufficiently small that the steepest incident angle is less than α_c. Such x-ray focusing mirrors, as shown in Figure 2.12, are obviously large beasts, as the 'footprint' of the x-ray beam on the mirror surface is magnified in the beam direction by a factor of $1/\sin\alpha$. X-ray mirrors are described in detail in Section 4.3.

Although x-rays are totally reflected from surfaces if their incident angle is less than α_c, a moment's reflection should make it obvious that there must be *some* sort of interaction between the x-rays and the material in order for reflection to occur at all. In actual fact, an *evanescent* standing wave of the x-rays does penetrate marginally into the refractive medium to a depth of a few nanometres only (see Figure 2.13). This is some three orders of magnitude smaller than the penetration depth of x-rays at angles well above the critical angle. It is therefore possible to use x-rays as a *surface-sensitive probe* down to the atomic level, using grazing-incidence radiation, as will be described in detail in Section 5.13.

Above the critical angle, the reflectivity for an x-ray in vacuum impinging at an angle α on a perfectly flat surface of a material with a real part of the refractive index $n_R = 1 - \delta$, drops off rapidly (see Figure 2.13) and is given by the Fresnel equations, dealt with in more detail in Section 5.15.

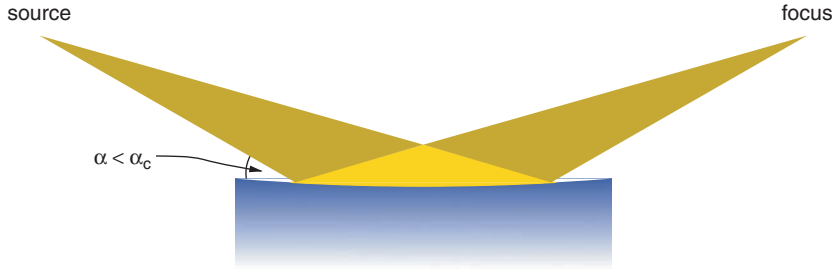

Figure 2.12 *An x-ray focusing mirror has a radius of curvature such that the variation in the incident beam angle on the mirror surface is much less than the critical angle α_c.*

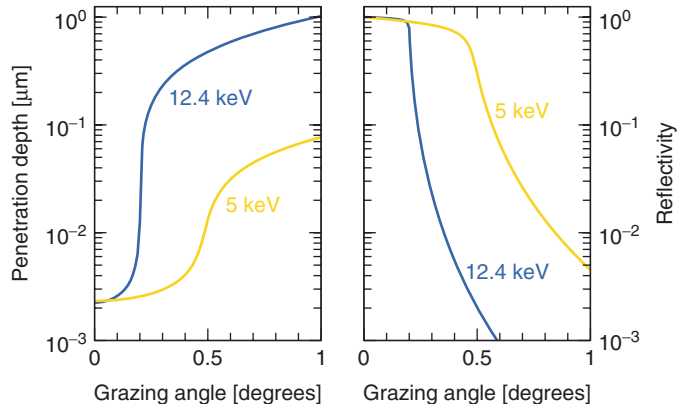

Figure 2.13 *The penetration depth and reflectivity of 5 keV and 12.4 keV photons in $SrTiO_3$ ($\rho = 5.12\,g\,cm^{-3}$). Below the critical angle, the reflectivity is close to unity, while the evanescent wave penetrates the surface by only 2 to 3 nm. Above the critical angle, the x-rays rapidly begin to penetrate deeply into the material, which is accompanied by a catastrophic reduction in the reflectivity.*

2.6.3 Absorption

2.6.3.1 The Absorption Coefficient

From Equations (2.13) and (2.14), it can be seen that the absorption index β is related to f'' by

$$\beta = \frac{r_0}{2\pi}\lambda^2 f''. \tag{2.19}$$

We have already stated that away from absorption edges, f'' is proportional to the inverse square of E, hence from Equation (2.19) it should be apparent that β is proportional to the inverse fourth power of the energy.

Let us imagine an x-ray beam travelling in vacuum and then penetrating some sort of absorbing medium (Figure 2.14). We begin with the classical description of the electric field vector of a linearly-polarized electromagnetic wave propagating in the z-direction through vacuum

$$E(z,t) = E_0 \exp\left[i(k_0 z - \omega t)\right], \tag{2.20}$$

where $k_0 = 2\pi/\lambda_0$ is the wavevector in vacuum and λ_0 is the wavelength in vacuum.

When the wave travels through the medium, however, the wave function changes to

$$E(z,t) = E_0 \exp\left(ink_0 z - \omega t\right), \tag{2.21}$$

where we have now included the complex refractive index n. In other words, the wavelength of the x-rays is changed by the medium, although their frequency (ω) is, obviously, unaffected. Substituting Equation (2.12) into Equation (2.21), we obtain

$$E(z,t) = E_0 \exp(-n_I k_0 z) \exp\left(in_R k_0 z - \omega t\right). \tag{2.22}$$

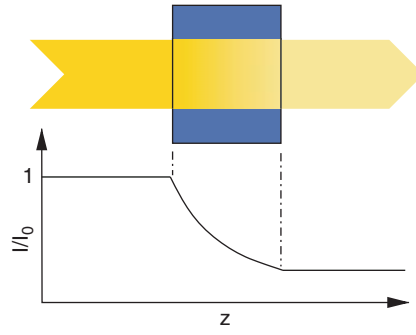

Figure 2.14 *The transmission of x-rays through a material falls off in an inverse exponential manner with penetration depth.*

The first exponential term describes the attenuation of the amplitude as the wave travels through the medium. One measures, however, the intensity, which is proportional to the square of the amplitude, and therefore drops by an amount $1/e$ over a depth $1/(2n_I k_0)$. This is the absorption coefficient. The exponential decay behaviour of the transmitted signal

$$\frac{I}{I_0} = \exp(-\mu z) \tag{2.23}$$

is the Beer–Lambert equation for linear absorption. We can now equate the negative absorption index $n_I = \beta$ to the absorption coefficient μ by

$$\mu = 2n_I k_0. \tag{2.24}$$

For any given material, therefore, its absorption coefficient μ is defined as the reciprocal of the thickness d of the material which is needed to reduce the intensity of an impinging electromagnetic beam by a factor of $1/e$.

Clearly, μ is element-specific and a function of the x-ray energy. In general, it is strongly dependent on the atomic number of the element (it varies approximately as Z^4, see Figure 2.15) and decreases with increasing photon energy, for which it varies as approximately E^{-3} (Figure 2.16(a)).

The absorption coefficient is essentially an indication of the electron density in the material and the electron binding energy. Therefore, if a particular chemical substance can assume different geometric ('allotropic') forms and thereby have different densities, μ will be different accordingly. The allotropes of carbon, diamond, graphite, C_{60}, and glassy carbon, have densities of 3.53, 2.23, 2.0 and 1.69 g cm^{-3}, respectively. Accordingly, their attenuation lengths, for example 1.56, 2.47, 2.76 and 3.26 µm, respectively, at 284 eV, are inversely proportional to their densities.

Conversely, compounds that are chemically distinct but contain the same number of electrons per formula unit and have similar mass densities will have similar absorption properties (except close to absorption edges). For example, the attenuation lengths of NaCl, MgS, AlP and Si (normalized to their mass densities) are shown in Figure 2.17. The average number of electrons per atom is 14. The small differences in absorption at

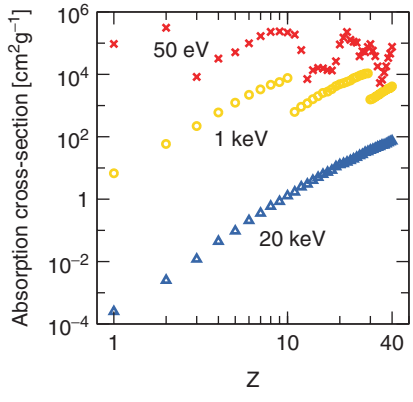

Figure 2.15 *The photoabsorption cross-section ($\sigma_a = \mu/\rho_m$) as a function of atomic number Z for three different photon energies. On this double logarithmic scale, the curves are close to being linear with a gradient of four, except for occasional jumps, due to shifts in absorption edges with Z. For example the drop in absorption between $Z = 10$ (neon) and $Z = 11$ (sodium) for the 1000 eV curve is a result of their respective strongest absorption edges (Kα) lying at 870.2 eV and 1070.8 eV.*

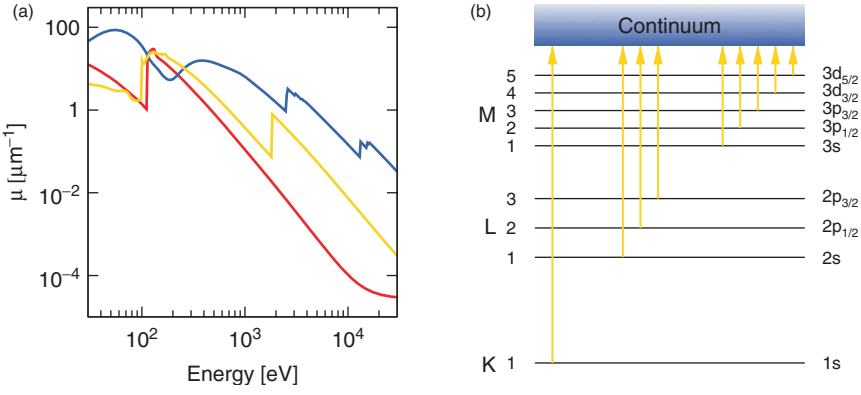

Figure 2.16 *(a) The x-ray absorption coefficient μ for beryllium (red curve), silicon (yellow) and lead (blue) as a function of the photon energy. (b) Sharp increases in the absorption occur when the photon energy is just sufficient to eject the electron from the electronic orbital to the continuum. The x-ray absorption (left) and atomic orbital labellings (right) are shown.*

high photon energies are due to differences in the average isotopic ratios of protons to neutrons from element to element. Note also how all the absorption edges are centered around that of Si.

A more fundamental parameter is therefore the *atomic* absorption cross-section of the element, σ_a, briefly introduced in Equation (2.10). It has units of area, such as cm^2, or

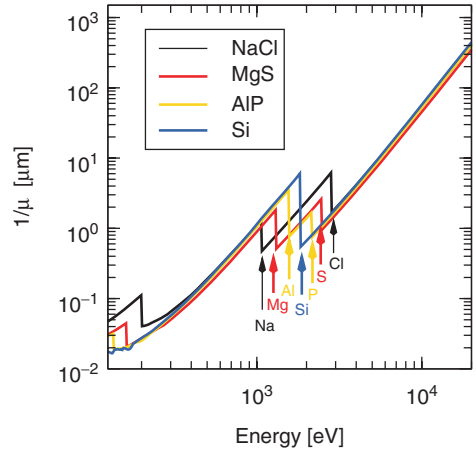

Figure 2.17 *The attenuation lengths of NaCl, MgS, AlP and Si (normalized to their mass density) as a function of x-ray photon energy. Far from absorption edges, the curves have very similar values. Note also how the positions of the K-edges of the component elements converge to that of Si as the number of electrons in those elements approach that of Si (Z = 14).*

the barn,[4] where 1 b is equal to 10^{-28} m^2. The relationship between σ_a and μ is

$$\mu = \rho_a \, \sigma_a = \left(\frac{\rho_m \, N_A}{A} \right) \sigma_a \tag{2.25}$$

where ρ_a, N_A, ρ_m and A are the atomic number density, Avogadro's number, the mass density and the atomic mass, respectively. More commonly, however, the absorption cross-section is given in cm^2 g^{-1}, in which case

$$\sigma_a \, [\text{cm}^2 \, \text{g}^{-1}] = \frac{N_A}{A} \, \sigma_a \, [\text{cm}^2] = \frac{\mu}{\rho_m}. \tag{2.26}$$

2.6.3.2 *Absorption Edges and Nomenclature*

As can be seen in Figures 2.16 and 2.17, the absorption spectra show occasional sudden sharp increases, known as absorption edges. These are produced by resonances between the x-ray photon energy and the minimum energy needed to ionize the atom. The electronic state of an electron in an atom is defined by the principal, orbital angular momentum, and spin quantum numbers n, l and s, respectively. In atomic physics, these states are labelled as $nl'_{\mathbf{j}}$. The angular momentum $l = 0$, 1, 2 and so on is denoted by the letters $l' = s, p, d, f, \ldots$ and $\mathbf{j}$ is the total angular momentum given by the vector sum of the orbital and spin momenta, that is

$$\mathbf{j} = \mathbf{l} + \mathbf{s}. \tag{2.27}$$

Note also that $\mathbf{s} = 1/2$ for an electron.

[4] The etymology of the unit barn originates from the days of wartime nuclear research, when physicists were so flabbergasted by the enormous neutron scattering strength of the uranium nucleus that they exclaimed its nuclear core to be 'as big as a barn'. One barn is indeed about the size of a uranium nucleus.

Table 2.1 *Correspondence between x-ray absorption edges and their electronic configurations.*

Edge	Configuration	Edge	Configuration
K	$1s$	N_1	$4s$
L_1	$2s$	N_2	$4p_{1/2}$
L_2	$2p_{1/2}$	N_3	$4p_{3/2}$
L_3	$2p_{3/2}$	N_4	$4d_{3/2}$
M_1	$3s$	N_5	$4d_{5/2}$
M_2	$3p_{1/2}$	N_6	$4f_{5/2}$
M_3	$3p_{3/2}$	N_7	$4f_{7/2}$
M_4	$3d_{3/2}$	O_1	$5s$
M_5	$3d_{5/2}$	O_2	$5p_{1/2}$

The nomenclature for absorption edges in x-ray spectroscopy is more pragmatic. The first 18 and their 'translation' to atomic-physics electron configuration notation are given in Table 2.1.

2.7 X-ray Fluorescence and Auger Emission

The two dominant processes occurring subsequent to absorption of an x-ray and the ejection of an x-ray photoelectron are fluorescence and Auger electron emission (see Figure 2.18).

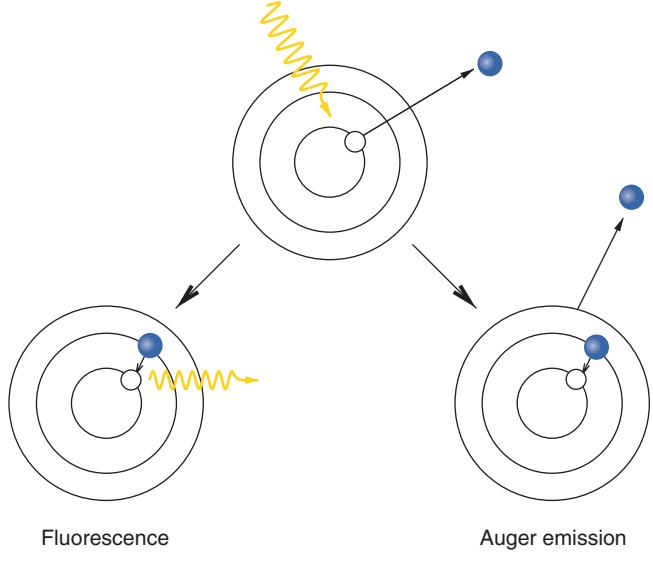

Figure 2.18 *The absorption of an x-ray photon proceeds via the ejection of a core electron. This can then be filled by an electron from a shell further out. The excess energy can either be emitted in the form of characteristic x-ray radiation or by the ejection of an Auger electron.*

In fluorescence, all the excess energy remaining in the atom, after relaxation of an electron from a shell further out to the hole left behind by the photoelectron emission, is carried away by an x-ray photon. In contrast, Auger emission is a nonradiative process, in which the excess energy is carried away by a second electron being ejected. We discuss each process in more detail now.

2.7.1 X-ray Fluorescence

Any material can be made to fluoresce by bombarding it with x-rays or high-energy electrons. If electrons are used, the energy spectrum of the emitted electromagnetic radiation is composed of two parts. A broad background signal is produced as electrons collide with the atoms and are strongly decelerated because any charged particle emits radiation when a force acts upon it. This signal is called 'Bremsstrahlung' (German for braking radiation). It has a maximum cut-off energy corresponding to the highest kinetic energy of the electrons impinging on the target. Superimposed on the Bremsstrahlung are intense and sharp fluorescence lines, which are characteristic for the type of atom from which they originate, as shown in Figure 1.6. In the case of excitation by x-rays, the Bremsstrahlung is missing.

These characteristic x-ray lines result from the transition of an outer-shell electron relaxing to the hole left behind by the ejection of the photoelectron from the atom. This occurs on a timescale of the order of 10 to 100 fs. As the energy difference between the two involved levels is well-defined, these lines are exceedingly sharp. Indeed, from Heisenberg's uncertainty principle, $\Delta E\,\Delta t \sim \hbar$, the natural linewidth is therefore of the order of 0.01 eV, although this depends on the element and the transition. Commonly, values for the energies of characteristic radiation are given to an accuracy of 0.1 eV.

Not all transitions are allowed, as the relaxation of an electron from one bound electronic state to another within an atom (or ion) follow the selection rules for electric dipole radiation (or, in other words, for the emission of an x-ray photon). Photons have an angular momentum of $1\hbar$, hence emission of an x-ray photon in fluorescence must be accompanied by a change in the orbital angular momentum l by one, that is

$$\Delta l = \pm 1. \tag{2.28}$$

Hence, for example, dipole transitions from the $2s$ to the $1s$, or from the $3p$ to the $2p$ level are forbidden.

The spin of the electron has no effect on the spatial distribution of the electron in the atom, and therefore cannot affect the orbital angular momentum. Using this somewhat reverse-logic argument, we arrive at the second selection rule, namely

$$\Delta s = 0. \tag{2.29}$$

From our definition of $\mathbf{j}$ Equation (2.27), we immediately obtain the third selection rule

$$\Delta j = 0, \pm 1; \text{ but not } 0 \rightarrow 0. \tag{2.30}$$

2.7.1.1 Nomenclature and Emission Energies

The nomenclature in x-ray absorption lines of K, L, M and so on described above, actually originates in the early studies of Barkla concerning x-ray emission spectra. In

1911, he described his findings that narrow emission lines characteristic for each element increased in their energy (and thereby, also their penetrating power) with atomic number. He wrote in *The Philosophical Magazine* [3]

> It is seen that the radiations fall into two distinct series, here denoted by the letters K and L*.

He justified the use of this nomenclature in the footnote indicated by the asterisk:

> Previously denoted by letters B and A. The letters K and L are, however, preferable as it is highly probable that series of radiations both more absorbable and more penetrating exist.

Barkla speculated that there might be more energetic series than the K-series, hence his decision to label the observed series in the middle of the alphabet, reserving the letters A to J for up to ten higher-energy series of transitions. In this one instance, he was wrong, although the nomenclature has remained.

Two years later, Henry Moseley recorded the energies of characteristic radiation by dispersing the spectra using diffraction crystals. He discovered a systematic relationship between the energies of the most intense characteristic radiation (which turned out to be the Kα-line, see below) and the atomic number, which he defined in what is now known as Moseley's law:

$$\nu = K(Z-1)^2, \tag{2.31}$$

where $K = 2.47 \times 10^{15}$ Hz, and ν is also given in Hz. He was thus able to show that the characteristic radiation followed the Bohr model of the atom, and that the lines had energies corresponding to the difference in energy of the discrete orbitals [4].

The nomenclature for x-ray fluorescence lines is more complex than that for absorption lines, as they involve two bound orbitals, not one (see Figure 2.19). The K-series involves

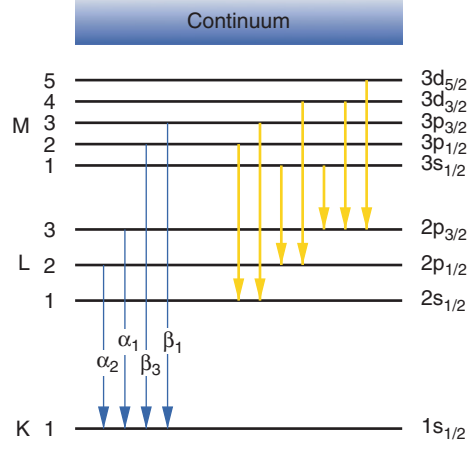

Figure 2.19 *The nomenclature of x-ray fluorescence lines.*

transitions in which the level to which the electron relaxes is the K- (or $1s$) state. Similarly, L-lines have L-states ($n = 2$) as the lower energy level, and so on. The longest wavelength (smallest energy-difference) transition within a series is labelled α, with successive lines denoted β, γ and so on. These lines are actually further split due to the fact that the energy depends subtlely on the orientation of the spin momentum vector with respect to the orbital angular momentum (so-called 'spin-orbit' interactions), labelled by the suffices 1, 2 and so on.

2.7.2 Auger Emission

Auger emission is a three-electron process first discovered by Lise Meitner in 1923, but was named after the French physicist Pierre Victor Auger, who independently discovered and explained the Auger process three years later.

Auger electrons are produced when an outer shell electron relaxes to the core-hole produced by the ejection of a photoelectron. The excess energy produced in this process is $|E_c - E_n|$, whereby E_c and E_n are the core- and outer-shell binding energies, respectively. Instead of being manifested as a fluorescence x-ray photon, the energy can also be channelled into the ejection of another electron if its binding energy is less than the excess energy. The process is shown in Figure 2.20 for ejection of an Auger electron from the same shell as that of the electron which relaxed to the core-level hole.

In this case, (ignoring relaxation effects) the electron energy is $|E_c - 2E_n|$. More generally, the kinetic energy is $|E_c - E_n - E'_m|$, where E'_m is the binding energy of the Auger electron. The prime shows that the binding energy of this level has been changed (normally increased) because the electron ejected from this level originates from an already ionized atom.

The notation of Auger electrons is ABC, where A, B and C relate to the three electronic processes involved in Auger emission, and assume Barkla's notation of K, L, M and so on. An example would be a KLL Auger electron. A is the core-level from which the photoelectron originates (with energy E_c), B is the level from which the electron that fills the core-hole originates (energy E_n) and C is the original level of the Auger electron (energy E'_m).

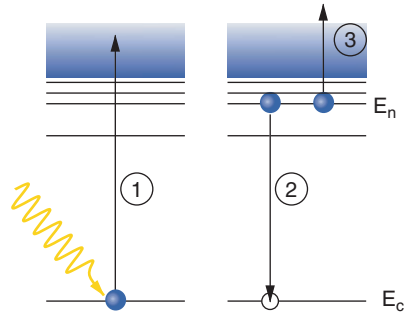

Figure 2.20 *The three-electron process in Auger emission. First, a core-electron is ejected from an atom by absorption of a photon. Next, the system relaxes by an outer electron falling into the core-level, releasing an energy $|E_c - E_n|$. In the Auger process, this energy is channelled into the ejection of a second, or Auger, electron.*

Typical Auger electron energies are in the range of 100 to 500 eV, which, as we shall see in the Chapter 6, have escape depths of only a few nanometres, hence Auger spectroscopy is very surface sensitive.

Of particular importance is the fact that, in contrast to photoelectrons, the energies of Auger electrons ($|E_c - E_n - E'_m|$) are independent of the incident photon energy, although the amount of Auger electrons emitted is directly proportional to the absorption cross-section in the surface region.

2.7.3 Fluorescence or Auger?

Auger-electron emission and x-ray fluorescence are competitive processes [5]. The rate of spontaneous fluorescence, given by the Einstein A-coefficient, is proportional to the third power of the energy difference between the upper and lower state. Hence, for a given atom, K-emission lines are more probable (and therefore more intense) than L-emission (assuming L-radiation is allowed at all), while fluorescence is stronger for heavier atoms, which have a more attractive positive nuclear charge and therefore a larger energy difference separating adjacent shells (see Figure 2.2). Figure 2.21 shows the normalized yields for Auger and fluorescence processes as a function of the atomic number, Z, and defined as $Y = n/N$, whereby n is the number of emitted electrons (Auger) or photons (fluorescence), and N is the number of core-shell ionizations.

In contrast, the probability of an Auger electron being emitted increases with decreasing energy difference between the excited atom and the atom after Auger emission. Hence, LMM events are more likely than KLL events. Also, low atomic-number atoms have higher Auger yields than do heavier atoms. Similarly, high atomic-number elements

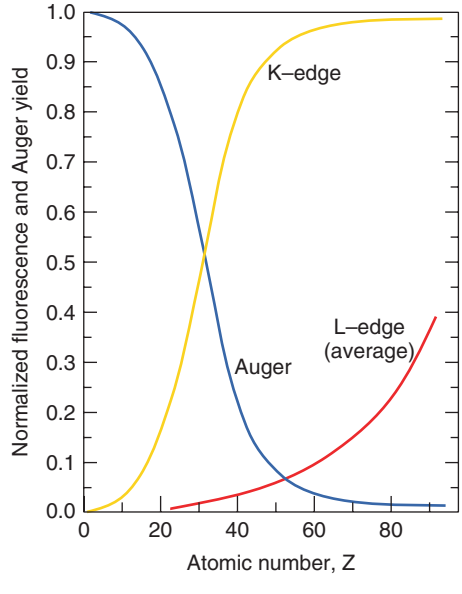

Figure 2.21 *The fluorescence yields of the elements for K- and L-radiation and the Auger yield for K-shell vacancies as a function of atomic number, A.*

have a large positive charge at the nucleus, which binds electrons more tightly, reducing the probability of Auger emission.

2.8 Concluding Remarks

In this chapter, the basic concepts of the interaction of x-rays with matter have been introduced. The most important scatterer of x-rays is the electron. Because x-rays have such high energies compared to the binding energies of all but the most tightly bound core electrons, their interaction with matter is relatively weak – that is why medical radiographic images through the torso of a human can be recorded, not something that could be achieved using visible light!

We have discussed how the phenomena of refraction, reflection and absorption are all intimately related to one another through the complex property of the atomic scattering factor. The knowledge we have gained here will provide the basis for understanding both x-ray hardware, discussed in Chapter 4, and also in fathoming the various experimental methods described in the second half of the book.

References

[1] J. Als-Nielsen, "X-rays and matter – the basic interactions," *C. R. Physique*, 9, pp. 479–486, 2008.

[2] A. H. Compton, "A quantum theory of the scattering of x-rays by light elements," *Phys. Rev*., 21, pp. 0483–0502, 1923.

[3] C. G. Barkla, "The spectra of the fluorescent Rontgen radiations," *Phil. Mag*., 22, pp. 396–412, 1911.

[4] H. G. J. Moseley, "The high-frequency spectra of the elements. Part II.," *Phil. Mag*., 27, pp. 703–713, 1914.

[5] W. Bambynek, C. D. Swift, B. Craseman, H.-U. Frend, P. V. Rao, H. Mark, and R. W. Fink, "X-ray-fluorescence yields, Auger, and Coster-Kronig transition probabilities," *Rev. Mod. Phys*., 44, pp. 716–813, 1972.

3

Synchrotron Physics

3.1 Introduction

For many synchrotron-facility users, a detailed knowledge of the working principles behind the generation of synchrotron radiation may seem of secondary importance. In general, however, both users trying to optimize their experimental setup, and prospective users trying to decide which facility or beamline might be appropriate for a planned experiment, will benefit greatly from having a working knowledge of how synchrotrons perform. Importantly, the demand for access to synchrotron radiation in modern facilities far outstrips availability, hence a submitted proposal which might otherwise be ranked positively by a Review Committee set up to decide which users will obtain beamtime, will be in danger of summary rejection if the foreseen experiment demands operational parameters outside the capabilities of the facility in question. Understanding something of synchrotron physics is therefore advantageous.

In this chapter, I will convey the most important aspects of synchrotron physics, without bogging down the reader with unnecessary mathematical derivations. This should provide essential rules of thumb and simple equations to determine what any given facility has to offer a potential user. For a deeper understanding of the myriad details of 'machine physics', as it is also known, the reader is referred to excellent primers by Herman Winick [1], Helmut Wiedemann [2] and Philip Duke [3].

3.2 Overview

Third-generation synchrotrons are characterized by their ability to produce highly parallel and narrow beams of x-rays with high intensity [4]. These properties can be expressed in a single quantity called 'brilliance', described in detail below. Another important property is the energy of the electrons within the storage ring which generate the x-radiation, which influences the range of photon energies that any one facility can practically cover. The most relevant parameters of a selection of third-generation facilities are listed in Table 3.1.

An Introduction to Synchrotron Radiation: Techniques and Applications, First Edition. Philip Willmott.

Table 3.1 *Important properties of selected third-generation synchrotrons, listed in order of their storage-ring energy.*

Facility, Country	Storage-ring energy [GeV]	Current [mA]	Circumference [m]	Emittance (x, y) [nm rad × pm rad]	Brilliance [ph/s/mrad^2/mm^2/0.1% BW]	Top-up
BESSY II, Germany	1.7	100	240	6 × 100	5×10^{18}	n
ALS, US	1.9	400	198	6.8 × 8	3×10^{18}	n
Elettra, Italy	2–2.4	320	260	7 × 70	10^{19}	n
SLS, Switzerland	2.4	400	288	5 × 2.8	4×10^{19}	y
Anka, Germany	2.5	200	110	50 × 400	10^{18}	n
Soleil, France	2.75	500	354	3.7 × 11	10^{20}	y
Diamond, England	3.0	300	562	2.7 × 27	3×10^{20}	y
ESRF, France*	6.0	300	846	3.8 × 10	8×10^{20}	y
APS, US	7.0	100	1104	3.0 × 25	8×10^{19}	y
SPring8, Japan	8.0	100	1436	2.8 × 6	2×10^{21}	y
Petra-III, Germany	6.0	100	2304	1.0 × 10	2×10^{21}	y
MAX-IV, Sweden†	3.0	500	528	0.17 × 9	2.2×10^{21}	y
NSLS-II, USA†	3.0	500	792	0.6 × 8	3×10^{21}	y

*After upgrade 2010.

†Design/construction phase in 2010.

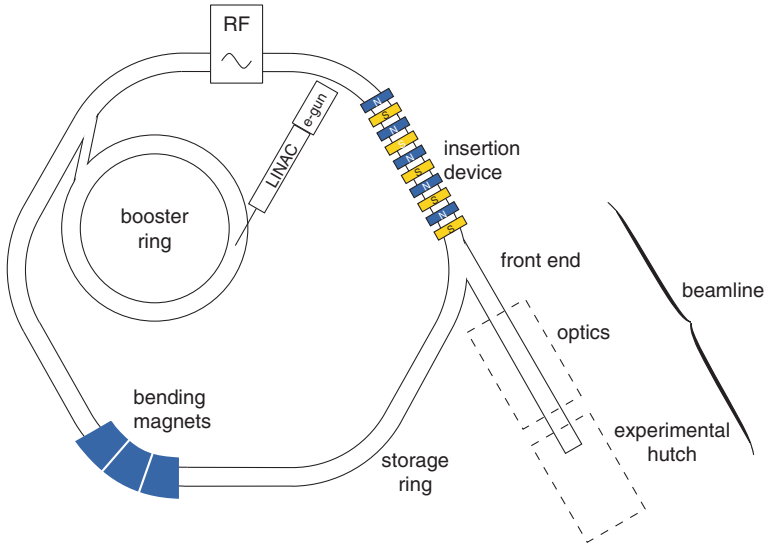

Figure 3.1 *A schematic of the most important components of a modern synchrotron source. Electrons from a source (e.g. a heated filament in an electron gun) are accelerated in a linear accelerator (linac) into an evacuated booster ring, where they undergo further acceleration. They are then injected into the so-called storage ring. There, they are maintained in a closed path using bending magnets at arc sections. The beamlines use the radiation emitted from insertion devices (wigglers or undulators) and from the bending magnets, and therefore are positioned downstream, on the axis of emission. The energy lost by radiation of synchrotron light by the electrons is replenished by a radio frequency (RF) supply.*

A synchrotron consists of five main components (see Figure 3.1). These are

1. A source of electrons (normally generated by thermionic emission from a hot filament) in an electron gun. The electrons are accelerated using a linear accelerator (linac) to about 100 MeV. A regular supply of electrons is required, as they are always being lost in the machine, due to collisions with residual gas particles in the storage ring.
2. A booster ring into which the electrons are injected from the linac and further accelerated. They may either be accelerated to the energy of the electrons in the main storage ring, or (less commonly, especially for modern facilities) to a somewhat lower energy. They are then periodically injected into the storage ring, so that the specified storage ring current is maintained. This is traditionally performed when the storage ring current drops to about $1 - 1/e \approx 70\%$ of the initial current. If the energy to which the electrons are accelerated is not that of the storage ring, then experiments at the beamlines must be suspended during injection, measured in minutes. Even if the electrons attain the storage ring energy in the booster ring, this normally also requires 'downtime' at the beamlines, as the insertion devices are opened (because their magnetic fields must be minimized) as otherwise the perturbation to the electrons' orbit is too strong during injection, resulting in a spontaneous 'beam dump'.

   There have been two recent important developments in synchrotron technology. Firstly, booster-ring diameters have been increased in order to 'gently' accelerate the

electrons and minimize their emittance (described below) in this process. This innovation is optimized for the second development, namely so-called 'top-up' operation. In top-up, the booster ring operates at the same energy as the storage ring (nothing new there), but importantly, the stored current in the ring is quasi-continuously 'topped up' by small injections of electrons, allowing uninterrupted user operation. Typically, the storage-ring current is allowed to drop by approximately 1 mA, or less than a percent, before it is topped-up. For typical machine losses, this means injection every two minutes. Top-up injection greatly enhances stable operation at the beamlines, because it provides a constant heat load on x-ray optical components, which, once optimized, are not subsequently distorted by a change in the intensity of the incident synchrotron radiation.

3. The storage ring contains the electrons and maintains them on a closed path by the use of an array of magnets, commonly referred to as the 'magnet lattice' of the ring. The magnets are most commonly of three types: dipole- or bending magnets cause the electrons to change their path and thereby follow a closed path; quadrupole magnets are used to focus the electron beam and compensate for Coulomb repulsion between the electrons; and sextupole magnets correct for chromatic aberrations that arise from the focusing by the quadrupoles (see Figure 3.2).

   The electrons have kinetic energies measured in GeV,[1] and their velocities are highly relativistic, that is, only *very* marginally less than the velocity of light.

   The ring is a structure consisting of arced sections containing bending magnets, and straight sections used for insertion devices (IDs), which generate the most intense synchrotron radiation. The bending magnets (BMs) used to deflect the electrons round the arced sections that connect the straight sections are also often used to provide bending-magnet radiation – although their brilliance is significantly lower than that produced by IDs, even monochromated BM-radiation is still orders of magnitude more intense than that which can be provided by laboratory-based sources.
4. Energy is lost by the electrons, due to emission of synchrotron radiation. This must be replenished if the electrons are not to spiral into the inner walls of the storage ring and be lost. This is achieved by a radio frequency (RF) supply, which supplies the electrons just the right amount of extra energy every time they pass through it.
5. The beamlines run off tangentially to the storage ring, along the axes of the insertion devices and tangentially at bending magnets. The first section of a beamline, referred to as the 'front end', has several functions and safety features – it isolates the beamline vacuum from the storage ring vacuum; it monitors the position of the photon beam; it defines the angular acceptance of the synchrotron radiation via an aperture; it blocks, when required, the x-ray and Bremsstrahlung radiation during access to the next section containing the optics and experimental hutches; and it filters out, if necessary, the low-energy tail of the synchrotron radiation spectrum, which is strongly absorbed by matter and can damage optical components. The beam is then normally focused

[1] Among other parameters, storage rings are characterized by the electrons' kinetic energy, as in 'the 6 GeV storage ring at the ESRF'. This should not be confused with the energy of the photons produced by deflecting these electrons in a magnetic field, that is, the synchrotron radiation, which is several orders of magnitude smaller. The highest-energy synchrotron storage ring presently in existence is at SPring8, which stores electrons at 8 GeV, hence its name.

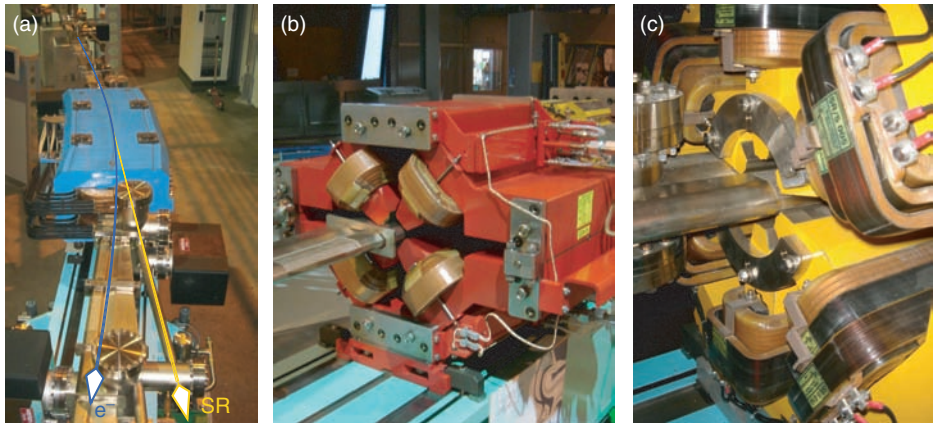

Figure 3.2 *Elements of the magnet lattice. (a) bending- or dipole magnets force the electrons to execute a curved path, thereby emitting synchrotron radiation. (b) A focusing quadrupole magnet and (c) a correcting sextupole magnet.*

and/or monochromated in the optics hutch, before it enters the experimental hutch. For those beamlines generating high-energy x-rays, the hutches are shielded using lead-lined, thick concrete walls, to protect users from not only x-rays, but also from gamma rays and high-energy neutrons, which can be produced in the storage ring when the relativistic electrons collide with stray gas particles. The maximum energy of such radiation is equal to the storage ring energy, of the order of a few GeV. Such gamma-rays penetrate very deeply in matter, hence effective radiation shields may consist of tens of centimetres of lead blocks. Experiments in the hutch are therefore performed remotely, from outside the radiation area.

In this chapter, we will consider the generation of synchrotron radiation in more detail. Beamline components are dealt with separately in Chapter 4. To begin with, we should understand some aspects of electrons moving at relativistic velocities.

3.3 Radiation From Relativistic Electrons

The radiation emitted by accelerated electrons moving at a small fraction of the speed of light has a dipole distribution, as shown earlier in Figure 2.5. In contrast, radiation from a relativistic electron moving along a circular arc is like a sweeping searchlight; a narrow beam of light radiates in the same direction as the electron motion (see Figure 3.3). The features of the radiation depend on two parameters: the angular frequency ω_0 (that is, the number of radians turned by the electrons per second), and on the energy of the electrons in the storage ring, given by the dimensionless parameter known as the Lorentz factor

$$\gamma = \frac{\mathcal{E}}{mc^2}, \tag{3.1}$$

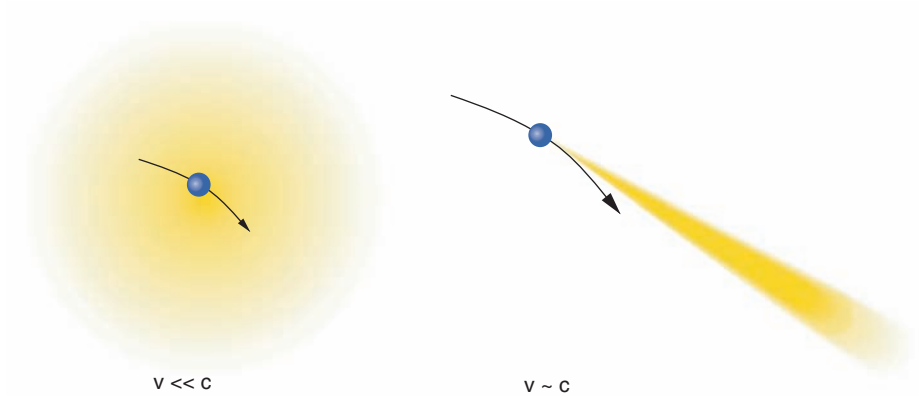

Figure 3.3 *Electromagnetic radiation emitted by acceleration of a charged particle travelling at velocities that are small compared to the speed of light is isotropic. In contrast, relativistically moving charged particles emit radiation in a pencil-like beam in the direction in which they are travelling.*

where mc^2 is the rest mass energy of the electron, equal to 511 keV. This can be re-expressed simply in terms of the storage ring energy in GeV as

$$\gamma = 1957\,\mathcal{E}[\text{GeV}]. \tag{3.2}$$

The natural opening angle (or divergence) of the narrow radiation cone in the plane perpendicular to the storage ring is $\theta \sim \gamma^{-1}$ which, for typical storage ring energies of 1–8 GeV, is equal to 0.5–0.06 milliradians (mrad), respectively (1 mrad is approximately equal to 0.057 °).

In calculating the angular frequency, it is important to remember that only a part of the storage ring path is curved. Hence, a storage ring with a 300 m circumference, of which 50 m are straight sections, has an angular frequency of $2\pi \times c/(300 - 50) = 7.53\,\text{mrad}\,\text{s}^{-1}$.

The emitted bending-magnet spectrum from the orbiting electrons is very broad (see Figure 1.3). However, for photon frequencies higher than about $\gamma^3\,\omega_0$, the spectral intensity falls off sharply. The bending-magnet radiation brilliance spectrum for a 2.4 GeV storage ring running at 400 mA, containing bending magnets, each with a magnetic field strength of 1.41 T, is shown in Figure 3.4.

As mentioned above, the radiation emitted by relativistically moving electrons points along the tangent of their path. This relativistic effect is qualitatively explained here. Consider Figure 3.5. The velocity of the electron is very close to c, so an observer looking at an electron that is turning into his line of sight sees an enormous relativistic Doppler shift, as the distance between radiation wavefronts is hugely relativistically compressed (the electrons are moving almost as fast as the wavefronts).

At other positions on the orbit, for which the angle θ between the observer's line of sight and the electron velocity vector is larger than $1/\gamma$, the relativistic compression of the wavefront quickly falls off. The on-axis rapid and large Doppler shift is synonymous with a larger acceleration. This means that in the observer's frame of reference, radiation is

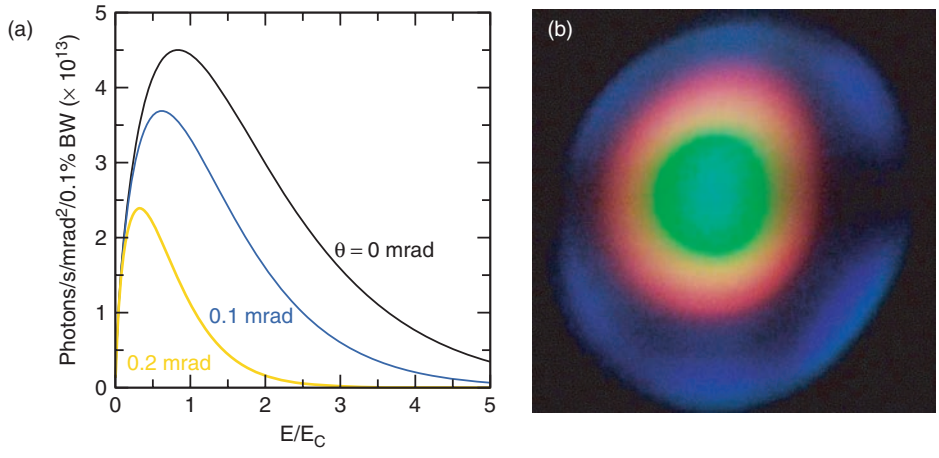

Figure 3.4 *The spatial distribution of synchrotron radiation. (a) The spectrum from a 1.41 T bending magnet in a 2.4 GeV storage ring. The abscissa is in dimensionless units of the photon energy E divided by the characteristic, or critical, energy $E_c = \hbar\omega_c$, defined in the text. The weight of the spectrum changes according to the angular position above or below the electrons' orbital plane. On axis, there is a larger contribution from high-energy photons compared to that in the spectra off axis. (b) This is vividly demonstrated in the image taken on-axis of the visible/UV spectrum of the wiggler at the Femto Project of the Swiss Light Source. Courtesy Gerhard Ingold, Paul Scherrer Institut.*

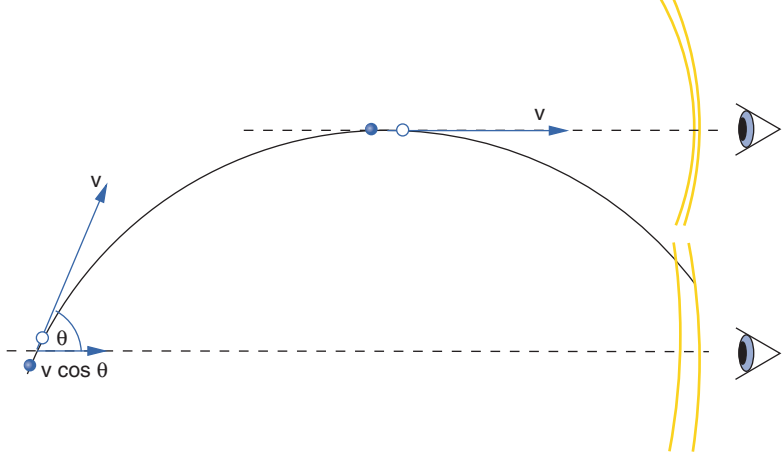

Figure 3.5 *The rapid and intense compression of the wavefronts of the radiation emitted from electrons moving along the line of sight due to a blue Doppler shift means the radiation appears to the observer to be far more intense here than at other positions on the orbit.*

only significantly observed along the observation axis, as emission of radiation requires the acceleration of the electron. Hence, the less pronounced Doppler effect observed slightly off-axis means that here, synchrotron radiation has its maximum intensity at lower energies compared to that of on-axis radiation (shown in Figure 3.4).

3.3.1 Magnetic Deflection Fields

According to special relativity, the kinetic energy of any particle of rest mass m travelling at a velocity v is given by

$$\begin{aligned} \mathcal{E} &= \frac{mc^2}{\sqrt{1-(v/c)^2}} \\ &= \frac{mc^2}{\sqrt{1-\beta^2}} \\ \rightsquigarrow \gamma &= \frac{1}{\sqrt{1-\beta^2}}, \end{aligned} \tag{3.3}$$

where $\beta = v/c$ and we have used our definition of γ given in Equation (3.1). As the particle approaches the velocity of light, it gains mass by a factor of γ. As it would have infinite mass exactly at $v = c$, it can only asymptotically approach this limit.

We manipulate Equation (3.3) such that

$$\beta = \left(1 - \frac{1}{\gamma^2}\right)^{1/2}. \tag{3.4}$$

For typical γ values of 10^4, $1/\gamma^2$ is very small, so we can ignore all but the first two terms in the Taylor expansion of Equation (3.4) and thereby obtain

$$\beta \approx 1 - \frac{1}{2\gamma^2}. \tag{3.5}$$

How are the electrons in the storage ring deflected into a closed orbit? A magnetic field[2] will induce a Lorentz force on a charged particle perpendicular to both the motion of that particle and the magnetic field vector (It is proportional to the cross-product of **v** and **B**). Classically, we equate the Lorentz force to a centripetal force such that

$$e\,\mathbf{v} \times \mathbf{B} = \frac{mv^2}{\rho}, \tag{3.6}$$

where ρ is the orbital radius of the arc. Let us assume that the magnetic field is perpendicular to the electron path, so that the cross-product $\mathbf{v} \times \mathbf{B}$ is merely $|v|\,|B|$. Here, we are dealing with highly relativistic electrons, and so we must replace m with the relativistic mass γm. We can also to a high degree of accuracy replace v with c. Therefore

$$ecB = \frac{\gamma mc^2}{\rho}$$

[2] Note that the magnitude of a radial electric field required to induce a comparable centripetal force as the magnetic-field-induced Lorentz force would be an impractical $10^9\,\mathrm{V\,m^{-1}}$!

$$\rightsquigarrow \rho = \frac{\gamma mc}{eB}$$
$$= \frac{\mathcal{E}}{ceB}. \tag{3.7}$$

If we express $\mathcal{E}$ in GeV, we can rewrite this equation using practical units so that

$$\rho[\mathrm{m}] = 3.3\,\frac{\mathcal{E}[\mathrm{GeV}]}{B[\mathrm{T}]}. \tag{3.8}$$

As typical magnetic field strengths of bending magnets and insertion devices are of the order of 1 Tesla[3] and storage ring electron energies are normally of the order of a few GeV, the bending radius is typically a few metres.

We now define the characteristic, or critical, frequency ω_c associated with a given synchrotron source in terms of of γ and ω_0, such that

$$\omega_c = \frac{3}{2}\gamma^3\,\omega_0 = \frac{3}{2}\gamma^3\,\frac{c}{\rho}. \tag{3.9}$$

The critical energy, $E_c = \hbar\omega_c$ exactly divides the total emitted power from a bending magnet in half. As ρ is proportional to $\mathcal{E}$ (Equation (3.7)), we can re-express the above equation in practical units

$$\hbar\omega_c[\mathrm{keV}] = 0.665\,\mathcal{E}^2[\mathrm{GeV}]\,B[\mathrm{T}]. \tag{3.10}$$

3.3.2 Radiated Power Loss in Synchrotrons

The radiated power from a circular arc of length L is stated without derivation as

$$P[\mathrm{kW}] = 1.266\,\mathcal{E}^2[\mathrm{GeV}]\,B^2[\mathrm{T}]\,L[\mathrm{m}]\,I[\mathrm{A}]. \tag{3.11}$$

For a storage ring with arc sections of equal radius, the power loss of the entire storage ring due to bending-magnet radiation only is calculated by setting $L = 2\pi\rho$. The total loss in third-generation synchrotrons is more, because of the additional loss generated by the insertion devices installed in the straight sections of the storage ring.

The power per unit solid angle per 0.1% bandwidth can be calculated by weighting the curves of Figure 3.4, which show the brilliance in terms of the rate of emission of photons, by the photon energy, as shown in Figure 3.6. It is immediately obvious that, on-axis, the area under the curve is most certainly not divided equally left and right of $\mathcal{E}_c$. However, the maximum of radiation off-axis shifts more and more to lower energies with off-axis angle. Importantly, the weighted contribution to the total power for a spectrum at a given off-axis angle is proportional to the circumference of the associated cone, and hence to the subtended angle. Once these factors are taken into account, it can be shown that $\mathcal{E}_c$ indeed does divide the total emitted power from a bending magnet in half.

[3] As a comparison, the Earth's magnetic field is strongest around the magnetic poles in northern Canada and southern Australia, and in part of Siberia, at a little above 60 μT (0.6 Gauss), while a common-or-garden kitchen magnet may have a field strength at its surface of 0.03 T.

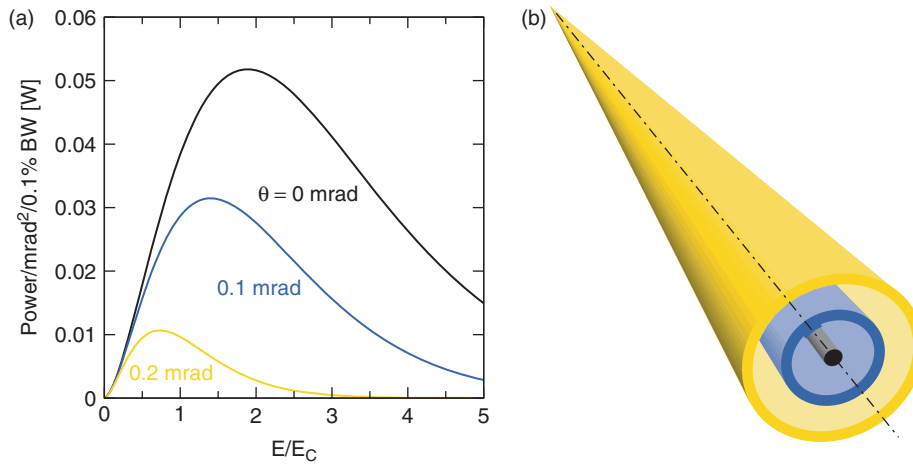

Figure 3.6 *Bending-magnet power spectra. (a) The power spectra per mrad^2 and 0.1% bandwidth of bending-magnet radiation as a function of the dimensionless parameter E/E_c for different angular positions out of the plane of the electron orbit. The power spectrum falls off and the positions of the spectrum maxima move to lower energies with increasing angle from the axis. (b) The contribution to the integrated power spectrum of a cone associated with a certain off-axis angle is proportional to that angle. The spectrum associated with the central (black) cone has a higher power density and is shifted to higher photon energies than that of the blue cone, and still more than that from the yellow cone. However, the weighted contribution of any cone to the integrated power is proportional to its subtended angle.*

3.4 Radio Frequency Power Supply and Bunching

The kinetic energy of the electrons dissipated due to emission of radiation at the bending magnets and insertion devices must be replenished before they spiral into the inner wall of the storage ring. This is achieved by giving them a small boost at every turn as they pass through a radio-frequency (RF) cavity (normally powered by a so-called klystron, or similar device, Figure 3.7(a)). How does one ensure that the electrons enter the cavity at the correct phase of the RF voltage cycle to obtain the correct boost in energy?

Well, one doesn't. Either the electron is in the right environment and it survives, or it isn't and it doesn't. In other words, the manner in which this is done is in fact an automatic, self-correcting process and, as we shall see, also results in the electrons settling into a series of small packets, or 'bunches'.

First, what is the drop in energy of the electrons for each cycle around the ring? This is simply the total power loss of the storage ring divided by the storage-ring current, and is equal to approximately 0.2 to 1 MeV, or of the order of 0.05% of the nominal electron energy of approximately 2 GeV. The RF cavity must therefore have a peak voltage in excess of this. Typical klystrons at synchrotrons have peak voltages of about 3 MV.

Consider Figure 3.7(b). On average, the electrons require a certain boost in order to maintain them on a stable path, given by an amount eV_{ref}. If an electron dissipates more than this amount of energy in the following turn around the storage ring, it will enter the RF cavity somewhat later at point A, and will experience a larger acceleration than if it

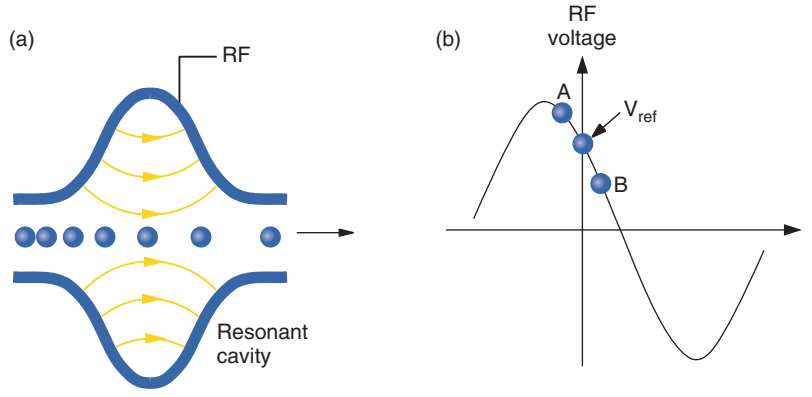

Figure 3.7 *Replenishing the energy of electrons in a storage ring. (a) Electrons entering the resonant cavity at the right moment in its voltage cycle are accelerated by the electric field within the cavity generated by the klystron source. (b) 'Slow' electrons entering the RF cavity at* A *will be given an extra boost, while 'fast' electrons (at* B*) receive less energy.*

were at the reference voltage. Likewise, if the electron is too fast (i.e., it has dissipated less energy than eV_{ref} in the last turn round the storage ring), it will receive less of a boost.

Electrons outside this range above and below the reference voltage will not gain the correct energy and will be lost to the system. The electrons therefore quickly bunch into packets associated with each cycle of the RF cavity. For example, electrons in a storage ring with a 300 m circumference complete one cycle every microsecond. A 500 MHz RF supply will therefore generate 500 bunches in the storage ring, separated from one another by 60 cm (or, in time, by 2 ns). Typical bunch lengths are measured in mm (or of the order of 10 ps). This ratio of bunch length to bunch separation of the order of 0.005 demonstrates how tightly defined are the conditions for an electron to survive a transit through the RF booster.

The short length of the bunches allows users to exploit the time structure of synchrotron radiation down to well below the nanosecond timescale. These time-resolved experiments are becoming increasingly important in modern synchrotron facilities, and several examples will be given in later chapters.

To perform time-resolved studies, storage rings can either be run in single-bunch mode, where only one bunch of electrons exists in the ring; in multibunch mode, the normal mode of operation; or indeed in a combination of the above, in a so-called 'hybrid mode' (see Figure 3.8). This last mode has the advantage that 'normal' users can operate at reasonable average photon fluxes (in contrast to when the single-bunch mode is running), while those beamlines wanting to perform time-resolved experiments can receive a trigger from the storage ring control centre, in order to synchronize their experiments with the isolated bunch.

Coulomb repulsion between electrons in a given bunch, which will defocus the electron beam in the ring, is ameliorated by relativistic effects. In the direction of movement, relativistic length-contraction means that the distances 'seen' by electrons moving together in the same frame of reference are γ times larger than in the laboratory frame of reference, and so the Coulomb forces are $1/\gamma^2$ weaker. The so-called 'beam stiffness'

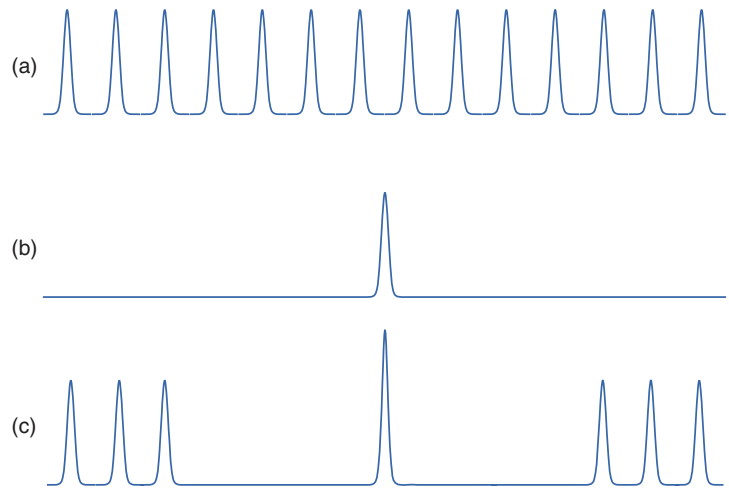

Figure 3.8 *Different filling modes in a synchrotron storage ring. (a) Normally, the ring is filled with bunches of electrons equally spaced from one another. (b) Some experiments may require that only a single bunch of electrons is in the storage ring, or (c) a more complex distribution of modes such as the hybrid mode, where some bunches are dropped to leave a region in which a bunch is isolated.*

therefore increases with energy. Nonetheless, the beam must be manipulated and nudged into the right dimensions and positions using quadrupole and hexapole magnets. The precision with which this is achieved fundamentally influences the emittance of the facility, described in the following section.

3.5 Photon Beam Properties

3.5.1 Flux and Brilliance

Third-generation storage rings optimize the photon flux and brilliance required for modern experimentation. Flux and brilliance indicate the quality of a synchrotron beamline facility. The spectral flux is defined as the number of photons per second per unit bandwidth (normally 0.1%) passing through a defined area, and is the appropriate measure for experiments that use the entire, unfocused x-ray beam. Brilliance essentially states how the flux is distributed in space and angular range (or, for the *cognoscenti*, in phase-space), and determines the smallest spot onto which an x-ray beam can in principle be focused. It is defined as

$$\text{Brilliance} = \frac{\text{photons/second}}{(\text{mrad})^2\,(\text{mm}^2\ \text{source area})\,(0.1\%\ \text{bandwidth})}, \tag{3.12}$$

and is therefore the flux per unit source area and unit solid angle. If the x-ray optics of a beamline are less than perfect (which is always the case to a greater or lesser extent), the brilliance at the experimental station will fall short of theory. It is the goal of any self-respecting beamline engineer to approach as best as possible this upper limit.

The brilliance of a third generation undulator is about 10^{20} photons/s/mrad^2/mm^2/0.1% bandwidth (see Figure 1.3), some ten orders of magnitude higher than that of a rotating anode Cu Kα line and only two or three orders of magnitude lower than high-quality visible laser sources. This is, above all, the reason why synchrotrons have become such important research tools.

Why do synchrotrons have such a high brilliance? Firstly, the size of the radiation source is primarily determined by the transverse size of the electron beam. This is determined below for undulators (Equation (3.33)) and yields values of the order of microns. Secondly, synchrotrons emit an enormous amount of light. The flux emitted by an electron can be shown to be proportional to the square of the electron's acceleration. The centripetal acceleration exerted on electrons in a storage ring is proportional to γ^2, hence the flux is proportional to γ^4. High-energy synchrotron rings therefore have correspondingly high brilliances.

3.5.2 Emittance

From Equation (3.12), it can be seen that the brilliance is inversely proportional to both the source size and the x-ray beam divergence. The product of the linear source size and the beam divergence in the same plane is known as the emittance. One would therefore like to obtain as low an emittance as possible, in which an exceedingly small source size emits x-rays that are almost perfectly parallel. This begs the question as to whether one can in principle make the emittance arbitrarily small, thereby enabling one to increase the brilliance without limit. For a given synchrotron storage ring, the emittance in a given transverse direction is a constant, though this constant is different for different facilities, being determined primarily by the degree of sophistication of the magnet lattice. The goal then is to make this constant as small as possible.

However, there is indeed a fundamental lower limit to the emittance, given by Heisenberg's uncertainty principle, in other words, by the fact that the electron recoils as a result of emitting a photon. It turns out that this so-called diffraction-limited emittance is

$$\epsilon_{\min} = \frac{\lambda}{4\pi}. \tag{3.13}$$

Hence, for 1 Å radiation, the diffraction-limited photon-beam emittance is some 8 pm rad. In practice, the emittance is always higher than this, and is a convolution of the emittance of the electron beam circulating within the storage ring, and that of the photon beam generated by a single electron passing through the source.

Consider Figure 3.9. We define here the two emittances within the plane perpendicular to the electron motion (the z-direction), one in a radial direction x (i.e. in the plane of the orbit), and the other, y, perpendicular to this, at 90° to the orbital plane, that is

$$\epsilon_x = \sigma_x \, \sigma_x', \tag{3.14}$$

$$\epsilon_y = \sigma_y \, \sigma_y'. \tag{3.15}$$

Here σ_x and σ_y are the standard deviations of Gaussians describing the beam profile in the x- and y-directions, that is

$$I = I_0 \exp(-x^2/2\sigma_x^2) \exp(-y^2/2\sigma_y^2). \tag{3.16}$$

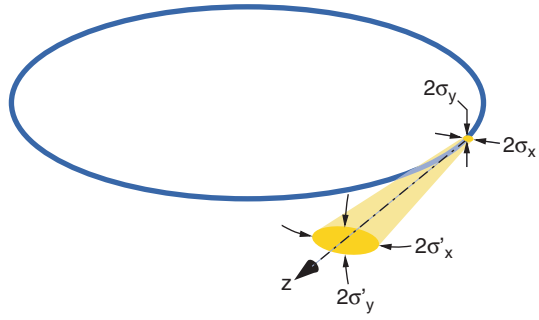

Figure 3.9 *The four parameters defining the emittances ϵ_x and ϵ_y of a storage ring.*

Hence the relationship between $\sigma_{x,y}$ and the full-widths at half-maximum ($\mathrm{FWHM}_{x,y}$) is

$$\mathrm{FWHM}_{x,y} = \sqrt{8 \ln 2}\, \sigma_{x,y}. \tag{3.17}$$

σ'_x and σ'_y are the corresponding angular divergences. The emittance in the x-direction is inherently larger because the Lorentz forces from the magnet lattice act in this direction and there is a spread of the electrons' energies.

If an electron deviates from the ideal reference orbit in the storage ring and possesses a nonzero transverse momentum, it cannot be maintained in a closed orbit using dipole bending magnets alone. In order to refocus the electrons and bring them back towards the ideal orbit, pairs of alternating vertical and horizontal focusing quadrupole magnets are employed. Hence, the beam shape, given by σ_x and σ_y, varies along the magnet lattice, although the emittances do not. A convenient measure of these variations around the storage ring is the so-called beta function, given by

$$\beta_x = \sigma_x / \sigma'_x, \tag{3.18}$$

$$\beta_y = \sigma_y / \sigma'_y. \tag{3.19}$$

Inserting Equations (3.18) and (3.19) into Equations (3.14) and (3.15), respectively, we obtain

$$\sigma_x = \sqrt{\epsilon_x \beta_x}, \tag{3.20}$$

$$\sigma_y = \sqrt{\epsilon_y \beta_y}. \tag{3.21}$$

Focusing the electron beam with quadrupole-magnet pairs results in the transverse momentum being minimized, hence σ'_x is small, and $\sigma_x = \epsilon_x / \sigma'_x$ and β_x are correspondingly large. The opposite is true at the sources (bending magnets and insertion devices, see below). An example of how the beta functions change is shown in Figure 3.10.

Coupling of the horizontal emittance to the vertical emittance can occur if the magnetic dipole fields contain some horizontal component. This induces a vertical dispersion of the electrons. A particularly common source of this sort of coupling is a misalignment of the focusing quadrupole magnets. This coupling should be minimized. At the Swiss Light Source, this cross-coupling of the electron-beam emittances is exceedingly low at

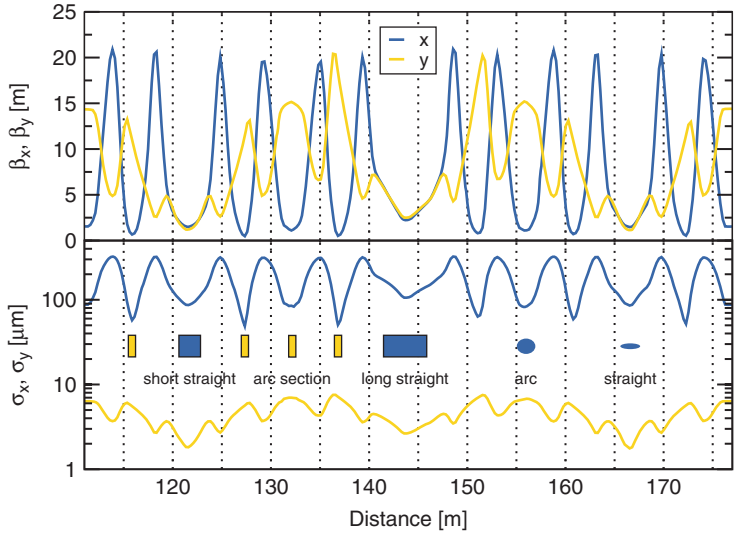

Figure 3.10 *Variations in the beta function and the electron-beam profile around a storage ring. Top: the beta functions in the* x- *and* y*-directions along a section of the magnet lattice at the Swiss Light Source. Bottom: the corresponding beam dimensions, calculated from* $\sigma = \sqrt{\epsilon\beta}$*. The insertion devices (blue blocks) are in the straight sections. Because the vacuum pipes of the storage rings need to be very narrow here to fit inside the ID gaps, the quadrupole magnets (not shown) refocus the beam in these regions to be as flat as possible (shown by the ellipse). In contrast, in the arc sections containing the bending magnets (yellow blocks), the beam is reshaped to be much rounder (shown by the circle). Courtesy Michael Böge, Paul Scherrer Institut.*

5×10^{-4}. The emittances are $\epsilon_y = 2.8$ pm rad and $\epsilon_x \leq 7$ nm rad. Note that this vertical emittance is lower than the theoretical limit given by Equation (3.13). This is not a contradiction, as the former refers to the electron-beam emittance, and the latter to the photon-beam emittance.

3.5.3 Coherence

Because no x-ray source has an infinitely narrow bandwidth, the different frequencies within the bandwidth of the beam will sooner or later get out of phase with one another. The time required for the phase between two sinusoidal waves differing in frequency by an amount $\Delta\nu$ to change by 2π radians (a full cycle) is simply $1/\Delta\nu$, and this is called the longitudinal coherence time, $\Delta\tau_c^{(l)}$. During that time, the waves have travelled in vacuum a distance $l_c^{(l)} = c\Delta\tau_c^{(l)}$, referred to as the longitudinal (or temporal) coherence length of the light source (see Figure 3.11(a)).

$$l_c^{(l)} = \frac{\lambda^2}{\Delta\lambda}. \tag{3.22}$$

The longitudinal coherence after a monochromator is usually determined by the quality of the crystal or grating used in the monochromator, which defines $\lambda/\Delta\lambda$. In crystal

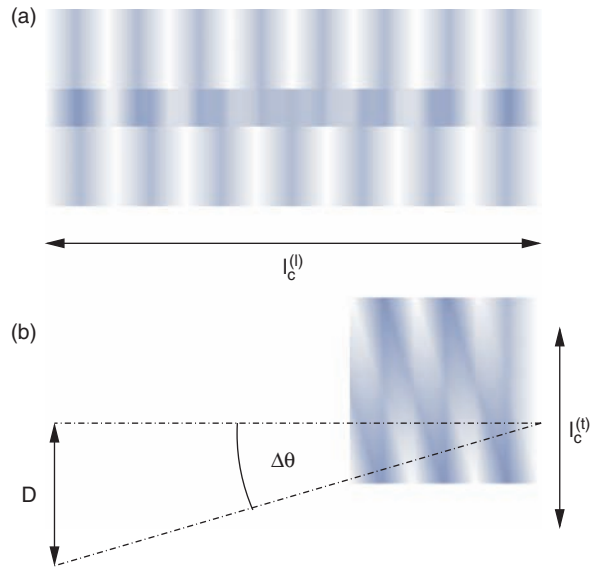

Figure 3.11 *Beam coherence. (a) The longitudinal, or temporal, coherence length is determined by the monochromacity of the source, while (b) the transverse (or spatial) coherence length depends on the beam divergence and source size.*

monochromators, this can exceed 10^4 and the longitudinal coherence length can be several microns, depending on the x-ray wavelength.

We can also define a *transverse coherence length* $l_c^{(t)}$ (also called the spatial coherence length) caused by the interference of waves of the same wavelength but with slightly different directions of propagation. This arises because all sources have a finite size, with a non-zero divergence $\Delta\theta$, as shown in Figure 3.11(b). In this case

$$l_c^{(t)} = \lambda/2\Delta\theta = \lambda R/2D, \tag{3.23}$$

where D is the lateral extent of the finite source and R is the distance from the source to the observation point D. Imperfections in optical components (such as the slope errors of x-ray mirrors, see Section 4.3.1) can further contribute significantly to the effective spatial coherence and blow this up far beyond the theoretical limit given by the divergence of the source.

3.5.4 Polarization of Synchrotron Radiation

The polarization of the x-rays emerging from a storage ring depends on the line of sight, illustrated in Figure 3.12. Viewed in the electrons' orbital plane, the polarization of the radiation is linear in that plane (assuming one has not installed a helical undulator, see below), as the electrons appear to oscillate in the horizontal plane. If the observer is above the plane of the storage ring, the electrons will appear to execute an elliptical orbit in a clockwise direction. This angular momentum from the observer's frame of reference is transferred to the emitted photons, which are left-circularly polarized (LCP).

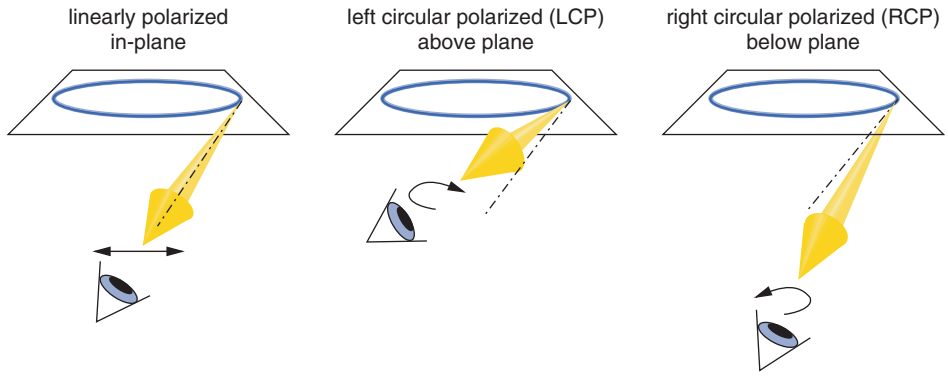

Figure 3.12 *The polarization of synchrotron light depends on the position of the observer relative to the plane of the electron orbit.*

The electric-field vector of the x-rays also rotates clockwise around the direction of propagation as viewed back down the beam from the observer. The angular momentum of the rotating electric field of the LCP photons is $-\hbar$, and is described by a vector of that magnitude pointing opposite to the direction of propagation. The opposite is true for an observer below the electrons' orbital plane – the polarization is right-circular (RCP) and the photons carry an angular momentum $+\hbar$. Note that the angular momentum of a photon can be transferred to systems which absorb that photon. This is an important fact for understanding (among other things) dichroism in spectroscopy, as we shall see in Section 6.8.

3.6 Bending Magnets and Superbends

The primary purpose of bending magnets is to circulate the electron beam in the storage ring on a closed path. Bending magnets have typical magnetic field strengths of around a Tesla. They produce bending-magnet radiation in a flattened cone with a fan angle equal to the angular change of the path of the electrons (plus the small extra amount due to the photon-beam divergence, equal to $1/\gamma$). Because of the relatively large subtended angle of bending-magnet radiation, measured in degrees (see Figure 3.13), it is possible to position more than one so-called 'bending-magnet beamline' at a single bending magnet using two or more apertures.

The critical energy and emitted power, defined by Equations (3.9) and (3.11), respectively, are determined by the storage-ring energy and the magnetic field strength. The bending-magnet spectra for three different combinations are shown in Figure 3.14. Particularly at low- and medium-energy facilities, the maximum values of both these quantities may not be normally sufficient to extend the spectrum far into the hard x-ray regime. By 'normally', the use of conventional magnetic-field sources is meant. However, if superconducting magnets are employed, one can begin to access harder x-radiation as the critical energy is proportional to the magnetic field strength. In addition, from Equation (3.11), one can see that there is much to be gained in radiative power by increasing the magnetic field strength B.

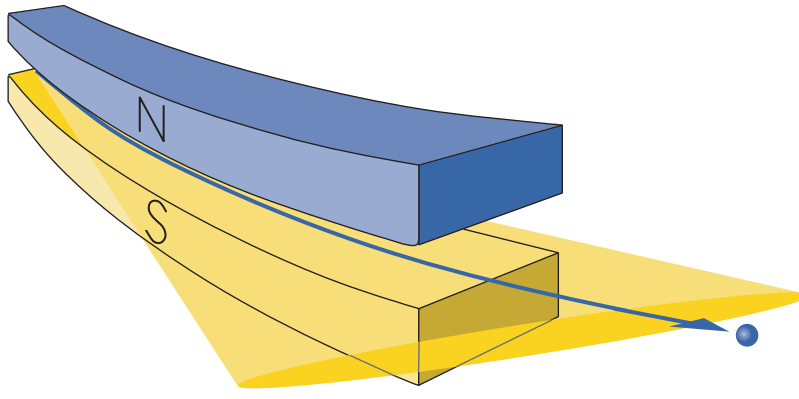

Figure 3.13 *A schematic of the fan of radiation produced by bending magnets.*

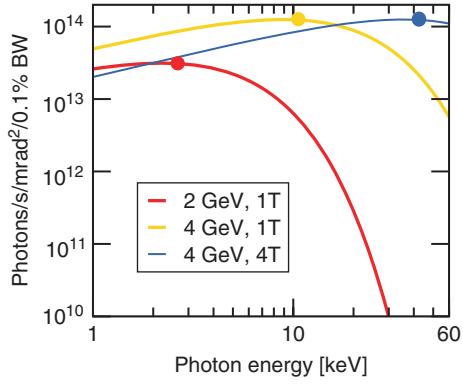

Figure 3.14 *The bending-magnet spectra for horizontally polarized radiation on axis for three different combinations of the storage-ring energy and magnetic-field strength. Also shown as circles for each curve is the critical energy $\hbar\omega_c$, given by Equation (3.10).*

Present technology limits the maximum field strength of permanent magnets to a little more than a Tesla – superconducting electromagnets, operating at cryogenic temperatures can provide field strengths in excess of 5 Tesla. Typical superconducting materials for these applications are based on liquid-helium-cooled niobium alloys. In this manner, the radiated power output (proportional to the square of B) compared to conventional bending magnets can be increased by approximately a factor of 20 or more, while the critical energy can be increased by a factor of four to five.

For certain experiments, a high brilliance and high degree of monochromacity are not needed – for example, in many computed tomography experiments (see Section 7.2), the sample might have to be illuminated by a uniform field of several square millimetres, while little is gained by highly monochromating the beam. In such instances, flux rather than brilliance is at a premium, and bending-magnet beamlines are very competitive. Yet more flux can be obtained by relaxing the bandwidth of wavelengths that are selected by

the monochromator. This can be simply achieved by using multilayer monochromators (see Section 4.3.3), which have bandwidths approximately ten to a hundred times larger than those of crystal monochromators.

3.7 Insertion Devices

Third-generation synchrotrons are characterized by the use of insertion devices (IDs), placed in the straight sections between the bending-magnet arc segments. IDs produce significantly higher fluxes and brilliances than do bending magnets. They operate by forcing the electrons to execute an oscillatory path in the plane of the storage ring, employing a set of magnets that produce magnetic fields that point alternately up and down (see Figure 3.15).

There are two types of ID, which are distinguished from one another by the degree to which the electrons are forced to deviate from a straight path. This seemingly subtle distinction has a fundamental effect on the nature of the radiation, however. For angular excursions larger than the natural opening angle γ^{-1}, the radiation cones from each 'wiggle' do not overlap and therefore the intensities are added and the ID is referred to as a *wiggler*.

For gentler excursions of the order of γ^{-1}, the ID is called an *undulator*. Here, the radiation cones emitted by the electrons as they execute the slalom motion overlap and interfere with one another. In this case, the field *amplitudes* are added vectorially (i.e. including the phase difference from each contribution) and the sum is squared to produce the intensity, which peaks at those wavelengths where interference is constructive (see Figure 3.16).

The maximum angular deviation $\phi_{\max}$ of the electron oscillations in an ID is defined by the dimensionless parameter K, given by

$$K = \phi_{\max}\,\gamma. \tag{3.24}$$

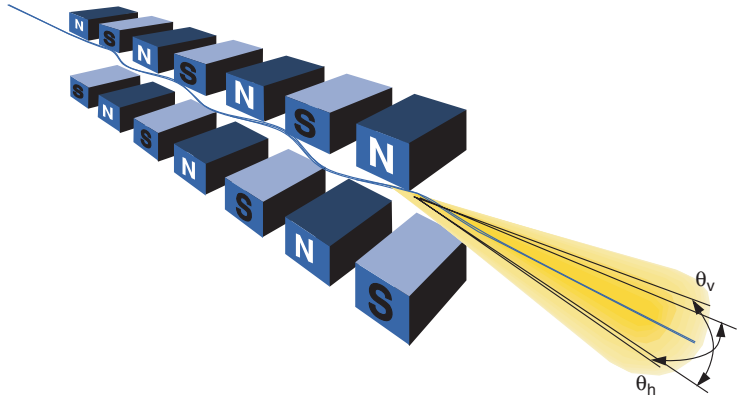

Figure 3.15 *A schematic diagram of an insertion device. The angular spread in radiation is given in the horizontal and vertical plane by θ_h and θ_v, respectively.*

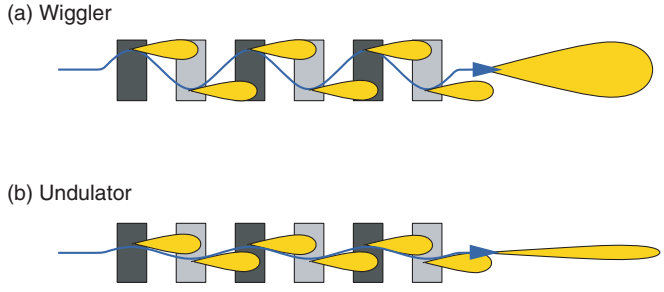

Figure 3.16 *Wigglers and undulators in principle differ only in the size of the excursions from a straight path that they force the electrons to execute. The angular excursions shown here have been greatly exaggerated.*

K can be expressed in terms of the maximum magnetic field B_0 as

$$K = \frac{eB_0}{mck_{u,w}} = 0.934\,\lambda_{u,w}[\text{cm}]\,B_0[\text{T}], \tag{3.25}$$

where λ_u and λ_w are the periods of the oscillations in the undulator or wiggler, respectively, and $k_{u,w} = 2\pi/\lambda_{u,w}$. For a wiggler, K is typically between 10 and 20, while for undulators, K is close to unity. The horizontal spread in radiation is

$$\theta_h = 2K/\gamma. \tag{3.26}$$

So, for example, a wiggler having $K = 10$, operating in a 4 GeV storage ring would have a horizontal divergence of 2.6 mrad (0.15 °).

3.7.1 Wigglers

A wiggler can be thought of as a series of bending magnets that turn the electrons alternately to the left and to the right, as shown in Figure 3.16(a). The maximum angular excursion from the wiggler axis is larger than the natural opening angle of the radiation, γ^{-1}. For each oscillation, the electrons are twice moving parallel to (and in reality also very close to) the wiggler axis. The radiation is therefore enhanced by a factor of $2N$, where N is the total number of periods of the wiggler. The spectrum from a wiggler has the same form as that from a bending magnet. Therefore the formula describing the emitted power is similar to that for a bending magnet, given by Equation (3.11). Because in a bending magnet the magnetic field is a constant, while in a wiggler the field strength oscillates along the ID axis, (B drops to zero exactly between magnet pairs), the square of the average field in a wiggler is $B_0^2/2$, where B_0 is the maximum magnetic field strength, and hence Equation (3.11) is modified to

$$P_w[\text{kW}] = 0.633\,\mathcal{E}^2[\text{GeV}]\,B_0^2[\text{T}]\,L[\text{m}]\,I[\text{A}]. \tag{3.27}$$

The excursions in Figure 3.16 are much exaggerated, and in fact L is to a very good approximation the linear length of the wiggler, which is typically of the order of 2 m. The amount of power radiated by a wiggler increases as the gap between the upper and

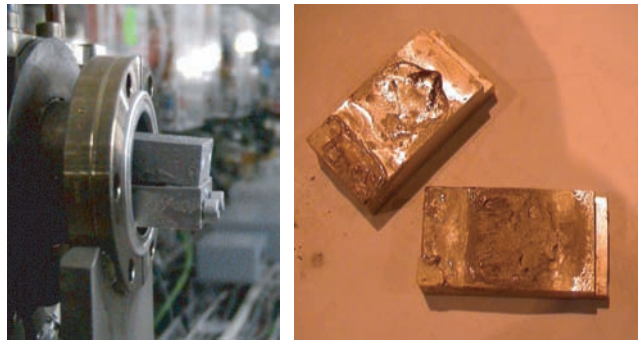

Figure 3.17 *The power of wiggler radiation. These images of a runaway failure of the beam-defining aperture in the front end of a wiggler beamline at the Swiss Light Source demonstrate the power delivered at synchrotron facilities. The aperture consists of two pairs of tungsten blocks, and is (or, rather, was) the first component in the beamline, and therefore 'sees' the full spectrum of radiation from the wiggler. This includes the soft x-ray component, which is very efficiently absorbed by matter, before this dangerous spectral component is removed by a carbon or diamond filter. Failure occurred because the upper block was not held sufficiently firmly in place, and dropped marginally into the beam path. The increased heat load caused the support structure, made of copper, to also become hot, which loosened further and allowed the W-block to sag more ... resulting in a runaway failure. The entire beam became occluded and the upper block became so hot that the two steel M8 screws at the front evaporated, and the upper and lower W-blocks fused together. The absorbed power was over 8 kW. Note that the boiling point of iron is 3134 K, while the melting point of tungsten is 3695 K!*

lower sets of magnets is reduced. At large gap distances, a wiggler will eventually become an undulator, as will become clear below. Wigglers are not designed for this purpose, however, and the power (which is proportional to the square of the magnetic field, which drops off with increasing gap size) becomes too low to be of any practical use.

In the case of a wiggler, where there is no overlap of the radiation emitted by successive dipoles, $\phi_{\max} = K/\gamma$ is equal to the ID's horizontal divergence σ'_x. In the vertical direction the divergence $\sigma'_y = 1/\gamma$ remains unchanged compared to a bending magnet.

Wigglers therefore provide high-power x-ray beams with broad spectral ranges, as dramatically demonstrated in Figure 3.17.

The heat load on optical equipment downstream from a wiggler, such as monochromators, beryllium windows, and x-ray mirrors can be very high, as matter generally absorbs efficiently for x-ray energies below about 5 keV. It is therefore often necessary in the case of beamlines using hard x-rays to install a low-energy absorbing filter in the front end to protect components downstream. This filter should therefore be transparent in the energy region of interest above approximately 5 keV, but absorb the longer wavelength radiation, without being damaged itself. This limits the choice of possible material to one only – carbon. Carbon has a low Z-value, is very robust, and has a very high vaporization temperature. It can be heated in vacuum to 1800 K before its vapour pressure becomes problematic. As graphite or glassy carbon, it has a high emissivity, allowing efficient radiative heat dissipation. In the form of diamond, it has an excellent thermal conductivity. High-pass filters of this nature are discussed in more detail in Chapter 4.

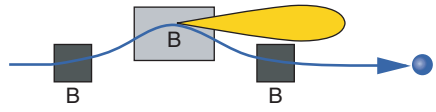

Figure 3.18 *A three-pole wavelength shifter, consisting of one strong and two weak dipoles.*

So-called 'wavelength shifters' (WSFs) are very short wigglers that provide the opportunity to shift the spectral distribution of photons for a given storage ring to higher energies than might otherwise be easily accessible. They normally only consist of either three or five magnetic poles. A three-pole WSF is a wiggler consisting of only two outer, weak, magnet pairs and one central strong (usually superconducting) magnet dipole, which, together, force the electrons to execute a single oscillation (see Figure 3.18). The strong central magnet emits radiation with a high critical frequency (see Equation (3.10)). WSFs are most often employed when high-energy photons are required in a beamline attached to a storage ring that is optimized for lower-energy radiation, and are conceptually very similar to superbends.

3.7.2 Worked Example: The SLS Materials Science Beamline Wiggler

The light source for the Materials Science beamline at the SLS was in operation from April 2001 until October 2010 and is a so-called hybrid minigap wiggler, which combines NdFeB permanent magnet blocks with Permendur pole pieces. These high-performance materials, together with a minimum magnetic gap of 7.5 mm, produce a maximum magnetic field of 1.97 T. The wiggler is 2 m long with 65 poles, ($N = 32.5$ oscillations) and a periodicity $\lambda_w = 61$ mm.

The storage-ring energy is $\mathcal{E} = 2.4$ GeV, and hence $\gamma = 2.4 \times 10^9/511 \times 10^3 = 4697$ and the natural opening angle of the bending-magnet radiation is $\theta_v = 0.213\,\text{mrad} = 0.0122^\circ$. From Equation (3.25), we calculate K to be 11.22. Note that this is a small K value for wigglers, as a result of the unusually small magnetic pole periodicity, and this ID has thus been dubbed a 'wundulator'. With our value for K, we can determine the horizontal angular spread of the wiggler radiation to be $\theta_h = K/\gamma = 2.39$ mrad. The critical energy of the wiggler radiation is calculated using Equation (3.10) to be $E_c = 7.55$ keV Finally, the emitted power given by Equation (3.27) yields $P_w = 11.32$ kW at 400 mA ring current.

In actual fact, the real values of K, E_c, and P_w are 9.2, 7.9 keV and 8.4 kW, respectively. These deviations from the above-calculated values are due to the fact that the magnetic field distribution is not perfectly sinusoidal and that there are stray fields at the two ends of the wiggler. However, the agreement is within about 20%, which is enough to ascertain the quintessential properties of the ID.

3.7.3 Undulators

Undulators differ from bending magnets and wigglers in that their spectra consist of regularly separated, narrow bands of radiation. This is a result of the fact that the radiation fields shown schematically in Figure 3.16(b), overlap and interfere with each other.

Only certain wavelengths interfere constructively, and the undulator spectrum therefore consists of a fundamental frequency plus a series of higher harmonics.

Although the concept of undulators harks back to the 1940s, and the first successful demonstration of undulator radiation was in the early 1950s (albeit producing microwave radiation), the first practical device to operate in the x-ray regime was constructed by Klaus Halbach and coworkers at the Lawrence-Berkeley Laboratory and tested at the SSRL synchrotron at Stanford in 1981. This breakthrough was largely thanks to the development of novel magnetic alloys, such as $SmCo_5$ [5], which allowed the construction of magnet arrays with the required small periodicity and high magnetic-field strength [6].

The basic parameters for undulator radiation are the relativistic energy parameter γ, the undulator spatial period λ_u, and K. For an undulator, K is about unity (i.e. the maximum angular deviation is about the same or smaller than the natural opening angle $1/\gamma$). Here, both the horizontal and vertical divergences in radiation of the fundamental frequency are given by

$$\sigma_h^p = \sigma_v^p \approx 1/\sqrt{nN}\,\gamma, \tag{3.28}$$

where N is the number of periods in the undulator ($2N$ being the number of magnet poles) and n is the harmonic number (see Equation (3.29) below). For a typical undulator consisting of one hundred poles or more, $\theta_{h,v}^p \sim 10\,\mu\text{rad}$.

It should be noted, however, that invariably the divergence measured at a beamline is significantly larger in the horizontal direction than predicted by Equation (3.28). This is because the observed divergence is a convolution of the photon-beam divergence with that of the electron beam, the latter being considerably larger. This is why the photon divergences in Equation (3.28) have been labelled with the superscript p. Undulators therefore have horizontal divergences more typically of the order of $\sigma_x' = 100\,\mu\text{rad}$.

The transformation from wiggler to undulator radiation is achieved in practice not by reducing the lateral excursions by decreasing the magnetic field strength between the magnetic pole pairs – this would result in an unacceptable drop in flux – but instead by reducing the magnetic pole spatial periodicity λ_u (Equation (3.25)).

We now derive the longest (fundamental) wavelength for constructive interference in an undulator in terms of the undulator period λ_u. Consider Figure 3.19. Radiation is emitted at time $t' = 0$ at point A. After a time $t' = T'$ the electron has moved one undulation downstream to point B, and the radiation that originated at A has propagated a distance cT'. The difference in these two distances is $cT' - \lambda_u$, and only that radiation with a wavelength λ_n equal to an integral fraction of this will interfere constructively, that is

$$n\lambda_n = cT' - \lambda_u. \tag{3.29}$$

The fundamental wavelength λ_1 satisfies Equation (3.29) for $n = 1$. What we need to do now is determine T'. For low-amplitude oscillations of the electron beam path, it is marginally greater than λ_u/v. However, we cannot ignore the oscillatory path and assume it is straight, as the electron velocity v is anyway so close to c, so this small deviation has a significant impact on the condition for constructive interference. Consider Figure 3.20.

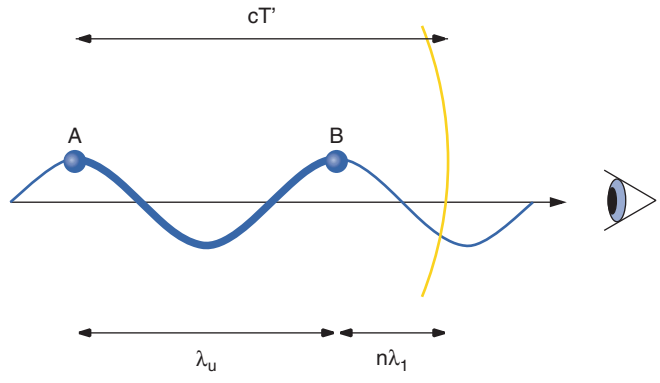

Figure 3.19 *Constructive interference between wavefronts emanating from equivalent points on the undulations occurs when the wavelength is an integral fraction of their separation. The curved path S taken by the electron travelling at the relativistic velocity v is highlighted in blue.*

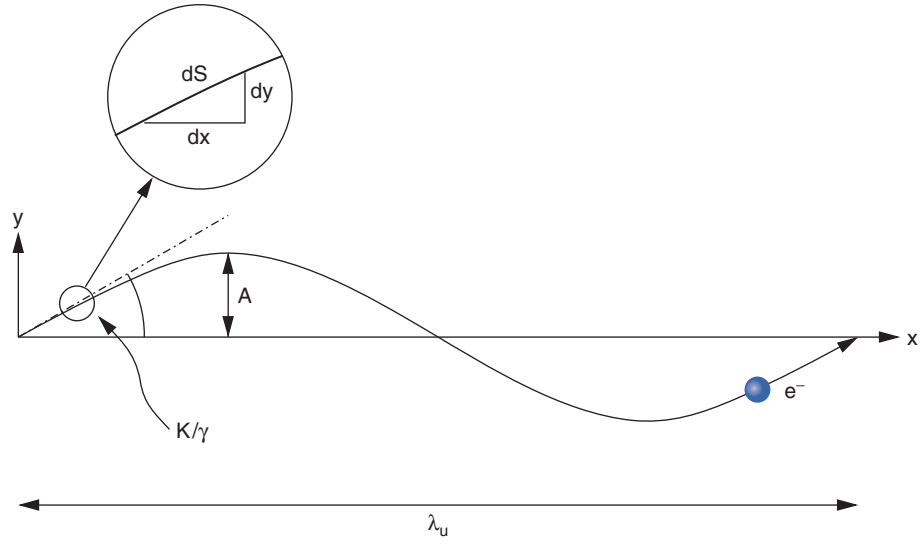

Figure 3.20 *The path taken by an electron of velocity v during one cycle in an undulator insertion device. The y-scale is much exaggerated.*

The path $S = vT'$ taken by the electron in one cycle in the undulator is

$$S = \int_0^{\lambda_u} dS. \tag{3.30}$$

But by Pythagoras' theorem,

$$dS = \sqrt{dx^2 + dy^2}$$

$$= \left[1 + \left(\frac{dy}{dx}\right)^2\right]^{1/2} dx$$
$$\approx \left[1 + \frac{1}{2}\left(\frac{dy}{dx}\right)^2\right] dx. \tag{3.31}$$

Also,

$$y = A \sin\left(\frac{2\pi x}{\lambda_u}\right),$$
$$\frac{dy}{dx} = \frac{2\pi A}{\lambda_u} \cos\left(\frac{2\pi x}{\lambda_u}\right) \tag{3.32}$$

and $dy/dx \ll 1$. What is the value of the oscillation amplitude A? For small angles, $\tan\theta \approx \theta$, and hence

$$\tan\left(\frac{K}{\gamma}\right) \approx \frac{K}{\gamma} = \left.\frac{dy}{dx}\right|_{x=0}$$
$$= \left.\frac{2\pi A}{\lambda_u} \cos\left(\frac{2\pi x}{\lambda_u}\right)\right|_{x=0}$$
$$\rightsquigarrow \frac{K}{\gamma} = \frac{2\pi A}{\lambda_u}$$
$$\rightsquigarrow A = \frac{K\lambda_u}{2\pi\gamma}. \tag{3.33}$$

Note that for typical values for K, λ_u, and γ, A is of the order of a micron, and that the lateral extent of the cone produced by the natural opening angle $1/\gamma$ after a pole separation of, say, 25 mm, is also of the order of a few microns – in other words, the radiation emitted by an undulator does indeed overlap if K is kept to around unity.

Inserting Equations (3.31), (3.32) and (3.33) into Equation (3.30), we obtain

$$S = \int_0^{\lambda_u} \left[1 + \frac{K^2}{2\gamma^2} \cos^2\left(\frac{2\pi x}{\lambda_u}\right)\right] dx. \tag{3.34}$$

We use the geometrical identity

$$\cos^2 A = \frac{1 + \cos 2A}{2}$$

to obtain

$$S = \lambda_u + \frac{K^2}{4\gamma^2} \int_0^{\lambda_u} \left[1 + \cos\left(\frac{4\pi x}{\lambda_u}\right)\right] dx$$
$$= \lambda_u + \left(\frac{K}{2\gamma}\right)^2 \left[x + \frac{\lambda_u}{4\pi} \sin\left(\frac{4\pi x}{\lambda_u}\right)\right]_0^{\lambda_u}$$
$$= \lambda_u \left[1 + \left(\frac{K}{2\gamma}\right)^2\right]. \tag{3.35}$$

The condition for constructive interference is therefore

$$n\lambda_n = \frac{\lambda_u}{\beta}\left\{1 + \left(\frac{K}{2\gamma}\right)^2\right\} - \lambda_u. \tag{3.36}$$

But we know from Equation (3.5) that $\beta \approx 1 - 1/2\gamma^2$. Inserting this into Equation (3.36), we obtain

$$\begin{aligned} n\lambda_n &= \lambda_u/2\,\frac{(4\gamma^2 + K^2)}{(2\gamma^2 - 1)} - \lambda_u \\ &= \frac{2\gamma^2\,\lambda_u + \lambda_u K^2/2 - 2\gamma^2\,\lambda_u + \lambda_u}{2\gamma^2 - 1}, \end{aligned}$$

which, once we recognize that $\gamma^2 \gg 1$, reduces to

$$n\lambda_n = \frac{\lambda_u}{2\gamma^2}\left(1 + \frac{K^2}{2}\right), \tag{3.37}$$

or in practical units

$$n\lambda_n[\text{Å}] = \frac{13.056\,\lambda_u\,[\text{cm}]}{\mathcal{E}^2[\text{GeV}]}\left(1 + \frac{K^2}{2}\right). \tag{3.38}$$

Inserting this into Equation (2.2), we obtain

$$E_n[\text{keV}] = 0.95\,\frac{n\mathcal{E}^2[\text{GeV}]}{(1 + K^2/2)\,\lambda_u\,[\text{cm}]}. \tag{3.39}$$

The more observant reader might have objected to the implicit geometrical assumption in Figure 3.19 – the condition for constructive interference seems to imply that the distances between adjacent pole pairs (positions A and B in Figure 3.19) must be adjusted to an accuracy better than a wavelength, that is, of the order of an Angstrom, which is, needless to say, an unreasonable engineering demand. Fortunately, in arriving at the above condition for interference, we needed to exploit Equations (3.4) and (3.5), which derive from special relativity. The consequence of this is that, from the frame of reference of the electrons, the magnetic lattice is strongly Lorentz-contracted (to be precise, by an amount equal to $\gamma^2/2$) and therefore the undulator period λ_u appears to be shrunk by this amount (of the order of 10^{-7}).

The undulator spectrum consists of a set of narrow lines equally spaced in energy, ΔE, which from Equation (3.37) can be shown to be given by

$$\Delta E = \frac{2hc\gamma^2}{\lambda_u\left(1 + \frac{K^2}{2}\right)}. \tag{3.40}$$

The spectrum can therefore be tuned by changing K. This is achieved by changing the gap between the two sets of magnetic poles and thereby the magnetic field strength B_0.

Note that a high-K undulator gives more intense higher-energy harmonics than a low-K device. The detailed appearance of the undulator spectrum depends on the K-value

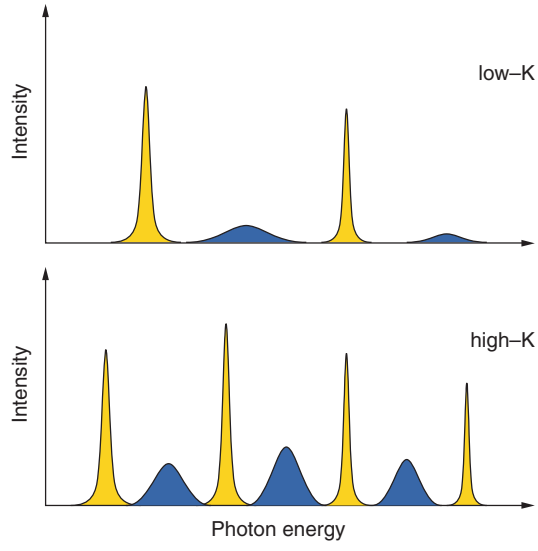

Figure 3.21 *Schematic spectra of low-K and high-K undulators. Note that the harmonics of the low-K undulator have larger photon energy separations and smaller contributions of the even harmonics (blue) compared to the odd harmonics (yellow) than those of the high-K device.*

and the point of observation.[4] We now discuss the form of an undulator spectrum as a function of K.

For an observer on-axis and for low K-values, the even harmonics are largely suppressed, due to symmetry considerations. The spectrum is dominated by $n = 1$. As one increases K (for example, by closing the undulator gap), the spectrum acquires an increasing content of higher harmonics (see Figure 3.21). In addition, the increased angular spread means that even harmonics also begin to appear, as the tight symmetry defining the undulator axis becomes relaxed to an extent. The spectrum of the in-vacuum undulator used at the Protein Crystallography beamline at the SLS is shown in Figure 3.22.

The spectral width of the undulator harmonics depends on the number of periods. As in any interference or diffraction setup, the condition for constructive interference becomes increasingly strict the larger the number of 'scatterers'. In other words, a small relative deviation $\delta\omega$ in the frequency from resonance will still produce a more or less strong positive interference if there are only a small number of participating waves. In contrast, for high N, constructive interference is only satisfied for a very narrow range of frequencies. It turns out that the monochromaticity $\lambda_n/\Delta\lambda_n = \omega_n/\Delta\omega_n$ is equal to the number of periods N multiplied by the harmonic number n.

The on-axis peak intensity I_n of the nth harmonic of undulator radiation is given by

$$I_n = \alpha N^2 \gamma^2 \frac{\Delta\nu}{\nu} \frac{I}{e} F_n(K), \tag{3.41}$$

[4] If the line of sight is off the undulator axis, the conditions for constructive interference must take these angular coordinates into account.

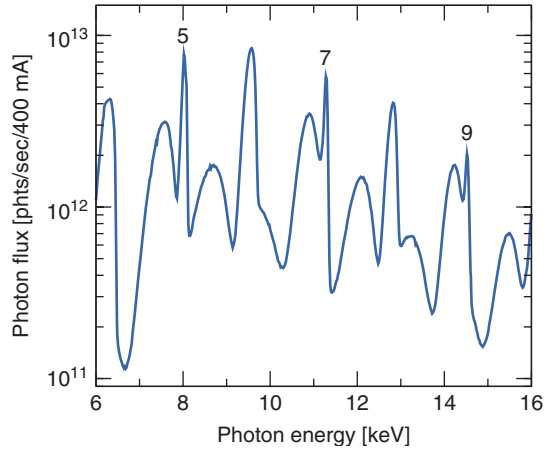

Figure 3.22 *The photon flux produced by the in-vacuum undulator at the Protein Crystallography beamline at the SLS. The undulator has 62 periods, each 24 mm long. Here, the gap size was 6.5 mm, resulting in K = 2.46. The sharp 5th, 7th, and 9th harmonics are labelled. Courtesy Clemens Schulze-Briese, Paul Scherrer Institut.*

where $\alpha = 1/137.04$ is the fine-structure constant, $\Delta\nu/\nu = 1/(nN)$ is the relative spectral bandwidth of the harmonic peak, I is the current, and

$$F_n(K) = \chi^2 K^2 \left[J_{(n-1)/2}(K^2\chi/4) - J_{(n+1)/2}(K^2\chi/4)\right]^2, \tag{3.42}$$

whereby

$$\chi = \frac{n}{1 + K^2/2} \tag{3.43}$$

and the J's are Bessel functions. These tuning functions $F_n(K)$ are plotted for the first six odd harmonics in Figure 3.23. In units of (photons·s^{-1}·mrad^{-2}·0.1% BW^{-1}),

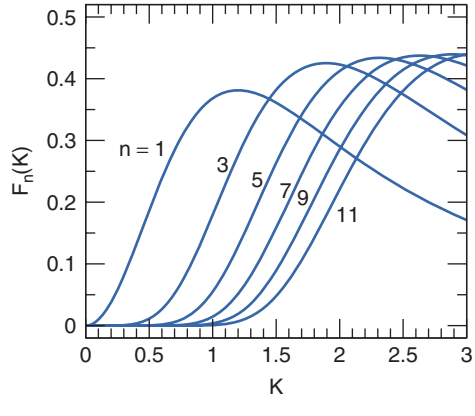

Figure 3.23 *The tuning function $F_n(K)$ for the first six odd harmonics in undulator radiation used to calculate the on-axis peak intensities.*

Equation (3.41) simplifies to

$$I_n = 1.74 \times 10^{14} N^2 \mathcal{E}^2 [\text{GeV}] I [\text{A}] F_n(K). \tag{3.44}$$

So, for example, an undulator with $K = 1.5$ containing 70 periods in a storage ring running at 3 GeV and 200 mA would provide an on-axis peak intensity for the 5th harmonic of 4.05×10^{17} photons·s^{-1}·mrad^{-2} · 0.1% BW^{-1}. For a typical undulator source size of $150 \times 20\,\mu\text{m}^2$, this equates to a brilliance of 1.35×10^{20} photons·s^{-1}·mrad^{-2}·mm^{-2} · 0.1% BW^{-1}.

To scan the photon energy produced by an undulator, one tracks the resonant energy E_n of the nth harmonic (Equation (3.39)) by slowly opening the undulator gap (decreasing K). Different harmonics are used in different energy ranges: above a certain photon energy, the intensity of a given (normally odd) harmonic, given by the function $F_n(K)$, drops below that of the next odd harmonic at the highest accessible K-value, given by the minimum undulator gap, and the undulator is closed accordingly and the new, higher, harmonic is used.

Finally, the polarization of undulator radiation can be controlled by laterally shifting the position of one set of magnet poles relative to the other, which can generate the necessary horizontal magnetic field in the undulator in addition to the vertical field. In a *helical* undulator, the upper set of magnets lies not directly above the lower set, but is shifted so the poles lie exactly in between those of the lower set. The electrons therefore describe a helical, screwlike motion, whereby their velocity in the direction of the undulator axis is invariant. In contrast, this component of the velocity of the electrons in a linear undulator varies (being greatest at the maximum displacement from the axis, and smallest when crossing the axis, see Figure 3.24) and there will therefore be an

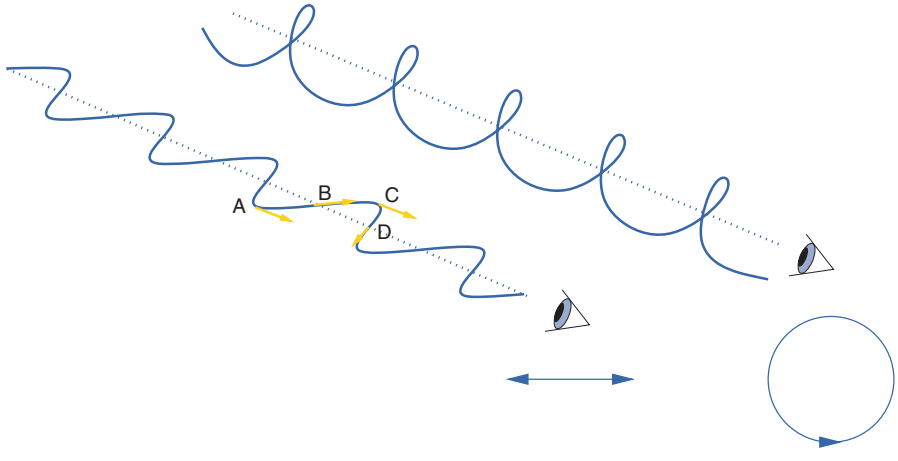

Figure 3.24 *Observed movements of the electrons within undulators observed on axis. In a linear undulator, the electrons appear to accelerate as they approach points A and C in the electron path and decelerate as they move towards the central axis. This change in speed and Doppler shift causes higher harmonics of the emitted radiation to be observed. In a perfect helical undulator, however, the longitudinal component of the electrons is constant, and the emitted radiation has therefore only one frequency.*

oscillatory fluctuation in the Doppler shift. Essentially, it is this which allows higher harmonics to be generated due to the introduction of Fourier components caused by the observed deviation from constant longitudinal motion (towards triangular motion). The Fourier components of a triangular wave are only odd harmonics, which is the reason why the odd harmonics dominate in linear-undulator on-axis spectra. However, in the case of a helical undulator, the magnitude of the magnetic field is constant and only its direction rotates at a constant rate around the undulator axis.

An on-axis observer sees the electrons executing circles, and at no point do they appear to be going faster or slower – the Doppler shift is constant. Therefore, on-axis, there are no higher Fourier components and only the fundamental frequency can be observed. The radiation is circularly polarized.

3.8 Future Sources of Synchrotron Light

3.8.1 The Energy Recovery Linac

Currently existing x-ray sources based on third-generation electron storage rings are limited in their peak spectral brilliance by their emittances and the pulse lengths of the electron bunches (see Section 3.5.1). More precisely, the brilliance is proportional to the flux and inversely proportional to the emittance, as, indeed, is the degree of coherence. Moreover, the peak brilliance is also inversely proportional to the electron bunch length, that is

$$\text{Brilliance} \propto \frac{\text{Flux}}{\epsilon_x \, \epsilon_y}, \tag{3.45}$$

$$\text{Peak Brilliance} \propto \frac{\text{Flux}}{\epsilon_x \, \epsilon_y \, \tau}. \tag{3.46}$$

Third-generation synchrotrons have brilliances of the order of 10^{21} photons·s^{-1}·mrad^{-2}·mm^{-2}·0.1% BW, while their peak brilliances are of the order of 10^{23} photons·s^{-1}·mrad^{-2}·mm^{-2}·0.1% BW. Hence any significant increase in flux or decrease in the emittance and/or bunch lengths would help extend x-ray experiments into new regimes [7]. The energy recovery linac (ERL) is a concept that is potentially capable of providing improvements in all of these properties by approximately three orders of magnitude.

The potential increased performance of ERLs is due to the fact that they circumvent one of the fundamental limitations of storage-ring facilities – an unavoidable consequence of electron storage is that Coulomb repulsion and perturbations by the magnet lattice degrade the electron bunches in all existing third-generation storage ring facilities to an equilibrium configuration after several thousand turns around the storage ring (i.e. in a few ms), limiting the brilliance, coherence, emittance, and pulse length.

An obvious solution to this problem would be to use the radiation from electrons emitted from a very low-emittance source before the emittance of the electron bunches degrades. This can in principle be achieved in a linear accelerator (linac), in which electrons are injected at one end, accelerated to the desired energy, pass through an insertion device and are then discarded. In this case, the quality of the injector determines

the emittances, pulse length, and electron current. Complex pulse structures can easily be programmed into the injector, allowing new types of time-resolved experiments.

The major disadvantage in using a simple linac lies in the fact that, for the acceleration voltages and beam currents of interest, typically 5 GeV and 100 mA, respectively, nearly a gigawatt of power is required to run such a facility, similar to the output of a large power station! Such a solution is therefore both economically and environmentally unsound.

In an ERL, the electrons are accelerated in the linac, then pass around a ring, in which they are used to produce synchrotron radiation. The length of the electrons' path is accurately controlled so that when re-entering the linac, their phase relative to the accelerating AC field is 180° shifted to that when they initially entered the linac from the injector. This means that now they are decelerated instead of being accelerated – in this manner, their kinetic energy is coupled to and recovered by the field of the linac (see Figure 3.25). As these decelerated electron bunches exit the linac, they are deflected by a modest magnetic field (remember, now they have low kinetic energies) into a dump. In summary, in an ERL the electron energy is stored, while in a storage ring the electrons themselves are stored.

Theoretical studies of ERLs predict an increase in flux and brightness of three orders of magnitude over present third-generation storage rings, while it is expected that the degree of coherence will reach several tens of percent, in contrast to storage rings, in which the coherent fraction is typically well under one percent.

Another exciting potential development associated with the short duration that the electrons remain within the ERL is that the minimum gap size of the undulators can potentially be significantly reduced. In a conventional storage ring, the electron-beam

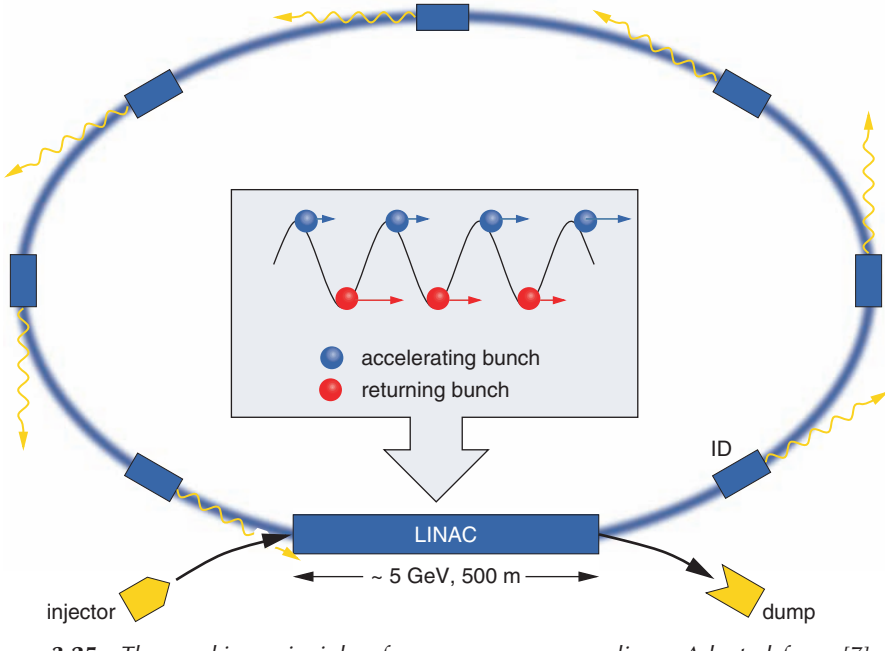

Figure 3.25 *The working principle of an energy recovery linac. Adapted from [7] with permission of Institute of Physics.*

emittance will soon become degraded by interaction between the electrons and the magnetic field if the gap size is too small, but in an ERL, they are removed before this becomes problematic. A smaller gap means a larger magnetic field strength B_0, and therefore a concomitantly shorter undulator period λ_u (see Equation (3.25)). ERL undulators may therefore be only a few tens of centimetres long.

ERLs are only just becoming, after the first decade of the twenty-first century, a feasible prospect. The major technological obstacles are the production of a very small emittance injector, the required high repetition rate and the need for a superconducting linac (otherwise, the resistance losses in the inner walls of the linac cavities would be unacceptably high). At the time of writing (2011), feasibility projects and funding programmes to design and construct the first ERL in the next decade are beginning at several locations around the world, most notably at Cornell University, the same institute where the potential of synchrotron light for materials research was first investigated by Tomboulian and Hartmann, half a century earlier.

3.8.2 The Free-Electron Laser

Femtosecond lasers operating in the near-infrared and visible regions have been available for well over two decades. This has made it possible to investigate in the form of 'snapshots' or 'movies' many processes with characteristic times under a picosecond, such as vibrational motion in molecules, exciton generation, and photochemical reactions. The *spatial* resolution of lasers is, however, severely limited by the wavelengths used to a few hundred nanometres. On the other hand, x-rays are able to provide spatial (i.e. structural) information on the atomic scale and can follow processes using the time-structure of third-generation synchrotrons down to approximately ten picoseconds.

A natural step forward for x-ray researchers would therefore be to develop femtosecond lasers that produce intense, perfectly focused, and coherent x-ray beams. This would enable them to study the dynamics, properties, and structure of materials with unsurpassed detail on a time-scale that is presently inaccessible. It would also make it possible to create holographic images of complex molecules such as proteins, potentially revolutionizing this field of research.

Unfortunately, conventional lasers cannot generate photons more energetic than about 8 eV (i.e. ultraviolet light of around 150 nm), which are far less energetic than x-rays. This is, in part, because conventional lasers rely on mirrors, which not only become less efficient as they reflect higher-energy light, but also absorb too heavily, increasing the cavity losses, thereby forcing the cavity below the lasing threshold. They will also begin to burn!

This problem has been resolved by developing the so-called free-electron laser (FEL), operating in the far ultraviolet and x-ray regimes. According to classical electrodynamics, a pointlike packet of charge $Q = N_e\, e$ moving at relativistic velocities radiates with a power which is proportional to N_e^2, whereby N_e is the number of electrons within the charge. Here, 'pointlike' means that the packet of charge in the direction of propagation is significantly shorter than the wavelength of the emitted radiation and the electrons all emit coherently (see Figure 3.26). We have already seen earlier in this chapter that the bunches formed by the electrons passing through the RF supply in third-generation synchrotron sources have a duration of the order of 50 ps, that is, they are a few millimetres long, or

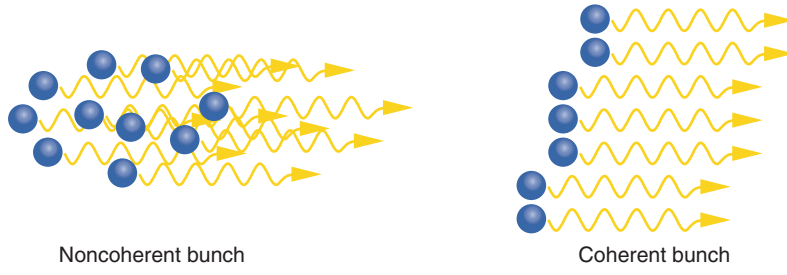

Figure 3.26 *If the size of an electron bunch in the direction of the propagation of radiation is large compared to the wavelength of the radiation, it is incoherent. If, however, the bunch is small compared to the radiation wavelength, radiation is in phase and therefore coherent, and the amplitudes of the radiation from each of the N_e electrons add linearly, resulting in an increase in intensity by N_e^2.*

some 10^5 to 10^8 longer than the x-rays they emit. Hence, third-generation synchrotrons are not efficient coherent sources.

3.8.2.1 The SASE Process

It is technically impossible to squeeze the electrons into such short bunches by using magnetic lenses. This seemingly insurmountable problem is solved by exploiting a process called 'self-amplified spontaneous emission' (SASE), first proposed by Anatoli Kondratenko and Evgeni Saldin in 1980 [8] and worked out in detail by Rodolfo Bonifacio, Claudio Pellegrini, and Lorenzo Narducci in 1984 [9]. Crucially, SASE does not require the mirrors used in a conventional laser cavity. Instead, a SASE free-electron laser uses a monoenergetic electron beam and a very long, high-quality undulator, typically of the order of a few hundred metres in length.

In deriving Equation (3.37) for the resonant fundamental x-ray wavelength ($n = 1$) generated by undulators, we started off by assuming that the difference in length between the path executed by the electron as it oscillates through a single undulator period λ_u and that of the electromagnetic field emitted by the electron was the length of a single wavelength (see Figure 3.19). Hence, from the perspective of the electron, the electromagnetic field has gained on it by a single wavelength and hence it (the electron) has experienced a single 2π cycle change of that field. Importantly, during the same period, the electron has itself executed one cycle of transverse motion due to the undulator's magnet array. We show this in Figure 3.27.

As the electrons begin to propagate down the undulator, they emit radiation independently and stochastically. However, they are bathed in the same light they generate. As they oscillate back and forth through the magnets, they also interact with the magnetic field component of this light. The transverse component of the electrons' motion (i.e., the sinusoidal motion induced by the magnet array which causes them to oscillate) and the magnetic field component of the emitted radiation create a Lorentz force which acts on the electrons in the axial direction. Depending on the phase, the transverse oscillatory motion of a given electron has relative to the electromagnetic field, the Lorentz force either causes the electron to accelerate, or otherwise, lose energy. As a result, the

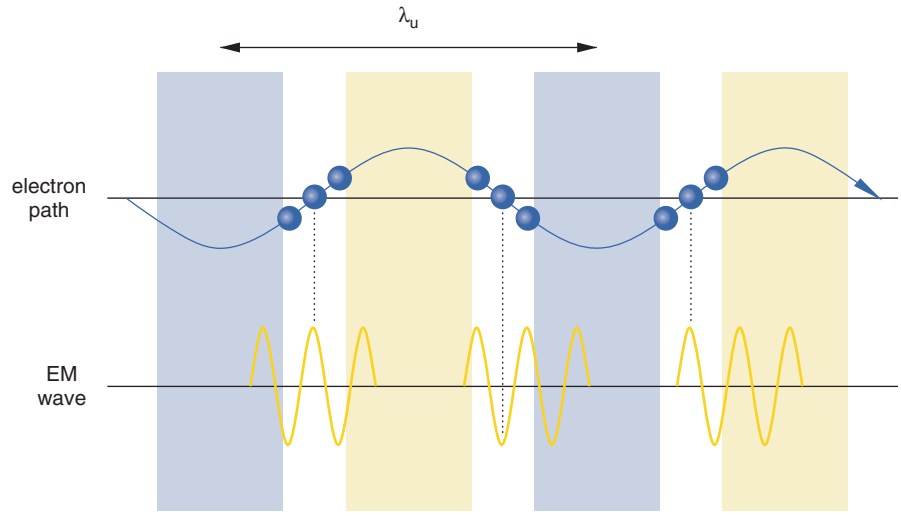

Figure 3.27 *The electron trajectory and the electromagnetic field propagation in an FEL undulator. The phase of the transverse motion of each electron relative to the phase of the electromagnetic wave remains constant along the length of the undulator.*

electrons begin to form microbunches (within the conventional bunches, described earlier in this chapter), separated by a distance equal to the wavelength of the light they generate. In the laboratory frame of reference, this is equal to the pole magnet periodicity divided by γ, hence is of the order of a few microns. Each microbunch therefore has a duration of a few femtoseconds to a few tens of femtoseconds. FELs therefore deliver femtosecond x-radiation.

Another intuitively pleasing way to visualize the microbunching process is as follows: in the laboratory frame of reference, the electron interacts with two fields, namely its own emitted radiation and the static undulator magnetic field. In the rest frame of the electron, however, relativistic effects transform the static undulator field into a counterpropagating wave with a wavelength equal to the relativistically (Lorentz) contracted undulator period, which is exactly equal to the radiation wavelength. This interferes with the copropagating emitted field, resulting in a standing wave which exhibits the spatial potential responsible for the bunching.

The light waves emitted by the electron bunches will line up in phase – meaning that the waves' peaks and valleys overlay each other – to reinforce and amplify the light's brilliance and intensity (see Figure 3.28). Eventually, a favourable runaway instability develops. The light intensity grows exponentially along the undulator until the process saturates, bringing the beam to its highest possible intensity [10]. This amplification operates only within a very narrow bandwidth around the undulator resonant wavelength, given by Equation (3.37).

By the time the light beam emerges, its initial peak intensity is amplified by more than a factor of a billion. This enormous factor results from the fact that the emission from the electrons in a given microbunch is now coherent and the amplitudes must be vectorially added. While the intensity from N incoherent emitters is proportional to N, that from a

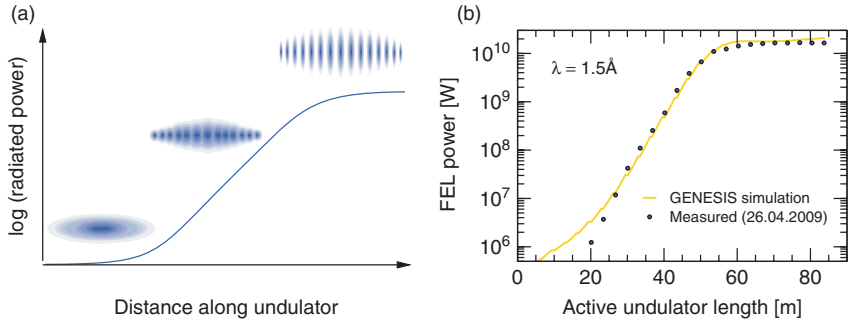

Figure 3.28 *The self-amplification spontaneous emission (SASE) process in an FEL undulator. (a) Because of the interaction of the electrons with the synchrotron radiation they emit, a longitudinal density modulation (micro-bunching) develops together with a resulting exponential growth of the radiation power along the undulator. Note that in reality the number of micro-bunches within a conventional electron bunch is much larger than shown ($\geq 10^5$ for x-ray FELs). (b) The gain/saturation curve measured at the LCLS in April 2009, along with the theoretical expected curve, calculated using the GENESIS simulation package. Reprinted from [10] courtesy Paul Emma and John Galayda, SLAC National Accelerator Laboratory, with permission of Macmillan Publishers Ltd.*

coherent source is proportional to N^2. The spacing between the microbunches is equal to the emitted wavelength λ, hence all the microbunches emit coherently with one another. As the number of electrons in all the microbunches is typically 10^8 to 10^9, the increase in brilliance is correspondingly high. Hence, FELs not only provide exceedingly intense radiation, but also radiation that is fully transversely (or spatially) coherent, in stark contrast to radiation from third-generation sources, which have typical coherent fractions of less than 10^{-3}. This enhanced coherence is one of the most important properties of FELs, as we will see in the worked example below.

The very short pulse duration of light emitted by all the microbunches means that atomic motions in condensed matter, that is, vibrations which have periods of the order of picoseconds down to tens of femtoseconds, can be tracked using x-radiation, a hitherto impossible task. Hence, in summary, FELs will provide peak brilliances, a degree of coherence, and pulse durations that are many orders of magnitude superior to radiation from third-generation sources [11, 12].

In contrast to synchrotron storage rings, the electron energy in linac-based sources such as FELs can more easily be changed. Also, because the undulators used in x-ray FELs are so long, it is impractical to tune the photon energy by varying the undulator gap (or K, given by Equation (3.25)). Instead, one changes the electron energy, expressed by γ, which in turn affects the condition for constructive interference in the undulator, given by Equation (3.37).

3.8.2.2 Seeded FELs

The output from a SASE FEL, although highly spatial and coherent in the transverse plane, can have a very spiky temporal structure, which seriously compromises the temporal (or longitudinal) coherence. The reason for this is the fact that the amplification

process starts from noise and the output therefore consists of a number of independent wave packets.

It can be easily shown that the ultimate limit (the so-called 'Fourier-transform limit') to the temporal coherence of a Gaussian packet of radiation of pulse length $\Delta\tau$ is given by

$$\Delta\tau\,\Delta\nu = \frac{2}{\pi}\ln 2 = 0.44, \tag{3.47}$$

whereby $\Delta\nu$ is the spread of frequencies given in Hz and $\Delta\tau$ is the pulse duration in seconds. Therefore the spectral purity is

$$\frac{\Delta\nu}{\nu} = \frac{2\lambda \ln 2}{\pi c\,\Delta\tau} \tag{3.48}$$

$$= 1.472 \times 10^{-4}\,\frac{\lambda[\text{Å}]}{\Delta\tau[\text{fs}]}. \tag{3.49}$$

A Fourier-transform limited 10 fs-pulse of 1-nm radiation would have a relative bandwidth of 1.472×10^{-4}, very close to the best monochromation possible using a Si(111) crystal. Therefore, any process which could clean up the spectral purity of FEL radiation would have an enormous technological benefit insofar that one could dispense with the need for a monochromator. This can be achieved to a large extent by seeding the amplification process with a highly temporally coherent source. Seeding has other added advantages: the transition from incoherent radiation at the beginning of the undulator to fully saturated radiation occurs earlier, allowing the use of shorter (and cheaper) devices; the stability of the resonant wavelength and the output power is much improved; and synchronization from the seed radiation with external signals for pump-probe experiments is simplified.

Seeding can be accomplished in one of several possible ways. We briefly discuss now some of the most important.

Although, in the vacuum-ultraviolet and x-ray regime, there are no laser seed sources of sufficient intensity, seed radiation can be produced in a so-called 'self-seeded FEL' by using a short SASE FEL to generate radiation at power levels approximately three orders of magnitude below saturation. The resulting photon pulse is then spectrally filtered ('cleaned up') by passing it through a narrow-band monochromator. As the range of wavelengths ($\Delta\nu$ in Equation (3.47)) is reduced by this process, the pulse becomes stretched in time. In the meantime, the electron beam is forced through a magnetic chicane. This has two purposes – the first is to delay the electron beam so that it enters a second, longer undulator at the same time as the seed photon pulse; the second is to destroy the density modulation of the electron bunches produced by the first SASE FEL. The high-purity photon pulse, which overlaps the electron bunch when they enter the second undulator, now provides the necessary seed to induce coherent amplification in this second FEL amplifier stage (see Figure 3.29(a)).

Another strategy for producing seeded radiation is similar to self-seeding and relies on the fact that in Equation (3.37), higher-harmonic radiation can be produced (where n is an integer greater than one). Radiation containing a rich higher-harmonic content of a relatively long-wavelength seed source of wavelength λ_s (which may be a laser or a harmonic selected by HHG, see below) can be generated in a so-called 'modulator' undulator which operates close to, but below, saturation. This is then fed into a

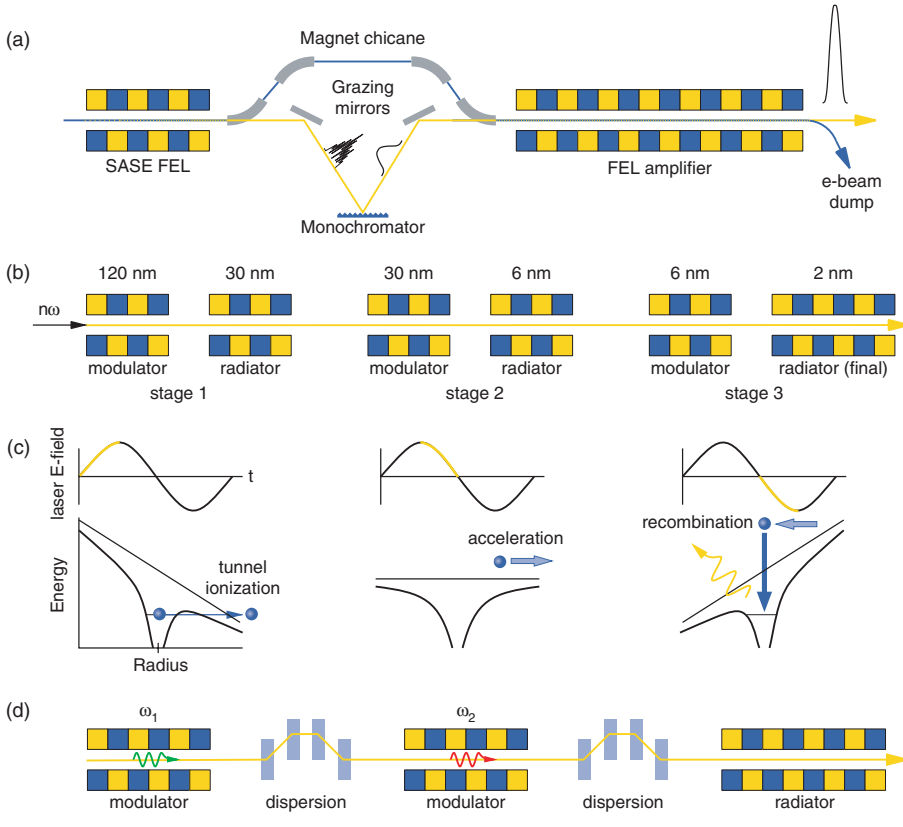

Figure 3.29 *Different seeding strategies in FELs. The setups for (a) self-seeded FELs and (b) HGHG. (c) The ionization and recombination processes responsible for HHG. (d) The configuration for EEHG.*

so-called 'radiator' undulator tuned to one of the higher harmonics, typically $n = 5$. This high-gain harmonic generation (HGHG) process can be cascaded for multiple harmonic up-conversions (Figure 3.29(b)). The performance of HGHG cascades is enhanced by inserting dispersion elements in between the modulator and radiator. HGHG is, however, a complex technique which is very sensitive to the gain at each intermediate stage and to the timing between the electron and photon pulses.

In high-harmonic generation (HHG), a noble gas is ionized in the intense electric field of a high-brilliance femtosecond laser. The ejected electron is then further accelerated in the electric field of the same laser pulse to very high energies. When the sign of the field reverses, the electron can return to and recombine with the parent ion with a certain probability in a radiative process, resulting in the release of a photon with a high harmonic frequency of the incoming radiation [13], as shown schematically in Figure 3.29(c). So, for example, it has been possible to generate a 5 nJ pulse of the 67th harmonic of an 800 nm driving laser pulse, which has a wavelength of only 12 nm.

This seed pulse contains 3×10^8 photons! It is noted, however, that the production of significant amounts of x-radiation in the nm- and sub-nm range using HHG is presently impractical.

Lastly, we briefly describe the newest and perhaps the most promising strategy for seeding, referred to as echo-enabled harmonic generation (EEHG) [14]. It lies beyond the scope of this introductory text to explain in detail how EEHG occurs, and we therefore only provide a qualitative description here. The setup for EEHG is shown in Figure 3.29(d) and consists of two modulator/dispersion element pairs plus a radiator. The electron bunch entering the first modulator interacts with a laser beam of relatively low frequency ω_1, whereby the modulator is also tuned to resonate at ω_1. Next, the electron beam, which now has a modulated density, passes through a compressing dispersive device, thereby shortening the modulation wavelength. The same process is repeated in a second modulator/dispersion pair, but this time at a frequency ω_2. After a certain 'sleep' time, an echo signal appears at an angular frequency $n\omega_1 + m\omega_2$, whereby n and m are integers. This echo occurs as a result of a recoherence effect as different frequencies in the beam come briefly back into phase. The key advantage of this scheme is that the high harmonic intensities of the echo signal decay only slowly with increases in the integers n and m.

3.8.2.3 Present and Planned FEL Facilities

Table 3.2 shows a list of the important SASE FELs in operation and under construction worldwide. The LCLS in Stanford, California, is the first true x-ray FEL to come online, in the Spring of 2009 and is the most intense source of x-rays by three orders of magnitude compared to the previous front-runner, FLASH, in Hamburg, Germany (providing radiation in the vacuum ultraviolet), and by eight orders of magnitude compared to the most powerful third-generation facility, SPring8, in Japan (see Figure 3.30).

The LCLS first lased on 10th April 2009. Ten days later, the first diffraction pattern using hard x-rays (8 keV) was recorded and is shown in Figure 3.31. It is of a 40 μm-diameter carbon fibre, imaged on a Ce:YAG plate 52 m downstream from the fibre. Note that the ratio between the fringe separation and the sample–detector distance (0.2 mm/52 m) is the same as that between the wavelength of the x-rays and the fibre diameter (1.5 Å/40 μm), as expected.

Table 3.2 *A selection of some of the important worldwide free-electron laser facilities operating in the vacuum-ultraviolet and the x-ray region and exploiting SASE, shown in order of decreasing wavelength ranges (increasing photon energy).*

Facility	Linac energy [GeV]	Wavelength range	Status*
Fermi, Italy	1.2	100–10 nm	C
FLASH, Germany	1	125–6 nm	O
SwissFEL, Switzerland	5.8	10–0.1 nm	P
SCSS, Japan	8	60–0.1 nm	C
LCLS, USA	13.6	1.5–0.14 nm	O
European XFEL, Germany	17.5	5–0.1 nm	C

*O = operational; C = design phase or under construction; P = planned

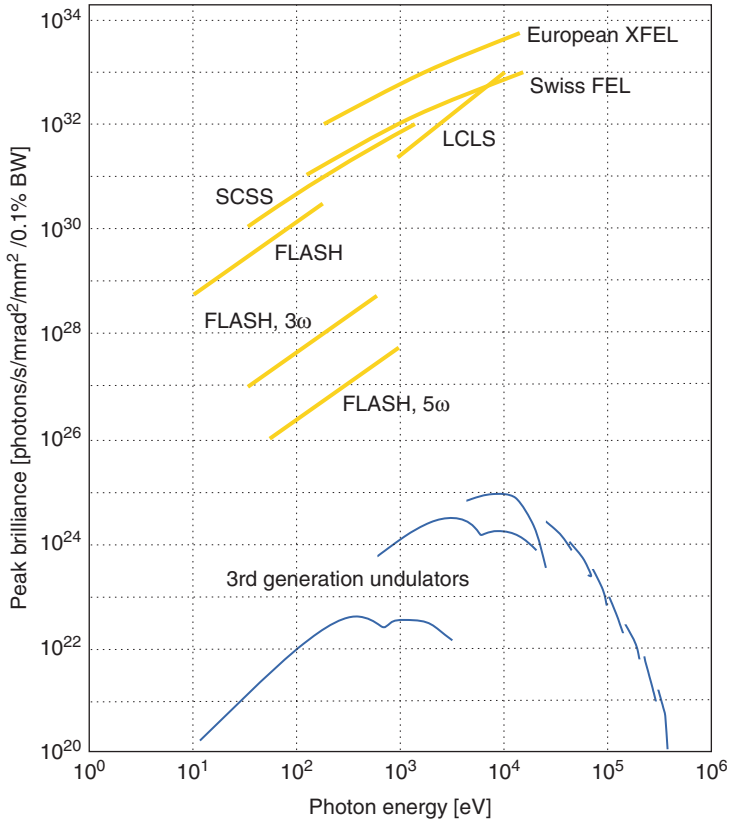

Figure 3.30 *Peak spectral brilliances of selected SASE FELs plus those of state-of-the-art third-generation synchrotron facilities shown for comparison. Note that the* peak *brilliances of the third-generation synchrotron sources shown here are approximately three orders of magnitude greater than the* average *brilliances shown in Figure 1.3.*

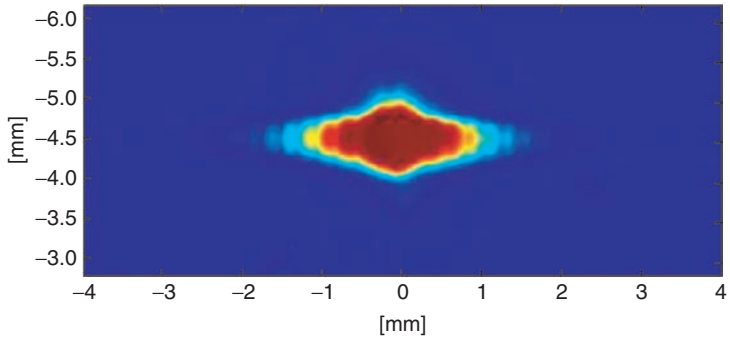

Figure 3.31 *The first hard x-ray diffraction image recorded at the LCLS free-electron laser. It is of a 40*μm *carbon fibre. See [15] for acknowledgements.*

The European XFEL, under construction in Hamburg, is expected to have first beam in 2015 and is predicted to have a maximum peak brilliance nearly an order of magnitude greater still than that of the LCLS, and will also provide 1 Å-radiation (see Figure 3.30).

The successful construction of FELs operating in the x-ray region requires that physicists and engineers must adopt a completely novel approach to beam control and manipulation and experimental procedures – the enormous increase in flux and brilliance means that essentially all experiments will be 'one shot', that is, the sample will be locally destroyed where it has been irradiated by a single FEL pulse. It is therefore unlikely that the incredible increase in brilliance offered by FELs will result in a drop in demand for synchrotron light from third-generation facilities (and very possibly in the future also from ERLs) for use in 'traditional' synchrotron experiments, which will certainly persist despite the development of FEL facilities.

3.8.2.4 Worked Example – Lensless Imaging at FLASH

It is known that the peak brilliance of focused FEL pulses is so high that any irradiated sample will be stripped of all its electrons, resulting in a rapid formation of a plasma and a giant Coulomb explosion, utterly destroying the sample where it is irradiated.

A crucial question, therefore, in using FELs is whether the sample under observation will survive longer than the duration of the femtosecond x-ray pulse with which it is being illuminated. How stringent this condition must be depends on the type of experiment being performed. When the atomic structure is to be investigated, the atoms must still have most of their electrons localized around the nuclear core for a large fraction of the pulse duration, as it is the electrons (not the nuclear framework) which scatter the x-rays, and the atoms' locations are given by the local electron density. Theoretical calculations indicate that pulse lengths of approximately 70 fs or less, depending on the integrated photon flux, are required to ensure that molecules preserve their structure long enough to allow meaningful diffraction data to be recorded [16]. Although such short durations are a serious technological challenge, they are certainly feasible.

In the case of scattering by mesoscopic structures, it is only required that the electrons have not strayed significantly from their original locations after irradiation, compared to the characteristic lengths of the average electron density contrast of the object. In order to establish whether this less exacting case of probing mesoscopic structures is possible using FELs, an experiment was carried out on a micronsized, nonperiodic image using a single 25 fs pulse, containing approximately 10^{12} photons at 32 nm (38.7 eV), from the FLASH FEL facility in Hamburg. The beam was focused down at the sample to approximately $30 \times 20\,\mu m^2$, resulting in a peak intensity of approximately 4×10^{17} W cm^{-2}. The sample, shown in the inset of Figure 3.32(a), was a simple pictogram produced by selectively etching through a semitransparent, 20-nm thick silicon nitride membrane. The pictogram caused the VUV pulse to be scattered. After reflection off a specially designed multilayer mirror, the far-field coherent diffraction pattern was recorded by a CCD x-ray detector (see Chapter 4), and is shown in Figure 3.32(b).

Using Fourier-transform techniques in conjunction with so-called 'phase-retrieval' algorithms, in a method called lensless imaging, discussed in Chapter 7, the original picture could be retrieved from the diffraction image, shown in Figure 3.32(c). Note also that subsequent diffraction images of the sample looked very different, while

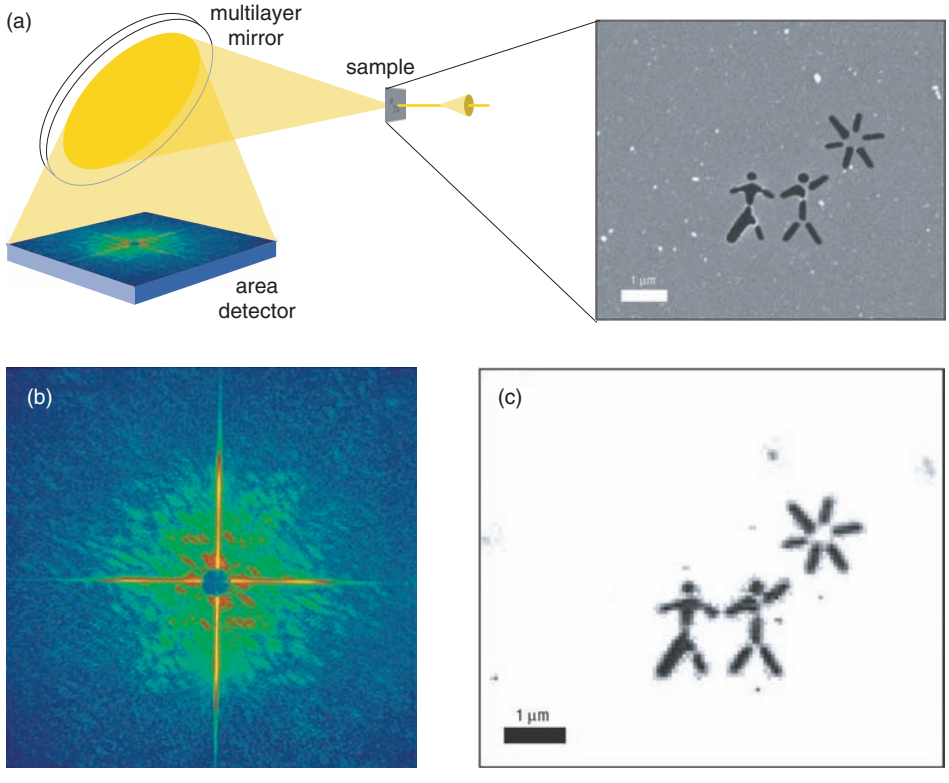

Figure 3.32 *Lensless imaging at FLASH. (a) Setup of the lensless imaging experiment, also showing the SEM image of the pictogram etched into the SiN membrane. (b) The coherent diffraction pattern captured by the area detector after the single-shot exposure of the sample. The sharp vertical and horizontal features are caused by the edges of the square window frame of the membrane. The more diffuse speckle features are generated by the contrast in the pictogram. (c) The reconstructed image retrieved from the diffraction pattern, using a phase-retrieval method. Adapted from [17] with permission of Macmillan Publishers Ltd.*

inspection of the membrane showed it to be utterly destroyed over an area of approximately 100 μm^2 [16].

Hence it was demonstrated that, using the diffraction image from a single FEL pulse, the original sample shape could be regenerated at a resolution of approximately 50 nm, paving the way for more challenging experiments in this field in the future.

3.8.2.5 Example – Femtosecond Protein Nanocrystallography at LCLS

In February 2011, the first successful structural determination of a macromolecule using an x-ray FEL was reported [18]. The approach was to record partial diffraction patterns from individual protein nanocrystals held within a narrow water jet and irradiated by 1.8 keV (6.9 Å) XFEL pulses (see Figure 3.33) focused down to a 7 µm diameter and running at a repetition rate of 30 Hz. As the nanocrystals (varying in linear dimension from approximately 200 nm to 2 µm) were oriented randomly in the jet, different parts of

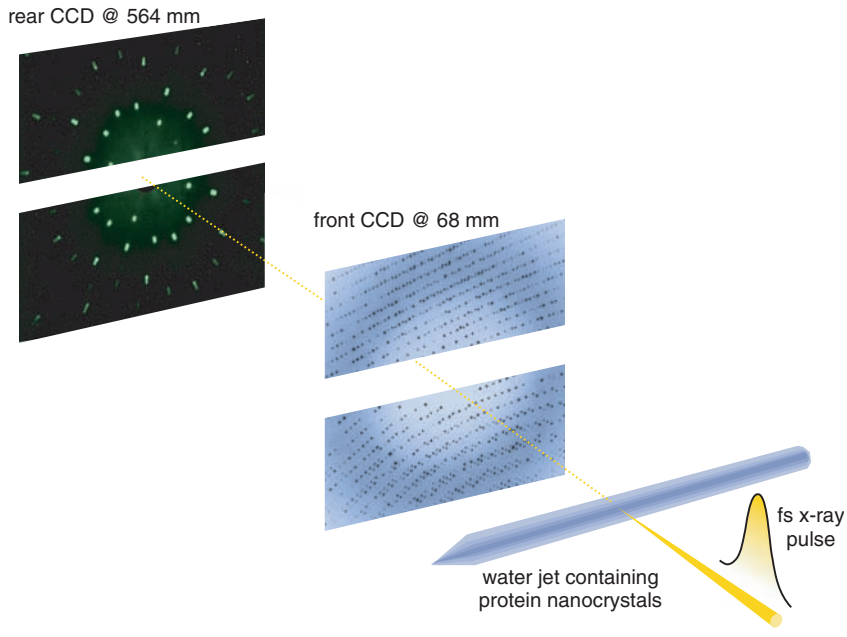

Figure 3.33 *Femtosecond nanocrystallography. Nanocrystals suspended in a water jet are irradiated with fs x-ray pulses from the LCLS XFEL. The highest resolution information is obtained from the diffraction signal scattered through the largest angles on the front CCD positioned at 68 mm. Information on larger scales provided by diffraction signal at low scattering angles was recorded by the rear CCD positioned at 564 mm from the water jet. Adapted from [18] with permission of Macmillan Publishers Ltd.*

the complete diffraction set were recorded for each XFEL pulse. Importantly, a significant fraction of the diffraction data, approximately 13%, could be indexed (that is, the positions in reciprocal space of the diffraction spots could be identified, see Chapter 5) and thereby a 'composite' complete diffraction set was generated from tens of thousands of such partial images.

The material investigated was the model membrane protein[5] 'photosystem I', which has a molecular weight of approximately 1 MDa, and for which, unusually for membrane proteins, its atomic structure has been solved. The crystals were of the fully hydrated protein, for which the solvent fraction by volume is 78%. In conventional protein crystallography, the crystal is cryocooled in order to suppress the effects of radiation damage. Fully hydrated crystals with such a large fraction of water are thereby destroyed by the formation of nanocrystals of ice, which therefore precludes their investigation using conventional techniques. This is another advantage of using an XFEL – the nanocrystals will certainly be completely destroyed in any case by the XFEL pulses (not merely radiation damaged), therefore cryocooling brings no benefit.

[5] At the time of writing, less than 300 structures of membrane proteins have been solved, despite their importance in vital biological functions at cell membranes. The main reason for this paucity of data is the great difficulty in synthesizing crystals of membrane proteins of sufficient quality and size for conventional protein-crystallography experiments.

Each x-ray pulse contained more than 10^{12} photons. Different experiments were performed at pulse durations of 10, 70, and 200 fs, and the diffraction-pattern quality compared. The patterns for 10 and 70 fs pulse duration were essentially indistinguishable, while that using a pulse length of 200 fs was significantly weaker, indicative that somewhere between 70 and 200 fs, the nanocrystals explode. The composite diffraction patterns from the shorter pulses agreed well with data recorded using conventional synchrotron radiation at 12.4 keV, albeit with lower resolution.

The concept of 'diffraction before destruction' using XFEL radiation and already proven down to resolution scales of approximately 60 Å on mesoscopic systems [16] seems therefore also to have been demonstrated at the near-atomic scale.

3.8.3 Tabletop Synchrotrons

Two novel compact sources of intense x-ray beams have been developed in the last decade or so [19, 20]. Although these machines are still in the developmental stage, it is foreseen that 'table-top synchrotrons' will provide fluxes in the hard x-ray regime somewhere in between that provided by a rotating anode laboratory source and that emitted from a bending-magnet beamline. Because the size of the equipment is just a few metres, such devices could be installed in a standard laboratory environment. This would open up a new vista of experiments, such as in protein crystallography, which could be carried out at a home laboratory.

3.8.3.1 The Compact Light Source

A novel compact source of intense x-ray beams has recently been developed in the United States [19]. Although still in the developmental stage, it is foreseen that this 'table-top synchrotron' will provide fluxes in the hard x-ray regime somewhere in between that provided by a rotating anode laboratory source and that emitted from a bending-magnet beamline. Because the size of the equipment is just a few metres, such devices could be installed in a standard laboratory environment. This would open up a new vista of experiments, such as in protein crystallography, which could be carried out at a home laboratory.

The compact light source (CLS) exploits the interaction of intense, near-visible infrared light with a beam of electrons in a small (ca. 2 m diameter) storage ring [19] (see Figure 3.34). The energy of the electrons is, at 25 MeV, two orders of magnitude smaller than that used in large synchrotron facilities, and can be provided by a small linear accelerator followed by injection into the ring.

The electric and magnetic fields of a mode-locked pulsed infrared laser (1 μm wavelength) cause the electrons to wiggle and hence emit radiation via inverse Compton scattering[6] with a spectrum similar to that produced by a long undulator.

From Equation (3.37), we know that the fundamental wavelength of undulator radiation is of the order of $\lambda_u/2\gamma^2$. Hence for conventional magnetic undulators with periodicities λ_u typically a few centimetres, γ needs to be of the order of 10^4 to obtain x-radiation. In a CLS, however, the infrared laser acts as the undulator, and $\lambda_u \approx 1\,\mu\text{m}$, which means

[6] In the inverse Compton effect, electrons are decelerated by photons, thereby increasing the latters' energy. This is the inverse of the process shown in Figure 2.6

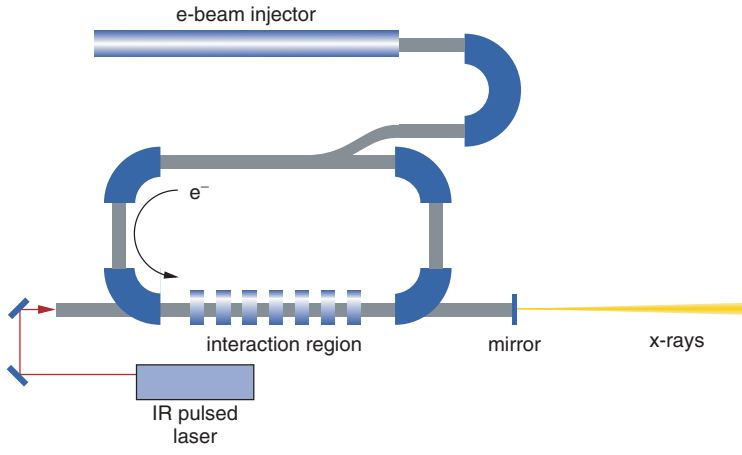

Figure 3.34 *The compact light source generates x-rays via electron oscillations induced by the electromagnetic field of a pulsed infrared laser propagating with the electron beam. Adapted from [19] with permission of the American Physical Society.*

one requires $\gamma \sim 100$. This is why the storage ring energy can be made to be so much smaller in a CLS, of the order of a few tens of MeV. This is, of course, associated with a larger beam divergence, which is typically a few tenths of a degree.

Large changes in the desired x-ray energy can be achieved by tuning the electron beam energy in the storage ring. The photon flux remains fairly constant up to x-ray energies of many tens of keV. Since a significant fraction of the x-rays lies within the energy bandwidth useful for most x-ray applications, the total x-ray power is a only fraction of a watt, instead of approximately a kilowatt of power produced at beamlines from third generation synchrotrons. This naturally narrow bandwidth simplifies the required x-ray optics, since component heating and additional radiation shielding can be avoided.

3.8.3.2 Worked Example – Imaging Soft Matter

Modern x-ray imaging devices rely on differential absorption of the specimen to be imaged – bone absorbs strongly whereas flesh is more transparent. This can be problematic when the features under investigation have a similar absorption strength to that of the material in which they are embedded. An obvious example would be the detection of tumours in a mammographic scan.

As will be discussed in more detail in Chapter 7, this problem can to a large extent be overcome by applying so-called 'phase-contrast imaging'. This technique requires, however, that the x-ray beam has some degree of spatial and temporal coherence and that the beam be large enough to cover a reasonable field of view. Although this is certainly offered by third-generation synchrotron facilities, their use in medical imaging is certainly impractical, as synchrotrons are far too costly, and are also incompatible with a clinical environment.

The CLS also meets the necessary requirements and is relatively cheap and small. In addition, the large divergence of the CLS compared to that typically found at synchrotrons is an advantage in imaging – a larger field of view can be obtained for short

Figure 3.35 *Three types of image contrast of a moth after different data processing of the same data set. (a) Standard absorption contrast, (b) differential phase contrast; and (c) dark field. The arrows indicate features which become much clearer in the phase-contrast and dark-field images than in the standard absorption-contrast image. Reprinted from [21] with permission of IUCr.*

source-to-sample distances. In this example, the natural divergence of the CLS was 20 mrad (1 °), although it was collimated using an aperture to 4 mrad. Hence, at a distance of 10 m from the source, the field of view was 4 cm.

Phase-contrast images of different biological samples were made possible by the use of a grating interferometer, described in detail in Chapter 7. An example is shown in Figure 3.35 [21]. Although exposure times were measured in minutes, making medical application with the present flux impractical, it is expected that once the CLS reaches its projected design specifications, exposures might be reduced to the order of one second.

3.8.3.3 The Laser Wakefield Accelerator

In late 2010, it was reported that bright, spatially coherent x-rays with a peak brilliance of 10^{22} photons·s^{-1}·mrad^{-2}·mm^{-2}·0.1% BW had been produced using a so-called laser wakefield accelerator (LWFA), a value which is only approximately one order of magnitude lower than that of third-generation synchrotrons, but generated by a machine which fits comfortably in a medium-sized laboratory [20].

The concept of wakefield accelerators was first proposed as long ago as 1979 [22] and first experimentally demonstrated in 1988 [23]. The basic ideas behind wakefield acceleration run as follows. When a laser pulse propagates through a dilute plasma, it induces through its electric-field component an oscillating plasma wave of angular frequency

$$\omega_p = \sqrt{\frac{n_0 e^2}{m\epsilon_0}}, \tag{3.50}$$

where e, n_0, m, and ϵ_0 denote the elementary charge, the plasma's free-electron density, the electron mass, and the permittivity of free space, respectively. A weak, or 'diluted' plasma typically contains 10^{25} free electrons m^{-3}, for which the plasma frequency is 1.7×10^{14} rad s^{-1} and the plasma wavelength is $\lambda_p = 2\pi c/\omega_p = 10^{-5}$ m. The local degree of ionization around the laser pulse can be very much higher if the electric field strength of the laser pulse is sufficient to strip the electrons from the nuclear core.

Excitation is most effective if the laser pulse length is shorter than λ_p. For $\lambda_p = 10^{-5}$ m, this equates to a pulse length of less than 33 fs [24]. Because the electrons are much lighter than the nuclear cores, they move much further during excitation, resulting in significant charge separation.

Although the individual particles (electrons and nuclear cores) move relatively slowly, macroscopically a wave, or 'wakefield' of very high charge separation and strong potential gradient is seen to travel through the plasma at the same velocity as that of the laser pulse, that is, the group velocity of the laser pulse, almost the speed of light c. An electron injected into this region of high ionization will be accelerated towards it and be 'dragged along'. Indeed, if the electron approaches the wakefield from an oblique angle, i.e., across the face of the wakefield, they move even faster, much like a surfer rides more quickly across the head of a breaker than the wave is moving forwards. They will also execute an oscillatory path due to the Coulomb attraction of the central positively charged region (see Figure 3.36), which thereby produces undulator-like synchrotron radiation.

The x-ray beam produced by LWFA in 2010 has a relatively large divergence of $\sigma'_x \times \sigma'_y = 13 \times 4$ mrad^2. The emittance, however, is low as the source radius defined by the extent of the wakefield is only 1 µm. In addition, the beam is spatially coherent, and inherently ultrafast, as it is produced by fs-laser pulses. The LWFA therefore has the potential of providing x-rays of similar peak intensity and quality as that produced by third-generation insertion devices. The major obstacle to it being widely used by the synchrotron community is the repetition rate of the laser. In the above experiment, the Hercules petawatt laser at the University of Michigan was used, which delivers a peak power density of 4.7×10^{19} W cm^{-2}, but has a repetition rate of the order of 0.1 Hz[7]. Even were this increased to, say, 1 kHz, the *average* photon flux from the LWFA

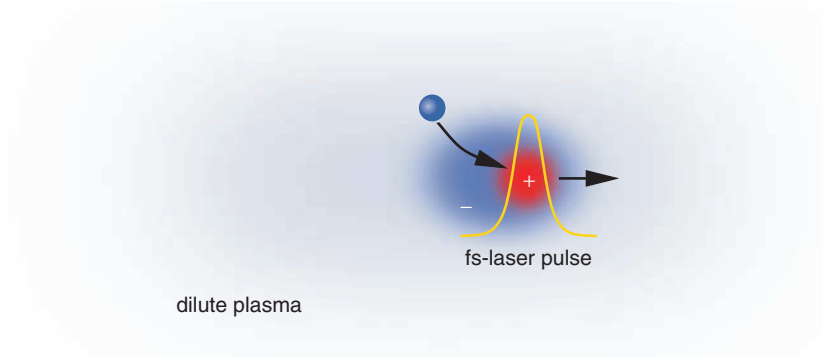

Figure 3.36 *Laser wakefield acceleration. A high-power femtosecond laser pulse travelling through a dilute plasma will locally induce a high degree of charge separation as the electric field of the laser pulse strips electrons from the nuclear cores of the atoms making up the plasma. Because the electrons are lighter, they move away more quickly, resulting in a high degree of charge separation and an associated very strong potential gradient. This wakefield propagates through the plasma with the same group velocity as the laser pulse and will accelerate any electrons that are injected into it.*

[7] The peak power of approximately 70 TW should be compared to the average global power consumption of approximately 10 TW!

would still be at least ten orders of magnitude smaller than that from a synchrotron. In addition, each fs-laser pulse has an energy of approximately 2 J, while the electrical power input required to produce this is much larger still. Running a LWFA at high laser repetition rates would present a formidable problem in power management. It is therefore more probable that LWFAs will find a niche application in the area of coherent and time-resolved studies.

3.9 Concluding Remarks

With the advent of third-generation synchrotron facilities, characterized by their use of insertion devices, particularly undulators, the brilliance available for researchers in the field of x-ray science jumped by some four orders of magnitude. This enormous increase in the number of x-ray photons in tightly focused bundles has revolutionized x-ray techniques, enabling the study of increasingly small structures over shorter and shorter timescales. This has driven a parallel thrust in the sophistication of x-ray optical components and detectors, which we discuss in some detail in the following chapter.

The number of dedicated, third-generation sources has burgeoned over the last decade and has opened the field of synchrotron physics, transforming it into a multidisciplinary enterprise, nowadays attracting some hundred-thousand scientists across the broadest spectrum of the natural sciences.

And now the fourth generation of x-ray sources is beginning to mature, with the first true hard x-ray free-electron laser already delivering first results. This ability to combine the time-resolution previously reserved only to the most sophisticated optical lasers with the atomic resolving power of x-rays will open a bright and new vista in the science of femtosecond dynamics on the sub-Angstrom scale.

References

[1] H. Winick, *Synchrotron radiation sources – a primer*. World Scientific, 1995.
[2] H. Wiedemann, *Synchrotron radiation*. Springer, 2003.
[3] P. Duke, *Synchrotron radiation – production and properties*. Oxford Science Publications, 2009.
[4] R. Clarke, "Incoherent sources: Synchrotrons," in *Encyclopedia of Modern Optics*, pp. 217–224, Elsevier, 2005.
[5] K. Halbach, "Physical and optical-properties of rare-earth cobalt magnets," *Nucl. Instrum. Methods Phys. Res*., 187, pp. 109–117, 1981.
[6] K. Halbach, J. Chin, E. Hoyer, H. Winick, R. Cronin, J. Yang, and Y. Zambre, "A permanent-magnet undulator for SPEAR," *IEEE Trans. Nucl. Sci*., 28, pp. 3136–3138, 1981.
[7] D. H. Bilderback, P. Elleaume, and E. Weckert, "Review of third and next generation synchrotron light sources," *J. Phys. B. At. Mol. Opt. Phys*., 38, pp. S773–S797, 2005.
[8] A. M. Kondratenko and E. L. Saldin, "Generation of coherent radiation by a relativistic electron beam in an undulator," *Part. Accelerators*, 10, pp. 207–216, 1980.

[9] R. Bonifacio, C. Pellegrini, and L. M. Narducci, "Collective instabilities and high-gain regime free-electron laser," *Opt. Commun*., 50, pp. 373–378, 1984.
[10] P. Emma *et al*. "First lasing and operation of an ångstrom-wavelength free-electron laser," *Nature Phot*., 4, pp. 641–647, 2010.
[11] J. Feldhaus, J. Arthur, and J. B. Hastings, "X-ray free-electron lasers," *J. Phys. B. At. Mol. Opt. Phys*., 38, pp. S799–S819, 2005.
[12] G. Margaritondo and P. R. Ribic, "A simplified description of X-ray free-electron lasers," *J. Synchrotron Rad*., 18, pp. 101–108, 2011.
[13] M. Ferray, A. L'Huillier, X. F. Li, L. A. Lompre, G. Mainfray, and C. Manus, "Multiple-harmonic conversion of 1064-nm radiation in rare-gases," *J. Phys. B. At. Mol. Opt. Phys.*, 21, pp. L31–L35, 1988.
[14] D. Xiang and G. Stupakov, "Echo-enabled harmonic generation free electron laser," *Phys. Rev. ST Accel. Beams*, 12, p. 030702, 2009.
[15] Courtesy David Dowell, Paul Emma, Yiping Feng, Clive Field, Alan Fisher, Jo Frisch, Heinz-Dieter Nuhn, (SLAC National Accelerator Laboratory); Daniel Ratner (Stanford University); Roy Clarke (University of Michigan); and Steve Durbin (Purdue University).
[16] R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu, "Potential for biomolecular imaging with femtosecond X-ray pulses," *Nature*, 406, pp. 752–757, 2000.
[17] H. N. Chapman *et al*. "Femtosecond imaging with a soft x-ray free-electron laser," *Nature Physics*, 2, pp. 839–843, 2006.
[18] H. N. Chapman *et al*. "Femtosecond x-ray protein nanocrystallography," *Nature*, 470, pp. 73–77, 2011.
[19] Z. R. Huang and R. D. Ruth, "Laser-electron storage ring," *Phys. Rev. Lett*., 80, pp. 976–979, 1998.
[20] S. Kneip *et al*. "Bright spatially coherent synchrotron x-rays from a table-top source," *Nature Physics*, 6, pp. 980–983, 2010.
[21] M. Bech, O. Bunk, C. David, R. Ruth, J. Rivkin, R. Loewen, R. Feidenhans'l, and F. Pfeiffer, "Hard X-ray phase-contrast imaging with the Compact Light Source based on inverse Compton X-rays," *J. Synchrotron Rad*., 16, pp. 43–47, 2009.
[22] T. Tajima and J. M. Dawson, "Laser electron accelerator," *Phys. Rev. Lett*., 43, pp. 267–270, 1979.
[23] J. B. Rosenzweig, D. B. Cline, B. Cole, H. Figueroa, W. Gai, R. Konecny, J. Norem, P. Schoessow, and J. Simpson, "Experimental observation of plasma wake-field acceleration," *Phys. Rev. Lett*., 61, pp. 98–101, 1988.
[24] A. Pukhov and J. Meyer-ter-Vehn, "Laser wake field acceleration: the highly non-linear broken-wave regime," *Appl. Phys. B*, 74, pp. 355–361, 2002.

4

Beamlines

4.1 Introduction

For users, the beamline is the part of a synchrotron facility of primary interest. Although for some, a beamtime at a synchrotron consists of bringing along a sample, placing it in the beam and starting an experiment (or getting the beamline scientist to do likewise), and then leaving again, clutching a CD ROM, USB stick or hard disk replete with experimental data,[1] many experiments require a deeper level of understanding of the beamline components in order to optimize the measurements. This can range from being able to change the x-ray photon energy, to preparing sophisticated time-resolved experiments that use triggering from the time structure of the electron bunches in the storage ring.

In this chapter, we will look at the main beamline components, in the same order as encountered by the photons emerging from the storage ring. Not all beamlines contain all the components described below, although the general features in Figure 4.1 are fairly standard.

4.2 Front End

The 'front end' of a beamline consists of the components in the ring tunnel, after the insertion device or bending magnet source. It functions are (a) to monitor the position of the photon beam; (b) to define the angular acceptance using a beam defining aperture; (c) to block, when required, some part of the x-rays and the Bremsstrahlung radiation; (d) to filter out, if necessary, the portion of the source spectrum in the soft x-ray region, which would otherwise interact very strongly with the beamline components; and (e) when possible, to isolate the beamline vacuum from the storage-ring vacuum.

[1] Indeed, it is becoming more common in protein crystallography that users simply send samples to the facility from their home institute for investigation, and await the results via Internet.

An Introduction to Synchrotron Radiation: Techniques and Applications, First Edition. Philip Willmott.

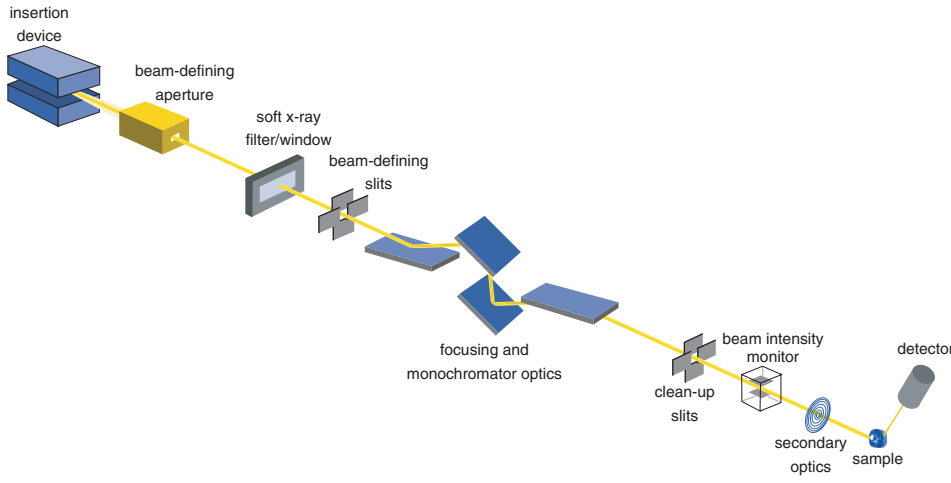

Figure 4.1 *Typical components found at an x-ray synchrotron beamline.*

4.2.1 Beam-Position Monitors

Beam-position monitors (BPMs) measure the lateral position of the x-ray beam at different positions along the beamline from the front end to the experimental hutch. There are several possible designs, with typical accuracies of the order of a micron. Wire monitors, shown in Figure 4.2(a), consist of a thin wire of, for example, tungsten, and can be scanned across the beam to record its profile in one direction, by measuring the photocurrent induced by the x-rays. Their resolution is, of course, limited by the diameter of the wire, and is typically of the order of 5 μm. Other common designs include blade BPMs (Figure 4.2(b)), in which four metallic blades which intersect with the outer halo of the x-ray beam produce four independent photon currents, from which the position in two dimensions can be ascertained using comparators; and CVD diamond BPMs (Figure 4.2(c)), in which a set of photocurrents from an array of ultrathin metallic electrodes deposited on thin and transparent synthetic diamond discs can provide not only the beam position, but also a two-dimensional profile of the beam intensity [1].

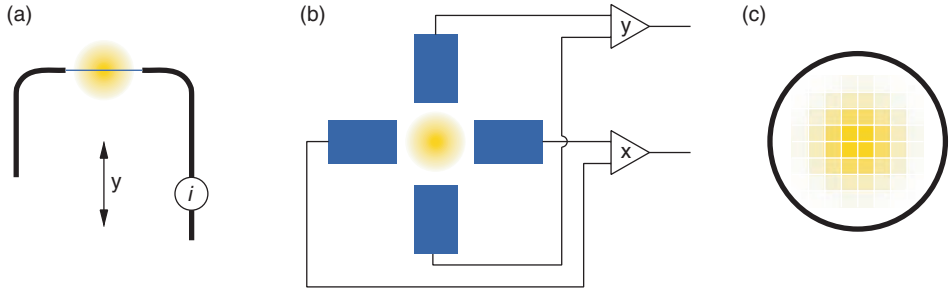

Figure 4.2 *Different designs for beam-position monitors: (a) a simple 1-D wire monitor, (b) a 2-D blade monitor, and (c) a diamond CVD profile monitor.*

4.2.2 Primary Aperture and Front-End Slits

The size of storage rings normally means that the nearest position to the x-ray source where components can be placed in a beamline branch is of the order of $l_{\min} = 10\,\text{m}$. Because the synchrotron radiation emitted from the insertion device or bending magnet has a finite size and divergence, as already discussed in Chapter 3, it will have a lateral size at this minimum distance of

$$l_{h,v} = l_{\min}\,\theta_{h,v}, \tag{4.1}$$

where $\theta_{h,v}$ is the full-width at half-maximum (FWHM) opening angle of the source in the horizontal (vertical) direction, given in radians. We have seen in the previous chapter that θ_v is typically of the order of 0.05 mrad for undulators. Hence, the typical size of undulator radiation at a distance of 10 m is 5 mm (H) by 0.5 mm (V).

The spectrum of synchrotron radiation away from the central axis peaks at lower energies, hence the outer parts of the radiation cone are more strongly absorbed by matter (see Figures 2.16 and 3.6), a phenomenon known as 'beam hardening'. It is therefore important in hard x-ray beamlines to define the beam using a fixed aperture at the front end, not just for purposes of collimation, but also to suppress the number of lower-energy photons incident on the beamline components, which could otherwise be destroyed. Because the power that needs to be absorbed can be considerable, the aperture is often in the form of a water-cooled rectangular funnel, so that the power is dissipated over a large area (see Figure 4.3).

The opening angle of bending-magnet sources in the horizontal direction is very large and in fact given by the angle subtended by the bending magnet itself (see Figure 3.13). Here, the need to restrict the beam to define the divergence is even more important.

In addition to the primary aperture, the divergence of the beam can be further reduced in the front end by two pairs of adjustable (and usually water-cooled) slits. These are normally protected from the high-flux heat-load experienced by the beam-defining aperture by one or more filters used to further remove soft x-rays. We describe these now.

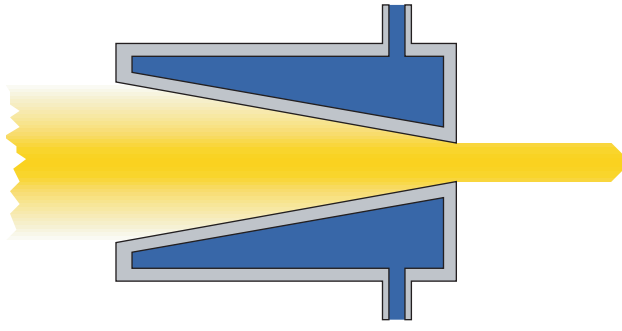

Figure 4.3 *Beam-defining primary apertures often have a rectangular cone shape, in order to distribute the thermal load over a larger area.*

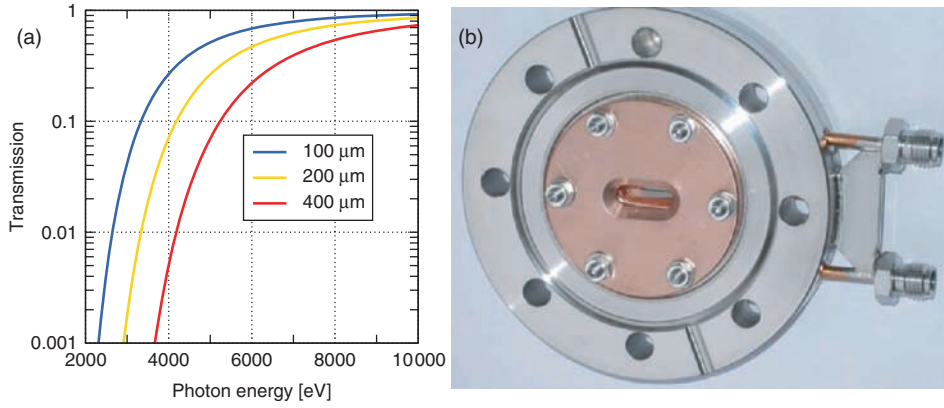

Figure 4.4 *(a) The transmission curves for various thicknesses of diamond filters, assuming a density of 3.5 g cm^{-3}. (b) A 100 µm-thick, ultra-high vacuum-compatible, water-cooled diamond window, used to remove soft x-rays and provide isolation between the storage ring vacuum and the beamline vacuum. Courtesy of Max Kleeb, Paul Scherrer Institut.*

4.2.3 Low-Energy Filters

The sources at high photon-energy beamlines nonetheless produce a large flux of soft x-rays, which, below approximately 5 keV, interact strongly with matter and thereby destroy beamline components. In order to filter out these soft-x-rays, one needs to use low-Z material which is robust, able to withstand high temperatures, have a good thermal conductivity and is preferably highly crystalline. The only realistic candidate is carbon, either in the form of graphite or glassy carbon (especially if the filter has to be physically large), or synthetically grown diamond, in the form of thin sheets (see Figure 4.4(a)).

Diamond has the highest known thermal conductivity of any material,[2] and is hence ideally suited to conduct away the absorbed power – commonly, such diamond filters are brazed onto a water-cooled copper block for this purpose (Figure 4.4(b)).

After the removal of the soft x-ray portion of the spectrum, other materials can be placed in the beam. So, for example, it is common to use thin beryllium windows[3] to isolate the vacuum of the storage ring from that of the beamline. It also further reduces mechanical distortions (or 'thermal bumps') due to localized thermal loads on the optical components used for focusing and monochromating. Caution must be exercised in using beryllium – firstly, the oxide of beryllium is exceedingly toxic if it enters the blood stream. Secondly, beryllium components are invariably polycrystalline, producing

[2] In contrast to other good thermal conductors such as silver, copper and aluminium, which conduct heat primarily via the kinetic energy of the conducting electrons, diamond conducts heat via quantized vibrations of the lattice, known as phonons. This is particularly efficient in diamond, because of the very high strength of its covalent bonds. The vibrational energy levels lie in a deep parabolic (that is, symmetric) potential well, where scattering of the phonons, given by *odd* terms in the potential-well description, is exceedingly weak. The phonons therefore move ballistically, resulting in very efficient transport of heat. The thermal conductivity of diamond is, at approximately 20 W cm^{-1}K$-$1, some four to five times larger than that of copper.

[3] A 0.5 mm-thick sheet of Be has a transmission of 70% at 5 keV, which rises to 95% at 10 keV.

diffraction rings, while microscopic inhomogeneities and inclusions produce significant and undesired diffuse signal in the forward direction.

4.3 Primary Optics

Optical components for x-rays are quite different than those used for visible light. The primary reason is that, as we have seen in Chapter 2, the refractive index of x-rays is very close to *and slightly below* unity. In addition, soft x-rays are very strongly absorbed by condensed matter, and can even be strongly attenuated by gas (e.g. air at one atmosphere, see Figure 4.5).

This makes it difficult to bend or redirect x-rays. The phenomena of total external reflection, refraction and diffraction are all used in x-ray optics to achieve these ends.

An x-ray optical system consists normally of a monochromator and one or two x-ray mirrors, plus sundry slits and filters. It performs up to three tasks; (a) photon energy selection using the monochromator; (b) optimization of the energy resolution using a collimating mirror upstream of the monochromator; (c) refocusing using either curved mirrors or gratings, or via bending one of the dispersing elements within the monochromator. An example of one possible configuration is shown schematically in Figure 4.6.

4.3.1 X-ray Mirrors

As we have already hinted, x-ray mirrors are in general very different beasts than the piece of aluminium-coated glass we peer into in the bathroom in the morning. As we have already discussed in Chapter 2, the reflectivity of x-rays drops catastrophically above the critical angle (see Figure 2.13). A useful rule-of-thumb relating the maximum allowable incident grazing angle to the energy of the synchrotron light is

$$\alpha_{max}[\text{degrees}] = \lambda[\text{Å}]/10. \tag{4.2}$$

This approximation derives from Equations (2.15) and (2.18) and assumes that all condensed matter has the same average electron density of 0.34 Å^{-3}, which is fairly accurate

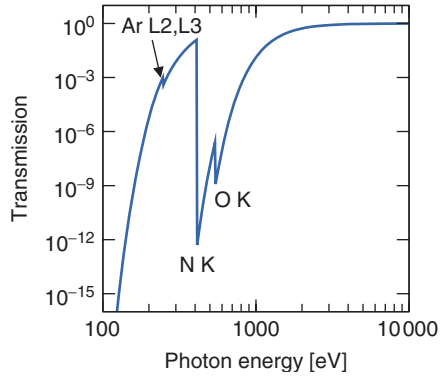

Figure 4.5 *The transmission of x-rays through 1 cm of air at room temperature and pressure. The composition of air was taken to be 78.5% N_2, 21% O_2, 0.03% CO_2 and 0.94% Ar.*

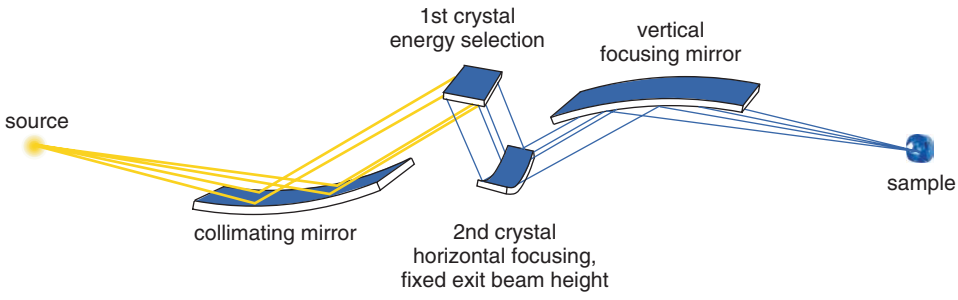

Figure 4.6 *Schematic of an x-ray optical system, consisting of a double-crystal monochromator and two mirrors. Note that the second crystal of the DCM can also be bent to provide horizontal focusing.*

for elements in the middle of the periodic table (for example, silicon, a commonly used x-ray mirror material, has an average electron density of $0.7\,\text{Å}^{-3}$). Hence, although hard ultraviolet photons at 50 eV (≈ 250 Å) are efficiently reflected up to a few tens of degrees incident angle, hard x-rays at around 1 Å can only be efficiently reflected at grazing angles of around 0.1° or less.

This highlights a major technical obstacle to producing high-quality reflective optics for hard x-rays. The extent of the mirror that is illuminated by a grazing-incidence beam is h/α, where h is the beam height perpendicular to the mirror surface, and the incident angle α is given in radians. Hence, for a typical beam height of 1 mm and an incident angle of 0.2°, the mirror length must be at least 290 mm. The atomic roughness of the mirror cannot exceed a few Angstroms for the mirror to be usable, while longer-range 'wobbles' in the flatness, given by the figure of merit known as the 'slope error', should be of the order of 1 µrad or less (Figure 4.7) in order to preserve the fidelity of the beam profile after reflection. This presents a formidable technological challenge, and mirrors can be the major contributor to an increase in the apparent size and consequently lower effective brilliance of the source.

The theoretically most perfect reflective focusing element is a two-dimensional parabolic mirror, which, ignoring diffraction effects, can focus a parallel beam to a

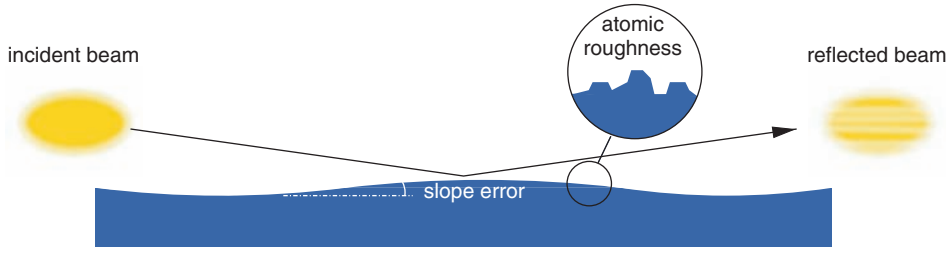

Figure 4.7 *A reflecting flat x-ray mirror has a residual 'wobbliness' to it, referred to as the slope error, which causes the beam profile to become more irregular. In addition, the roughness on an atomic scale must not exceed a few Angstroms.*

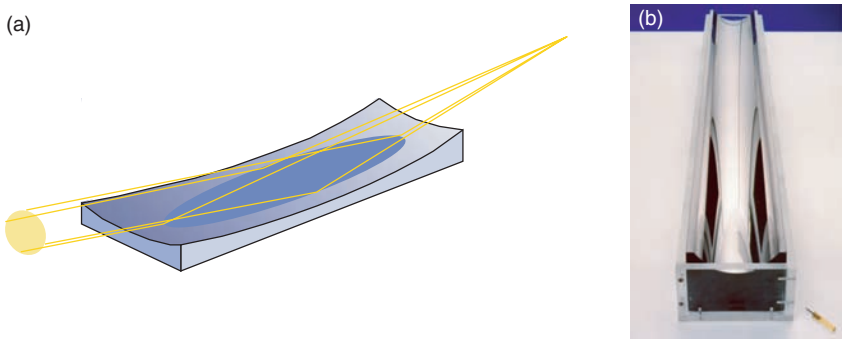

Figure 4.8 *(a) A two-dimensional parabolic mirror can focus a parallel beam to a point. Because the beam must approach the mirror at an angle approximately equal to, or below, the critical angle for total external reflection, the size of such mirrors must be of the order of 0.5 m. The shaded area on the surface represents where the x-ray beam strikes the mirror, and is commonly referred to as the 'footprint' of the beam. (b) A toroidal mirror, as used at the ID09B white-beam time-resolved beamline at the ESRF. The substrate is graphite and is coated with SiC. The average surface roughness is less than 2 Å. Courtesy Michael Wulff, European Synchrotron Radiation Facility.*

point (see Figure 4.8(a)). In reality, such surfaces are very hard to engineer, especially if one considers how large the surface needs to be.

A simpler engineering task is to make the mirror toroidal, that is, its shape is like a section of the inner surface of a bicycle tube (see Figure 4.8(b)). This approximates a parabolic surface to a good degree, and often the resulting 'spherical aberrations', that lead to a smearing out of the focus, can be kept acceptably small.

A still simpler solution is to divide the jobs of horizontal and vertical focusing between separate optical elements – by using two cylindrical mirrors, horizontal and vertical focusing can be achieved independently (see Figure 4.9). This arrangement is known as

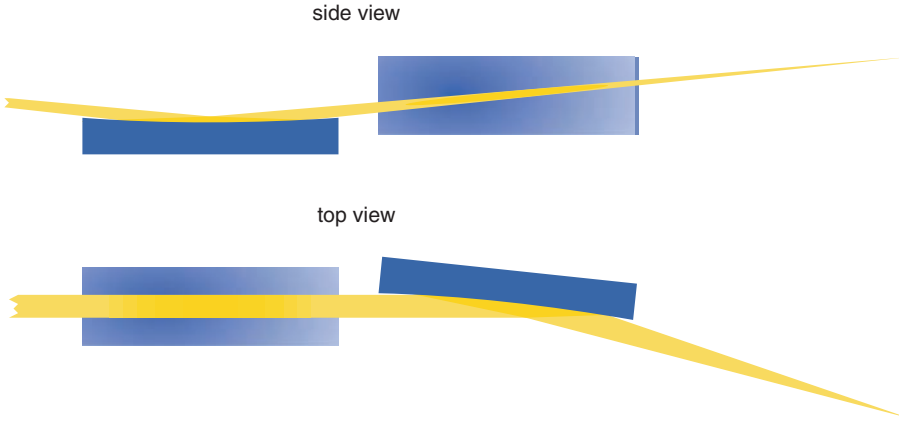

Figure 4.9 *By using two cylindrical mirrors in series, vertical and horizontal focusing can be achieved independently in the so-called 'Kirkpatrick-Baez' mirror configuration.*

a Kirkpatrick–Baez mirror, after its inventors, or, more simply, a K–B mirror [2]. It has the advantages that horizontal and vertical focusing are decoupled from each other, and that the focal lengths are independent of the photon energy. A disadvantage is that the length of the mirror system is effectively doubled.

4.3.2 Mirror Focal Lengths – The Coddington Equations

For a glancing incidence angle θ, the 'meridional' (vertical) and 'sagittal' (horizontal) focal lengths of a curved mirror (shown schematically in Figure 4.10) are given by

$$f_m = \frac{R_m \sin\theta}{2}, \tag{4.3}$$

$$f_s = \frac{R_s}{2\sin\theta}. \tag{4.4}$$

For a source–mirror distance p and a mirror-image distance q, the standard 'lensmaker' equation

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{q} \tag{4.5}$$

yields the Coddington equations for meridional and sagittal focusing

$$R_m = \frac{2}{\sin\theta}\left(\frac{pq}{p+q}\right), \tag{4.6}$$

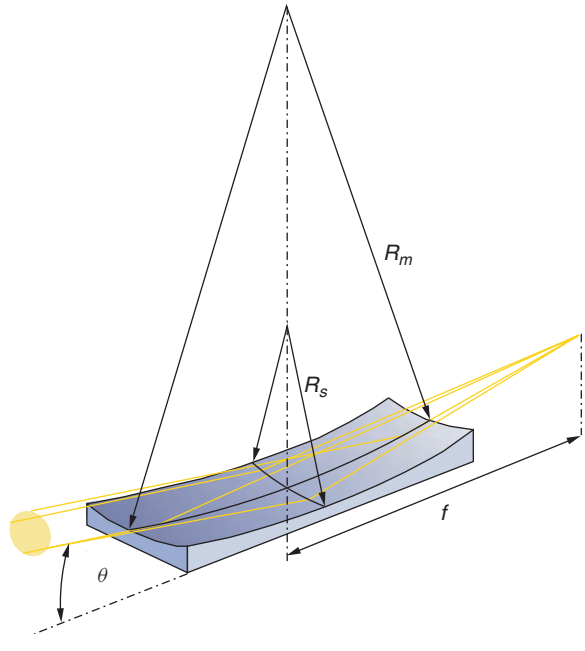

Figure 4.10 *The meridional and sagittal bending radii R_m and R_s of a curved mirror focusing a parallel x-ray beam both vertically and horizontally to a focal point f.*

$$R_s = 2\sin\theta\left(\frac{pq}{p+q}\right). \tag{4.7}$$

Remember that θ is, in general, very small for x-rays, as it is limited by α_c, the critical angle for total external reflection. For $\theta = 0.15^\circ$ (≈ 2.5 mrad), the meridional radius R_m is $\theta^2 \sim 10^5$ times larger than R_s and may be of the order of several kilometres, while R_s is typically only a few centimetres. Note also that if either p or q is infinite (corresponding, respectively, to a parallel incoming beam and a parallel outgoing beam), then the fraction $pq/(p+q)$ reduces to q and p, respectively, and we regain the mirror focal length equations, given above. Finally, the subtended meridional half-angle of the mirror is $l_{\text{mirror}}/2R_m$, where $l_{\text{mirror}} \sim 50$ cm is the length of the mirror. This is therefore of the order of 50 µrad, some two orders of magnitude smaller than typical critical angles for total external reflection. Hence, one can ignore any variations in reflectivity over the length of a curved mirror on which a perfectly parallel beam is incident.

4.3.3 Monochromators

Experiments performed at synchrotrons normally require a well-defined photon energy (Laue diffraction, described in Chapter 5, being one notable exception), and may also demand that the photon energy is changed (as in multiwavelength anomalous diffraction, MAD), or that it is scanned (as, for example, in x-ray absorption spectroscopy, XAS).

In general, radiation from insertion devices, even from a low-K undulator, is not sufficiently monochromatic to be used as a source for experiments without further energy dispersion, or monochromation. This is normally achieved by diffraction. For longer wavelength radiation, manufactured periodic gratings are used, in which the separation between grating lines can be several microns down to a few tens of nm, depending on the wavelength range of interest. For wavelengths shorter than a few nanometres, one uses instead crystal monochromators, which exploit the 'natural' gratings of crystal planes, with periodicities measured in Angstroms, or alternatively, multilayer monochromators which let through a larger bandwidth of wavelengths than do crystal monochromators, with a consequent increase in transmitted flux.

4.3.3.1 Grating Monochromators

Diffraction gratings used for synchrotron radiation are normally reflection gratings, at least for photon energies above the visible region of the electromagnetic spectrum, although a small fraction do utilize the transmitted beam. Gratings consist of a surface covered with a periodic array of many parallel grooves. There are several types of grating profiles and geometries. In particular, the most common profiles are lamellar, sinusoidal and blaze, shown schematically in Figure 4.11, while the geometrical surface of the grating might be planar or concave (spherical, elliptical or toroidal). Concave

Figure 4.11 *Three common grating profiles used to disperse ultraviolet light and soft x-rays.*

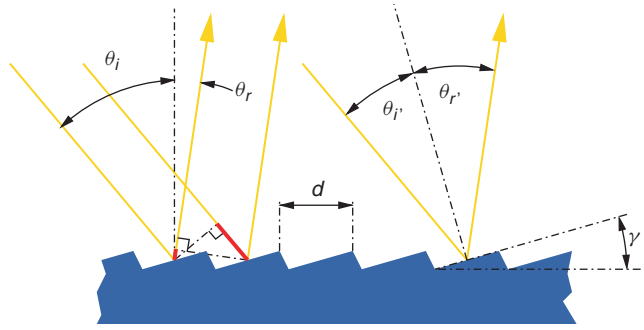

Figure 4.12 *Schematic of a blaze grating, showing the groove periodicity d, incident angle θ_i, reflected angle θ_r, and 'blaze' angle γ. The optical path difference between rays reflected by adjacent facets is the difference of the lines shown in red and is given by Equation (4.8).*

gratings are employed in order to combine focusing with dispersion, thereby dispensing with focusing mirrors. It lies beyond the scope of this book to detail all possible configurations, for which the reader is referred to reviews in the literature [3–5]. Here, we will consider one of the most efficient grating types, the so-called 'blaze' grating, in order to discuss the general properties of grating monochromators.

A schematic of a blaze grating and the relevant angles and parameters is shown in Figure 4.12. The optical path difference between reflections from adjacent facets is

$$d(\sin\theta_i - \sin\theta_r), \tag{4.8}$$

whereby d is the grating period (or 'pitch'), and θ_i and θ_r are the incident and reflected angles, respectively, relative to the *average* grating surface.[4] For constructive interference, Equation (4.8) must equal an integer number of wavelengths, $m\lambda$. The intensity of the monochromated light will be highest when both Equation (4.8) is satisfied, and when its direction is equal to that of light specularly reflected from the individual facets, that is, when $\theta_{i'} = \theta_{r'}$. For a given incident angle θ_i and order m, this double condition is satisfied for only one wavelength, known as the 'blaze' wavelength of the grating.

The narrowness of the range of wavelengths that satisfies the diffraction condition depends on how many facets are illuminated by the incoming beam. Consider the so-called Argand diagram of Figure 4.13. It should be immediately clear that the more facets are involved, the sharper is the drop-off from constructive interference – the change in the (squared) amplitude for the 4 facets in Figure 4.13 is more gradual than that involving eight facets. In mathematical terms, it can be shown that the resolving power, defined as the ratio of the wavelength λ to the bandwidth of the interference maxima $\Delta\lambda$, is given by

$$\frac{\lambda}{\Delta\lambda} = mN_f\,(\sin\theta_i + \sin\theta_r)\,, \tag{4.9}$$

whereby N_f is the number of (equally illuminated) facets. N_f can be of the order of several thousand, hence a resolution of the order of 10^4 is readily achievable. Gratings

[4] Equation (4.8) is sometimes expressed as $d(\sin\theta_i + \sin\theta_r)$, whereby the sign of the angles is positive for a clockwise rotation relative to the grating normal, and negative for an anticlockwise rotation. In the above, the absolute values are assumed.

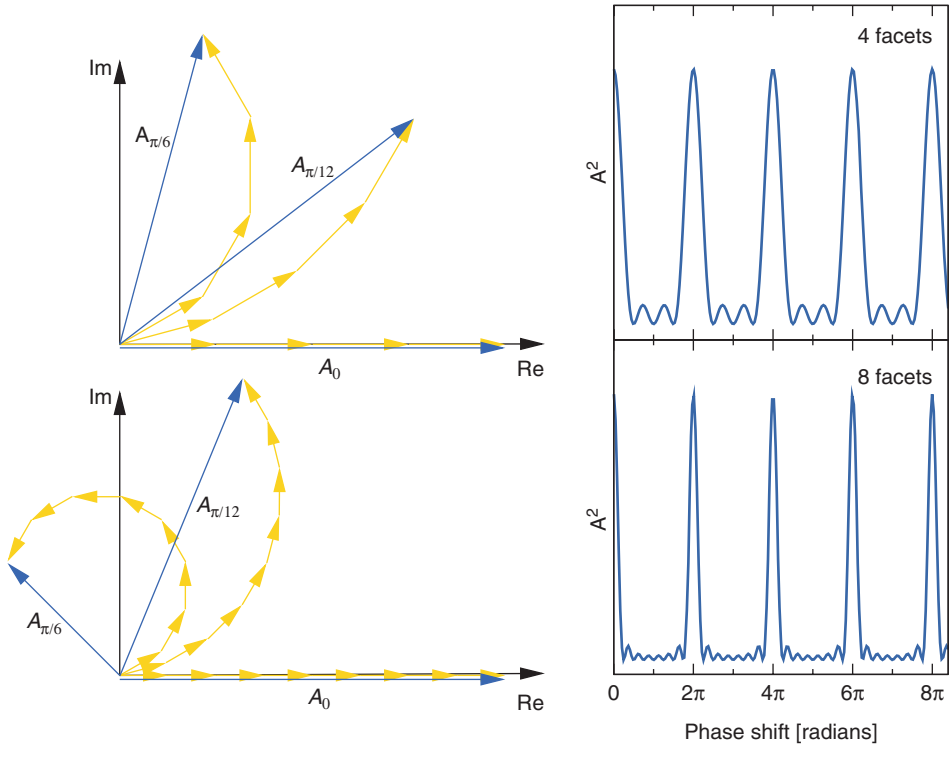

Figure 4.13 *Interference between grating facets. Left: Argand diagram showing how the amplitude changes with phase shift between adjacent facets. The vector sum of reflections off 4 (top) and 8 (bottom) facets are shown for phase shifts of 0, $\pi/12$ and $\pi/6$. Right: the change in intensity (A^2) as a function of phase shift, ignoring the drop off in intensity away from specular reflection off the facets.*

with lower numbers of lines per mm (the units in which this property is normally expressed) have a larger bandwidth and therefore provide more flux, while gratings with higher line densities provide better resolution, though at the cost of the number of photons available for experimentation.

4.3.3.2 Crystal Monochromators

As we have already intimated, the size of the separation of scatterers in diffraction phenomena needs to be comparable to the wavelengths that are being diffracted. Hence, crystalline solids can provide the regular array of scattering centres for hard x-rays available at synchrotron facilities, with wavelengths typically between 0.05 and 6 Å.

A crystal monochromator uses Bragg's law

$$m\lambda = 2d\ \sin\theta \tag{4.10}$$

to filter out a narrow band of radiation. Here, $m = 1, 2, 3, \ldots$ is the order, d is the lattice spacing, and θ is the angle of incidence on the relevant crystal planes. We will discuss the

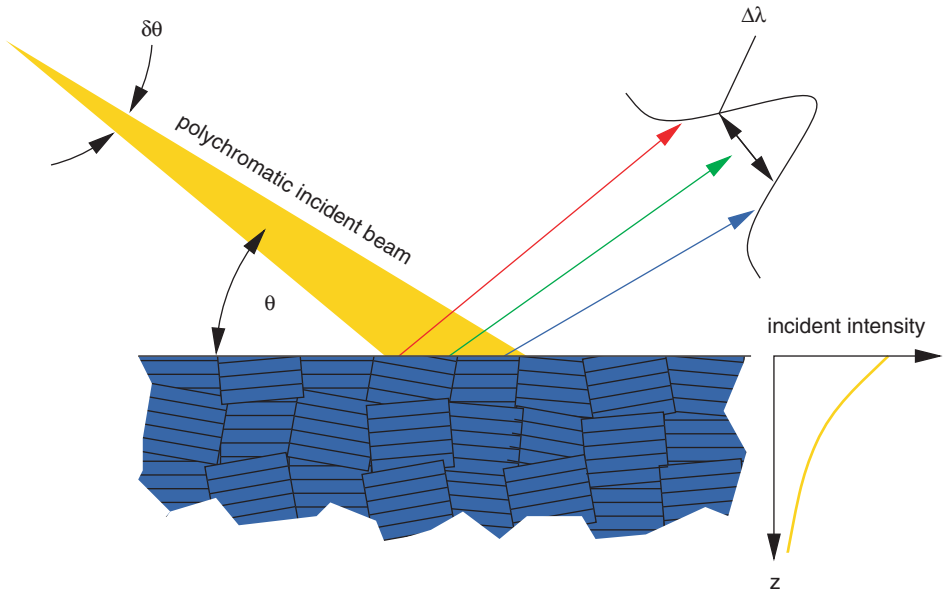

Figure 4.14 *Factors effecting the bandwidth of radiation diffracted by a crystal include the degree of collimation of the incident beam (δθ), the crystal perfection ('mosaicity') and the depth to which x-rays penetrate the crystal (the extinction and absorption depths).*

physics behind Bragg's law in more detail in Chapter 5. For the meantime, it is immediately clear that Bragg's law predicts that if one were to have a perfect, non-absorbing and very weakly scattering crystal of infinite depth, and a perfectly collimated incoming polychromatic beam, the bandwidth of the selected radiation would be infinitely narrow. However, residual beam divergence, slightly misaligned crystallites making up the (imperfect) single crystal (referred to the 'mosaicity') and finite absorption and extinction depths,[5] all contribute to a given crystal always having a monochromating bandwidth $\Delta\lambda$ greater than zero (see Figure 4.14).

The choice of crystal is dictated by the available crystal quality, the ability of the crystal to dissipate the thermal load associated with monochromation, and its resistance to radiation damage. By far the most commonly used material is silicon, and to a lesser extent also germanium. Silicon and germanium have reasonable thermal conductivities, and can be efficiently cooled by liquid nitrogen, in order to minimize the mechanical strain induced by a local thermal 'bump', where the incoming, polychromatic, beam impinges on the crystal surface. Importantly, they can be produced as near-perfect single crystals with linear dimensions of tens of centimetres, thanks to their use in the semiconductor industry, where large and exceedingly perfect wafers are produced en masse.

In the last decade, it has become possible to synthesize plates of diamond of high crystalline quality. The very high thermal conductivity of diamond makes it particularly

[5] The absorption depth is the depth to which the incident radiation is attenuated by $1/e$ due to photoabsorption, and is equal to $1/\mu$ in Equation (2.24). The extinction depth, on the other hand, is associated with attenuation due to *elastic* scattering of the incident beam by the electrons in the crystal, and is markedly smaller when the Bragg condition is met (Equation 4.10).

attractive for the use in high-flux, third-generation beamlines, where heat loads can be substantial. The cost of synthetic diamond single crystals of sufficient size means that diamond-based monochromators are still a rarity [6, 7]. With the advent of x-ray FELs, however, their future implementation is assured.

The two most common designs for crystal monochromators are the double-crystal monochromator (DCM) and the channel-cut monochromator (CCM).

Figure 4.15 shows a typical geometry for a double-crystal monochromator. The polychromatic light entering the monochromator is diffracted using a highly perfect single crystal. This first crystal absorbs that radiation which does not satisfy the Bragg condition, which normally corresponds to well over 99.9% of the flux. It is therefore actively cooled (using water or liquid nitrogen) and can additionally be dynamically flexed to compensate for the thermal bump caused by the high heat load.

The residual divergence of the radiation entering the DCM, crystalline imperfections in the monochromator crystals and their natural extinction depth, all contribute to there being a band of energies (or wavelengths) that are transmitted by the DCM. This beam is then allowed to impinge on a second crystal, which, because of the much reduced total

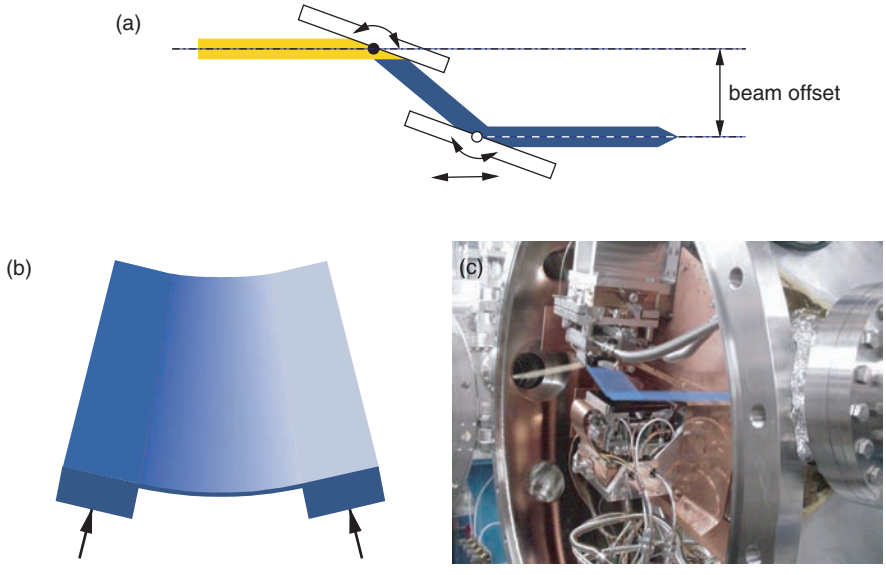

Figure 4.15 *Double-crystal monochromators. (a) The geometry of a DCM. The first crystal monochromates the incoming polychromatic light, while the second crystal redirects the monochromated beam parallel to the incoming beam. In order to keep the offset between the incoming and exit beam height constant for all photon energies (and monochromator crystal angles), the horizontal separation between the two crystals must be variable. (b) The second crystal can be flexed to focus the beam in the horizontal plane. The bending radius of the crystal depends on the angle of incidence (θ) and the desired focal position, and can be calculated using Equation (4.7). (c) A view of the DCM at the Materials Science beamline, Swiss Light Source. The path of the incident 'white' beam and the monochromated beam (in blue) is shown as semitransparent rays.*

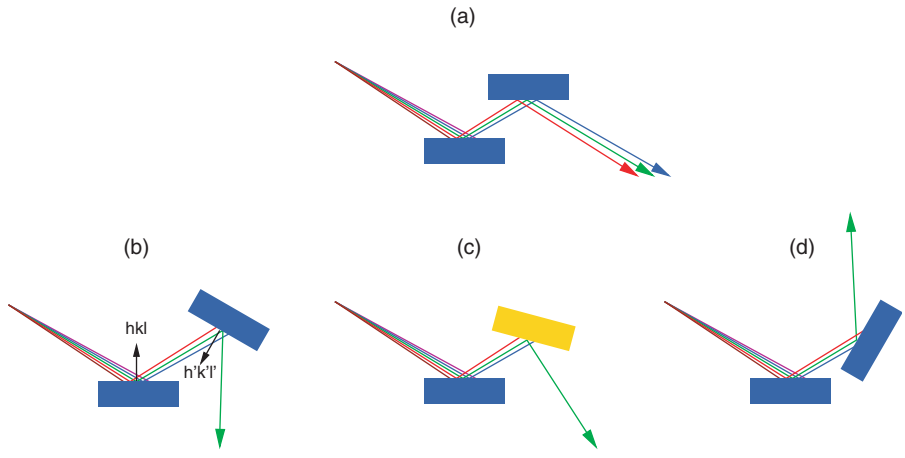

Figure 4.16 *Nondispersive (a) and dispersive (b)–(d) modes in DCMs. In case (a), the incident angle for any particular component of the bandwidth is the same for the first and second crystals. If another crystal orientation (b), another crystal type (c), or the same sense of reflection (d) is used for the second crystal, the Bragg condition becomes more stringent – the bandwidth will become narrower and the flux will decrease.*

flux, does not need to be cooled.[6] Horizontal focusing is also possible using the second crystal, by bending the second crystal sagittally, that is, with the plane of curvature perpendicular to the direction of the beam propagation.

According to the relative geometries, crystallographic orientation, and/or crystal types, the radiation emerging from the monochromator can be more or less monochromatic. Consider Figure 4.16. We assume that the crystal is perfect and that absorption and extinction are negligible. We allow, however, the incoming polychromatic beam to have a divergence of $\delta\theta$ in the plane containing both it and the crystal normal. According to Equation (4.10), the crystal will select longer wavelengths from that part of the beam that impinges more steeply (larger θ) on it than that part of the beam that strikes the crystal at a shallower angle. For example, a beam with a divergence $\delta\theta = 0.1$ mrad incident on a perfect Si(111) crystal ($d_{111} = 3.1356$ Å) at $\theta = 14^\circ$ will be filtered for a wavelength of 1.51714 ± 0.00061 Å, or alternatively, 8172.2 ± 3.3 eV.

In the normal 'nondispersive' configuration, shown in Figure 4.16(a), the monochromacity of the beam is not improved after the second crystal, as long as that second crystal is made from the same crystalline material and has the same orientation as the first. In this case, the divergence of the beam defined by diffraction at the first crystal is then the same as that after diffraction at the second.

However, if there is a desire to further improve the monochromacity, this can be achieved in one of three manners (see Figure 4.16(b)–(d)) – by the second crystal either

[6] Note, however, that if there is a temperature difference between the first (cooled) and second (uncooled) crystal, this must be taken into account in calculating the Bragg angle, as the crystals' lattice constants d depend on the temperature. It can be shown that, for a monochromator using two identical crystal types and crystal orientations, the difference in their Bragg angles (in radians) is $\delta\theta = (\Delta d/d)\tan\theta$, where $\Delta d/d$ is the fractional difference in the lattice constants of the two crystals, due to thermal expansion.

having a different crystal orientation than the first; by using a crystal of a different material (assuming it has a different lattice constant, which is invariably the case); or by the second crystal deflecting the beam in the same sense as the first crystal (both clockwise or both anticlockwise). Needless to say, all of these dispersive arrangements achieve this improved monochromacity at the expense of flux. This is succinctly explained using so-called 'Du Mond' diagrams [8], outside the scope of this text.

In channel-cut monochromators, the nondispersive geometry of Figure 4.16(a) is obtained by cutting the crystal surfaces out of a single, monolithic crystal (see Figure 4.17). This simple design has the advantage that both crystal surfaces can easily be maintained at the same temperature, and that the diffracting planes from the two surfaces are almost perfectly parallel, resulting in their 'automatic' alignment with each other. Only at high photon energies above around 20 keV, can residual strains in the crystal and possible distortions due to the mounting mechanism begin to cause the Bragg condition for the second crystal surface to drift outside the energy width of the beam produced by the first crystal face.

The 'beam offset' is the lateral shift in the x-ray beam between upstream of and downstream from the monochromator (see Figure 4.15(a)). A beam offset of several millimetres to a few tens of millimetres is required, due to the problem of Bremsstrahlung radiation: there is a cone of Bremsstrahlung radiation emanating from the storage ring at the height of the incoming beam, caused by collisions of the electrons with residual gas particles in the ring, which contains gamma-ray photons with exceedingly high energies, indeed up to that of the electrons in the storage ring (measured in GeV!). If this is not blocked it presents a serious health hazard, and will also produce a significant amount of background signal. In order to overcome this, a block of tungsten several centimetres thick is placed in the path of these gamma rays, while an aperture in this block at the height of the outgoing beam allows through the energy-filtered synchrotron radiation. The separation between the Bremsstrahlung and synchrotron radiation needs to be significant to ensure none of the Bremsstrahlung passes through into the experimental hutch.

In the case of CCMs, the beam offset varies with photon energy such that, for a Bragg angle θ and a channel height D, the beam offset is

$$\Delta y = 2D\cos\theta, \tag{4.11}$$

or, in other words, the exit beam height is not constant for all selected photon energies. In addition, the beam impinging on the second surface of the channel drifts more

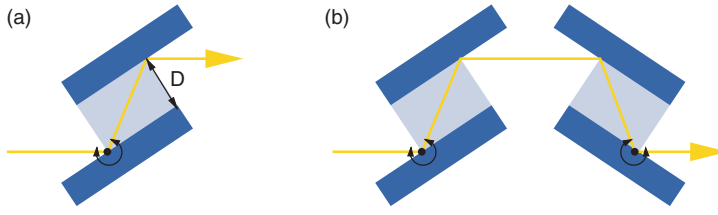

Figure 4.17 *(a) A channel-cut monochromator uses a monolithic crystal (normally Si) in which a channel is cut along a crystallographic plane (e.g. the (111)-plane), creating two opposing and parallel faces separated by a distance D. (b) A so-called four-bounce crystal-monochromator with a constant beam exit height consists of two CCMs in series.*

downstream with increasing energy (decreasing θ) as $D/\tan\theta$. Hence, a large beam offset is associated with a necessarily large channel-cut crystal, especially at higher photon energies.

The problem of a variable exit beam height can be resolved by using two identical CCMs in series, as shown in Figure 4.17(b). Such a configuration is referred to as a 'four-bounce crystal-monochromator' (4BCM). Note, however, that a 4BCM is a dispersive setup (as the x-rays travel from the second to third bounce), which results in lower flux.

4.3.3.3 Multilayer Monochromators

Lastly, so-called multilayers can be used for monochromating hard x-rays. Multilayers are produced by growing periodic layers of two sorts of thin film material, usually by a deposition process called magnetron sputtering. Each period normally consists of a thinner sublayer of high-Z material (e.g. tungsten) and a thicker sublayer of a low-Z (i.e. high transmission) material, such as carbon. Scattering of the incoming x-rays at the interfaces between the two sublayers results in diffraction maxima occurring in exactly the same manner as in crystals, whereby the 'lattice' spacing is not that between atomic planes in a crystal, but the multilayer period Λ. Multilayers can therefore be thought of as 'artificial crystals'. An example of the reflectivity of a W/C multilayer is shown in Figure 4.18.

Because there are relatively few scattering planes (at least compared to the number of atomic scattering planes in a crystal), multilayer monochromators are ideal for beamlines with less stringent requirements regarding the degree of monochromacity, and where photon flux is at a premium. The full-width at half-maximum (FWHM) of a multilayer peak (in radians) is

$$\Delta\theta = \frac{\lambda}{m\Lambda}, \tag{4.12}$$

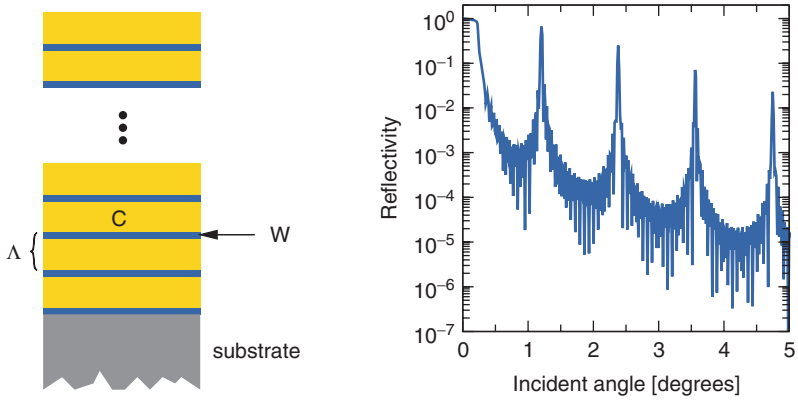

Figure 4.18 *The reflectivity curve for 10 keV photons as a function of incident angle on an ideal W/C multilayer, consisting of $m = 50$ bilayers, each having a thickness of 3 nm, whereby each bilayer is formed by 0.48 nm W and 2.52 nm C.*

whereby λ is the x-ray wavelength and m is the number of periods in the multilattice. In the example of Figure 4.18, this is approximately 0.02°, significantly larger than the natural (Darwin) bandwidth of Si(111), which lies in the region of 10^{-3} degrees. An example of where multilayer monochromators can be ideally used is in x-ray tomography.

4.3.3.4 Higher Harmonics

Referring once more to Equation (4.10), we see that for a fixed $2d \sin\theta$, there are a set of distinct wavelengths λ_1, λ_2, λ_3 and so on, that satisfy the Bragg condition, whereby $m\lambda_m$ is a constant, λ_1 is known as the fundamental, and the others are referred to as higher harmonics.

In general, however, we want only one wavelength (or narrow band of wavelengths), which is normally (though not always) the fundamental. How can we suppress the higher harmonics?

As we saw in Chapter 2, the critical angle for total external reflection increases as the square root of the electron density of the material from which the x-rays are reflecting, and linearly with λ (Equations (2.15) and (2.18)). One can therefore use an incident angle on an x-ray mirror which lies below the critical angle for the fundamental wavelength, but which is above the critical angles for all the harmonics (Figure (4.19)). As one tunes the monochromator to higher fundamental photon energies, one can lower the incident angle on the mirror so that it is below the critical angle for that photon energy and reflecting surface material. Above a certain photon energy, however, the incident angle might become so shallow that the footprint of the beam spills significantly over the front and back edges of the mirror. This of course depends on the size of the profile of the beam and on the mirror length.

In order to extend the range of energies that can be used by a given mirror, one can coat the mirror lengthwise with one or more stripes of different coatings, such as Au, Pt, Rh or Cr – for a given photon energy, heavy materials with high electron densities have steeper critical angles (Equation (2.18)) – and translate the mirror sideways from one stripe material to the other, depending on which photon energy range is desired. So, for example, by changing from an uncoated region of a silicon mirror to one that is coated with Pt, the critical angle can be increased by a factor of 2.7.

Higher harmonics are further suppressed by the fact that, in general, the flux at λ_1/m is usually lower than that at the fundamental λ_1, and also because, for certain crystals (of which Si(111) is an example), reflections of some of the higher harmonics are forbidden,

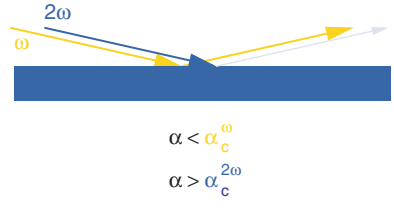

Figure 4.19 *Higher harmonics can be suppressed by adjusting α, the angle of incidence of the x-rays on the mirror, so that only the fundamental frequency has a critical angle that is larger than α.*

due to destructive interference between the scattering atoms within the unit cell of the crystal.[7]

Lastly, a slight detuning of the Bragg angle of the second monochromator crystal will suppress higher harmonics significantly more efficiently than the fundamental wavelength. This is because the width of the diffraction peak (mh, mk, ml) of the mth harmonic (filtering through the wavelength λ_1/m) is narrower than that of the fundamental reflection (h, k, l),[8] hence a slight detuning by rotating the monochromator crystal away from the Bragg condition by an amount $\delta\theta$ causes the flux of the higher harmonic to fall away more severely than for the fundamental. For example, the Darwin widths of Si(111) and its third harmonic Si(333) at 10 keV are 26 and 6.6 μrad, respectively.

4.3.3.5 Double-Crystal Deflectors

Certain experiments require that the incident x-ray beam impinges on the surface of a liquid, and therefore has to be tilted downwards. In experiments such as x-ray reflectivity (XRR, discussed in Section 5.15), the incident angle α must be varied while keeping the liquid sample stationary. This is achieved using a double-crystal deflector (DCD) [9], which is essentially a DCM using dissimilar crystals, such as the geometries shown in Figure 4.16(b) and 4.16(c). Consider Figure 4.20(a). Two crystals X1 and X2 have lattice spacings d_1 and d_2 perpendicular to their optical surfaces. The doubly diffracted beam therefore crosses the axis of the initial beam axis at an angle $\chi = 2(\theta_2 - \theta_1)$, where θ_1 and θ_2 are the Bragg angles for X1 and X2, respectively, at the energy of interest. The entire DCD (i.e. the construction on which the two crystal holders/goniometers are mounted) is allowed to rotate around the initial axis between $\phi = 0$ and 90°, whereby the lower limit of ϕ is when the diffracted beams lie in the same plane as the liquid-sample surface, and the upper limit is when they lie in the same plane as the surface normal. In this manner, the diffracted beam will impinge on the sample sitting at the intercept P of the initial axis and the diffracted beam at an angle

$$\alpha = \arcsin(\sin\chi \sin\phi)\,. \tag{4.14}$$

The change in α as a function of ϕ for different photon energies and using Si(111) and Si(220) crystals is shown in Figure 4.20(b). Note also that γ, the orientation of the footprint of the beam on the sample surface relative to the initial beam direction, is given by

$$\gamma = \arctan(\tan\chi \cos\phi). \tag{4.15}$$

[7] Such forbidden Bragg reflections are referred to as 'systematic absences'. Refer to standard textbooks on solid-state physics or crystallography for a detailed explanation of how they arise.

[8] The width of a diffraction peak of a perfect crystal is given by the so-called 'Darwin width', which accounts for dynamical (multiple scattering) effects. For a given incident Bragg angle (or constant $m\lambda$), the Darwin width w_D is given by

$$w_D = \frac{K}{m^2}|F|, \tag{4.13}$$

where K is a constant. The first term has therefore a strong narrowing effect for the mth harmonic compared to the fundamental, while the second term describing the structure factor of the crystal $|F|$ (see Chapter 5), drops off on average with $\sin\theta/\lambda$ (see Figure 2.8) in a less pronounced manner and hence has a more modest broadening effect.

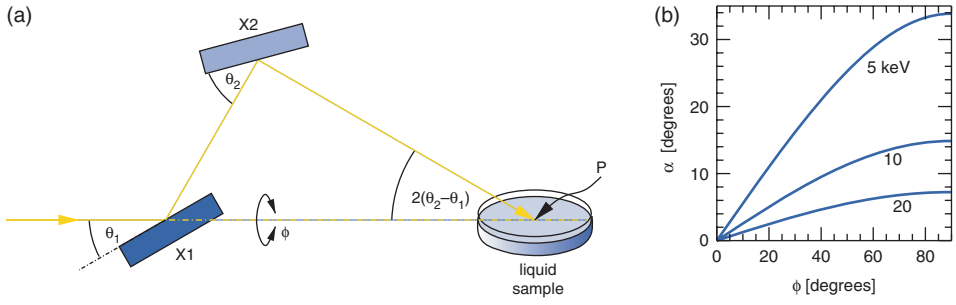

Figure 4.20 *Double-crystal deflectors. (a) the Bragg angles θ_1 and θ_2 of two dissimilar crystals X1 and X2 are adjusted to select a certain photon energy. Because $\theta_2 > \theta_1$, the doubly diffracted beam crosses over the original beam axis at P, where a liquid sample is placed. The incident angle α of the diffracted beam on the liquid surface can be adjusted by rotating the entire DCD around the original axis between $\phi = 0$ and 90°. The sample remains entirely stationary. (b) The change in incident angle α for a DCD using Si(111) and Si(220) crystals as a function of ϕ for three photon energies.*

4.3.4 Focusing Geometry

The position of the collimating and focusing mirrors and the DCM with respect to the x-ray source and sample decides the magnification, the divergence, and the degree of spherical aberration.

Wiggler- and bending-magnet radiation normally has a relatively large horizontal divergence and therefore a large cross-section at the primary optics. In order to minimize the radius of curvature of any optical components that need to be bent in order to collimate or focus, they should be as far as possible be positioned centrally between the source and sample (see Figure 4.21). This has the added advantage that this so-called

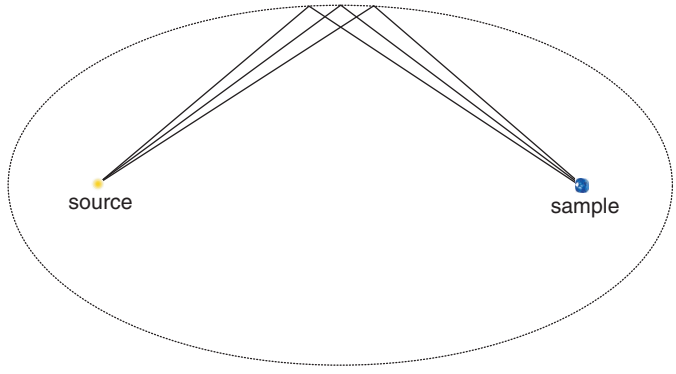

Figure 4.21 *By placing the beamline optics at a distance approximately in between the source and sample, any cylindrically or toroidally bent optical element most closely approximates an ideal elliptical surface, and also has the longest radius of curvature, thereby minimizing the amount by which it has to be flexed.*

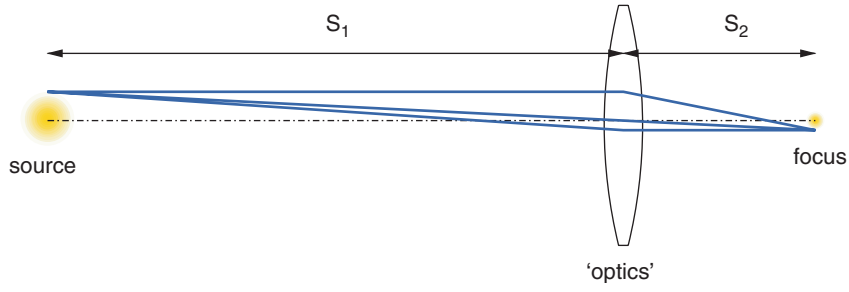

Figure 4.22 *The magnification factor of the x-ray source size is equal to S_2/S_1, which can also be expressed as a* demagnification factor of S_1/S_2.

'one-to-one' configuration is optimal for approximating ideal elliptical focusing surfaces by cylindrical ones, thereby minimizing spherical aberrations. Assuming near-perfect optics, the size of the focus should then be equal to that of the source.

Focusing undulator radiation is in general more straightforward, due to its intrinsic low divergence and small cross-section. Here, one can more easily place the optics away from the central position between source and sample without compromising the focus significantly. The main reason for moving away from the one-to-one configuration is to demagnify the source, and thereby obtain a tighter focus at the sample. This is important if the sample size is smaller than the 'natural' one-to-one image of the source.

Demagnification is simply achieved by placing the primary beamline optics closer to the sample than to the source. In general the magnification factor is equal to

$$M = \frac{S_2}{S_1} = \frac{1}{D}, \tag{4.16}$$

where S_1 is the source–optics distance, S_2 is the optics–focus distance, and D is the *de*magnification factor (Figure 4.22).

4.4 Microfocus and Nanofocus Optics

Demagnification by a judicious choice of the optics layout at the beamline can only provide focus spots down to the order of 1 μm in the vertical direction. Further demagnification to submicron dimensions normally requires secondary optics. The general approach to achieve this consists of focusing down the beam upstream of the sample at a pinhole which has a diameter significantly smaller than that of the primary focal spot, and using this as a virtual source for secondary optics (see Figure 4.23). For the tightest focusing it is important that one is able to decouple the position of secondary source from vibrations of the primary focus, caused by vibrations in the primary optics. This is achieved by mounting the pinhole, secondary optics and sample (at the secondary focus) on a quasi-vibrationless optical table.

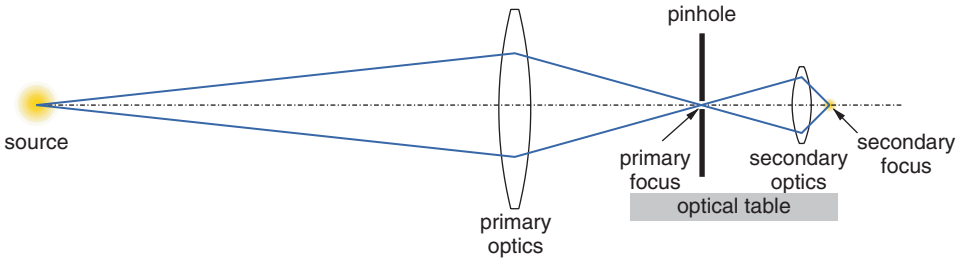

Figure 4.23 *Micro- and nanofocus beams are formed by selecting a fraction of the primary focus using a pinhole, and demagnifying this 'virtual source' using secondary optics.*

4.4.1 Lens Types

In this subsection, three of the most common lens types (excepting bendable mirrors and Kirkpatrick–Baez systems, described above) are described. They each use different optical phenomena, namely reflection, refraction and diffraction, to focus x-rays.

4.4.1.1 Compound Refractive Lenses

The fact that the refractive index for x-rays in solid material is generally lower than in vacuum or air can be exploited in the manufacture of x-ray optical elements, either by reflection or refraction. So-called compound refractive lenses (CRLs) are manufactured by a sequential array of holes drilled into high-transmission material (Be is a popular choice). Cylindrical focusing is achieved if all the hole axes are coplanar, while quasi-spherical focusing requires that alternate holes are drilled at 90° to one another (see Figure 4.24).

The focal length of CRLs is given by

$$f = R/2n_h\delta, \tag{4.17}$$

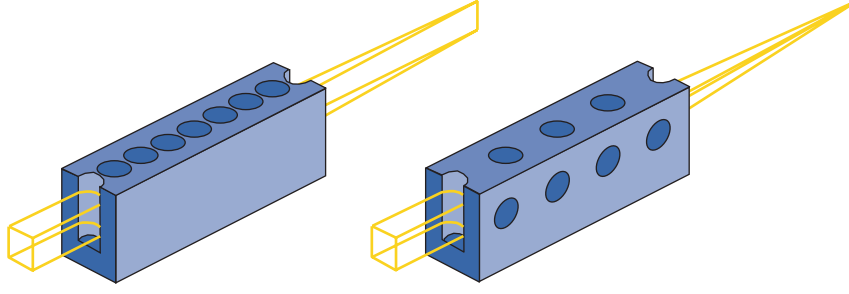

Figure 4.24 *X-ray focusing lenses can be constructed from high-transmission material drilled with arrays of holes, which act in much the same way as (very weak) cylindrical convex lenses do for optical light. These x-ray optical elements are called compound refractive lenses (CRLs).*

where R is the hole radius, n_h is the number of holes in the array and δ is the refractive index decrement (equal to 2.27×10^{-6} for Be at 12 keV). A lens consisting of an array of thirty 200 μm-diameter holes therefore has a focal length of approximately 73 cm.

The weak refraction of hard x-rays in condensed matter requires the radius of curvature R of the individual holes to be small. The spherical approximation to an ideal parabolic lens holds only for lenses with acceptance apertures for the x-rays $2R_0$ that are much smaller than R. However, in most CRL designs, $2R_0$ is comparable to or larger than R, and the spherical approximation breaks down. For the highest-quality imaging applications, these spherical aberrations are significant, limiting the performance of CRLs that use cylindrical geometries. Arrays of parabolic lenses have recently been fabricated using two techniques, electron-beam microlithography (EBM) and reactive-ion etching (RIE) at the Technical University in Dresden (see Figure 4.25). Using these lenses, formed in a silicon wafer, it was possible to generate a focal spot of approximately 50 nm for 21 keV photons [10].

CRLs have the advantage that they are easy to align in the x-ray beam, they do not divert the beam axis and elements can be stacked after one another to change the focal length. Their main disadvantages are (a) that the focal length is proportional to the square of the x-ray photon energy (see Equations (2.15) and (4.17)), which means that they must be repositioned, restacked or replaced if the photon energy is changed; and (b) that the transmission is in general fairly poor. A 100-hole CRL made of Be, with walls of 100 μm thickness separating adjacent holes has a transmission of 0.5 at 12 keV.

A further refinement of CRLs is illustrated in Figure 4.26. Here, one of the disadvantages of CRLs is effectively removed by decreasing the attenuation of the refracting medium via removal of sections of height Δ, chosen such that the phase difference ϕ between adjacent steps is 2π [11]. This is satisfied when

$$\Delta = m\lambda/\delta, \tag{4.18}$$

where m is an integer and δ is the refractive index decrement. For 1 Å radiation, $\delta \approx 3 \times 10^{-6}$ and hence Δ must be an integral multiple of 30 μm. Because δ is itself proportional

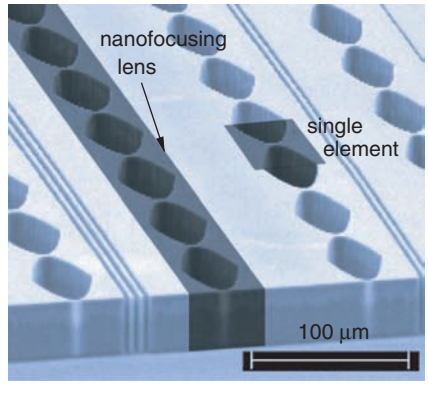

Figure 4.25 *A scanning electron microscope image of a nanofocusing lens array fabricated by lithographic and ion-etching techniques. Reprinted from [10] with permission of American Institute of Physics.*

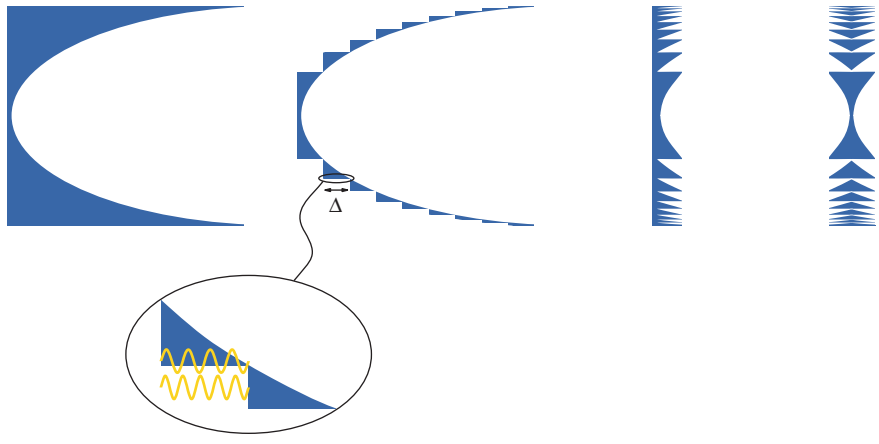

Figure 4.26 *Kinoform lenses. A parabolic CRL can be modified by removing sections of depth Δ, thereby reducing the the volume of attenuating material that does not actually refract the x-ray beam. The CRL can then be made to be more compact by collapsing all the elements to the same plane. The focusing power can be doubled by mirror imaging the CRL back-to-back. The inset shows that by carefully choosing the cutaway depth Δ, the phase difference between rays passing through the refractive material and those passing through the adjacent vacuum is a multiple of* 2π.

to the square of λ (Equation (2.15)), the disadvantage of these so-called kinoform lenses is that Δ is inversely proportional to λ and so they function only for discrete wavelengths.

4.4.1.2 Tapered Glass Capillaries

The capillary lens, first reported by scientists at Cornell University in 1994 [12], was the first lens type to produce submicron spot sizes and indeed a spatial resolution of 50 nm was achieved. Capillary lenses operate by 'funnelling' the light through a shallow cone by means of reflection.

The simplest taper is conical in shape. Consider the conical tube with a half-opening angle α shown in Figure 4.27(a). Parallel rays entering the capillary will be deflected by an angle 2α. This means that for a linearly tapered cone, the nth reflection impinges on the surface at an angle $(2n-1)\alpha$, which will continue until this value exceeds α_c, the critical angle for total external reflection, after which, the reflectivity drops precipitously. Beams entering the capillary further from the capillary axis will undergo more reflections and, depending on the geometry, may be lost. Hence, the effective acceptance aperture of the capillary depends on α, α_c and its length, and can be increased by coating the inner walls of the cone with a high-Z material with higher critical angle (Equation (2.15)).

More recently, tapered capillaries have been manufactured with parabolic or ellipsoid profiles in the plane containing the central axis (Figure (4.27(b))), which offer superior focusing properties and increased light-gathering power, on account of the increasing shallowness of the walls as the x-rays travel downstream.

An advantage of capillary lenses is that, because they operate on the principle of reflection, the focal length is independent of the photon energy, although the light-gathering

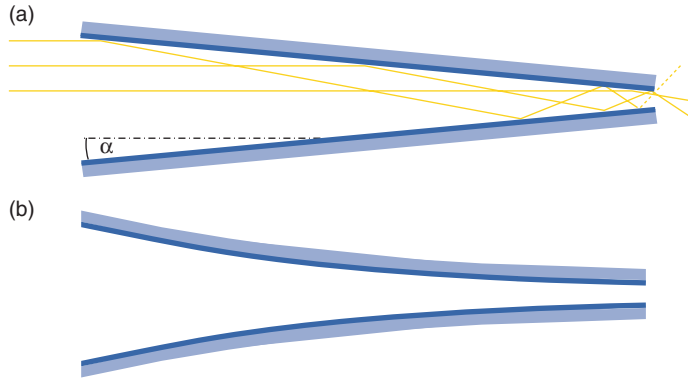

Figure 4.27 *Schematics of (a) a conically tapered glass capillary lens and (b) an ellipsoid capillary lens. The inner surface of the capillary can be coated with a high-Z material to increase the critical angle for total external reflection. Note in (a) that the outermost traced ray impinges on the conical surface at the fourth reflection at an angle that exceeds the critical angle for the inner coating, and is therefore lost.*

power decreases with increasing energy, because of the concomitant lowering of the critical angle for total external reflection (Equations (2.15) and (2.18)). A drawback of these optical elements is that the sample has to be placed very near the exit tip, generally within 20 to 100 times its inside diameter.

4.4.1.3 Fresnel Zone Plates

Fresnel zone plates (FZPs) focus x-rays using diffraction. A FZP consists of a set of radially symmetric rings, called Fresnel zones, which alternate between being opaque and transparent (see Figure 4.28). Constructive interference and thereby focusing occurs when the optical path difference between x-rays scattered through adjacent transparent

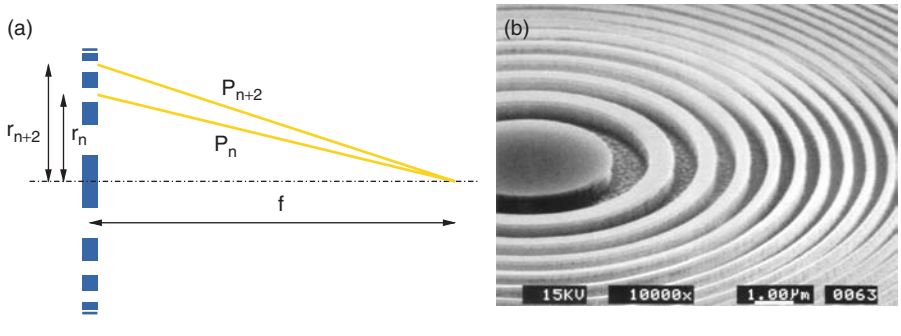

Figure 4.28 *Fresnel zone plates. (a) The path difference between adjacent transparent rings, $P_{n+2} - P_n$, in a zone plate should be equal to the wavelength of the x-rays being focused. (b) An electron microscopy image of a zone plate manufactured using electron-beam microlithography. Courtesy Christian David, Paul Scherrer Institut.*

rings is equal to one wavelength. Using Pythagoras' theorem,

$$\sqrt{f^2 + r_{n+2}^2} - \sqrt{f^2 + r_n^2} = \lambda. \tag{4.19}$$

Note here that n labels the ring and should not be confused with the refractive index. If we now assume that the radius of the largest Fresnel zone, r_N is much smaller than the focal length f, Equation (4.19) is accurately approximated by

$$\frac{r_{n+2}^2 - r_n^2}{2f} = \lambda. \tag{4.20}$$

This condition is met if

$$r_n = \sqrt{n\,\lambda f}, \tag{4.21}$$

from which it can be deduced that the areas of all the rings are constant and equal to

$$\pi\left(r_{n+1}^2 - r_n^2\right) = \pi f\,\lambda. \tag{4.22}$$

This is an important condition, needed in order to obtain complete constructive interference at the focus, as the integrated amplitude of the diffracted x-rays originating from each zone must remain the same for an evenly illuminated zone plate.

For plates containing many zones (i.e. large N), the focal length can be calculated from Equation (4.22) by remembering that $r_{n+1}^2 - r_n^2 = (r_{n+1} + r_n)(r_{n+1} - r_n)$. For an outermost zone of radius r_N with width Δr_N, the focus is given by

$$f = \frac{2r_N\,\Delta r_n}{\lambda}. \tag{4.23}$$

From Equations (4.21) and (4.23), it follows that

$$\Delta\, r_N = \sqrt{\frac{\lambda f}{N}}. \tag{4.24}$$

Hence, for a given x-ray wavelength and desired focal length, the product of the outermost zone radius and its width is a constant. A consequence of this is that the fabrication of FZPs for hard x-rays has in the past been impractical – for a focal length of, say, 5 cm, a wavelength of 1.5 Å, and $N = 100$, the outermost zone width would be less than 300 nm. One might therefore argue that the technological challenge to fabricate such narrow rings could be relaxed by choosing a smaller value for N. However, the ultimate resolution Δl of a zone plate depends on the zone widths according to

$$\frac{\Delta l}{\Delta\, r_N} = 1.22, \tag{4.25}$$

thus the smallest sized object that can be imaged using a zone plate is directly proportional to (and very close to, in absolute value) $\Delta\, r_N$.

Advances over the last decade in the fabrication process called electron-beam microlithography, however, have allowed FZPs to be produced with outer rings well

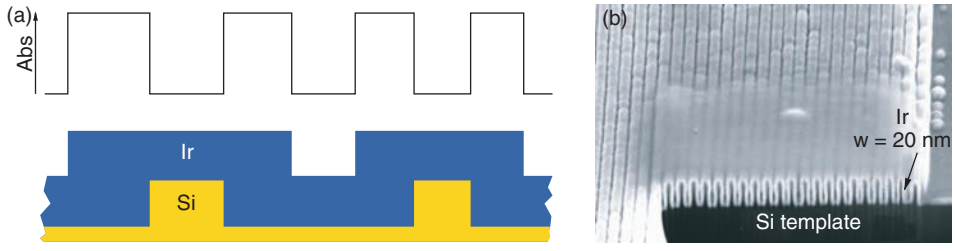

Figure 4.29 *Doubling the ultimate resolution of FZPs. (a) FZPs with ultra narrow zone ring widths can be manufactured by depositing high-Z material (e.g iridium) very homogeneously on top of a low Z-structured substrate (e.g. Si). (b) The rings of another zone plate were imaged using this novel design in a scanning transmission x-ray microscopy experiment. Features as small as 15 nm could be resolved. Reprinted from [13] with permission of the American Physical Society.*

under 100 nm, so that hard x-rays in the Angstrom range can now be focused down to just a few tens of nm.

A recent further improvement, essentially resulting in a doubling of the possible resolution, has been achieved using a simple post-processing of FZPs [13]. Consider Figure 4.29. A zone plate is etched out of silicon by electron-beam lithography such that the opaque (thicker) regions are narrower than they should be, according to Equation (4.21), and their height is small, making them only marginally opaque. This FZP is then homogeneously coated, atomic layer for atomic layer, with a very dense (and therefore x-ray opaque) material, such as iridium to a depth of a few tens of nm. The resulting transmission profile exhibits an effective increase in the number of rings by a factor of two, thereby improving the resolution accordingly. Using such a device, it was possible to resolve features of less than 10 nm.

Like CRLs, FZPs have the advantage that they are easy to align, while their focal length is linearly proportional to the photon energy (see Equation (4.23)). In general, they can potentially provide the tightest focus, although their transmission is typically only a few per cent.

4.5 Beam Intensity Monitors

It is important to monitor the incoming x-ray beam, in order to correct the recorded data for any fluctuations in beam intensity. As we saw in Section 3.2, for facilities not using top-up operations, the intensity of the beam incident on the beamline optics can drop by about one third over a period of some hours. This is liable to result in an even larger drop in the flux on the sample, because the thermal load on the beamline optics will change, resulting in thermal drift and detuning of the optical elements. The importance of having a reliable beam intensity monitor is therefore obvious.

The most common form of beam intensity monitor is the ionization chamber (Figure 4.30(a)). This operates by filling a vessel containing two electrodes with a gas (typically N_2 or Ar). The vessel is sealed at opposite ends with a transparent material

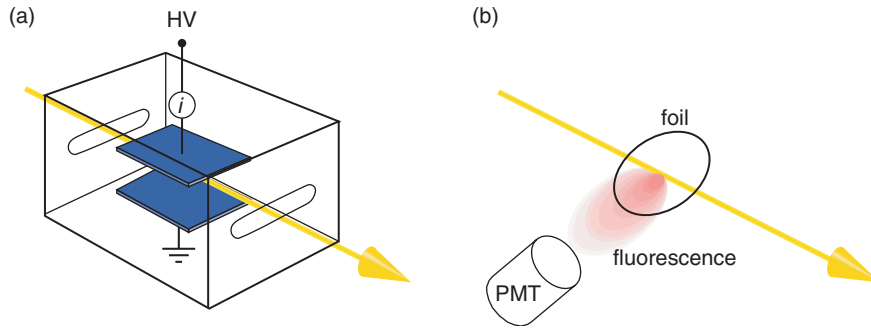

Figure 4.30 *(a) The intensity of the incoming beam in an ionization chamber is proportional to the current induced by photoionization of the gas in the chamber. (b) Fluorescence (or elastically scattered light) from a nearly transparent foil detected using a photomultiplier tube (PMT) can also be used to monitor the direct beam intensity.*

such as kapton foil, through which the direct x-ray beam can pass. One electrode is connected to a high-voltage source, of the order of a few hundred volts. An x-ray photon entering the chamber will be absorbed by the gas with a certain (low) probability. For example, about 0.5% of 16 keV photons are absorbed as they travel through 5 cm of N_2 at one atmosphere pressure. The gas molecule is thereby ionized and the photoelectron and ion accelerate to the anode and cathode, respectively. The electron, accelerated in the electric field, gains sufficient energy such that it frees one or more electrons upon collision with another atom or molecule of the medium. This avalanche process is measured as a current that is proportional to the x-ray flux.

There are several other methods of monitoring the direct beam intensity, for example, from the fluorescence or elastically scattered intensity from a nearly transparent foil (Figure 4.30(b)); from visible photoluminescence from a synthetic foil of diamond; or from the photocurrent produced in metallic meshes or ultrathin foils. The most important aspects of all these devices are that their response is linear with photon flux and that only an insignificant part of the beam is removed.

4.6 Detectors

In this section, we discuss the types of detectors used to detect x-ray photons and photoelectrons.

4.6.1 Photographic Plates

Although x-rays were first detected using photographic plates, which are still used in modern medical radiography, their use in the physical sciences has long waned, due to many drawbacks, including poor spatial resolution, nonlinearity, the difficulty in obtaining quantitative relative intensities, the relatively poor dynamic range and the very long readout times.

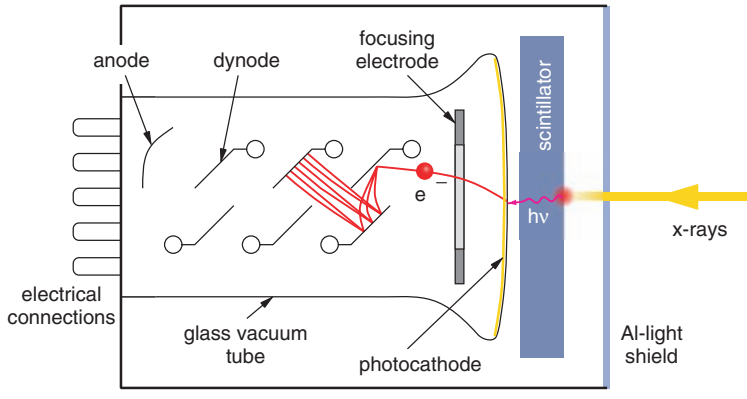

Figure 4.31 *Photons produced by x-rays absorbed in the scintillator material strike photocathode material deposited on the inside of the entry window of the tube. Electrons are produced by the photoelectric effect and are directed by a focusing electrode towards an array of electrodes, called 'dynodes'. Each successive dynode is held at a more positive voltage than the previous one. The electrons are therefore repeatedly accelerated towards each dynode and arrive with high energy. On striking a dynode, each electron produces several secondary electrons, which, in turn, are accelerated toward the next dynode, resulting in a cascade production of electrons. Finally, the anode is reached, where the accumulation of charge results in a sharp current pulse corresponding to the arrival of the original x-ray photon at the scintillator.*

4.6.2 Scintillator Detectors

Scintillation counters operate by the partial conversion of absorbed x-rays into visible or near-visible light, which can then be amplified using a photomultiplier tube (PMT). A schematic diagram of a typical detector is shown in Figure 4.31.

Typical inorganic scintillator materials are salts or metal oxides doped with high-Z materials. The host material is electronically excited by the absorption of an x-ray photon, and this excited state rapidly transfers its energy to nearby states of the dopant ion. These relax efficiently by nonradiative processes (i.e. those *not* involving the emission of a photon) to a much lower excited state only a few electron volts higher in energy than the final relaxed state, which is accessed by the emission of a photon in the visible or soft ultraviolet range (see Figure 4.32).

There are many types of scintillator material, both organic and inorganic. Here we mention just two inorganic scintillators, namely Tl-doped NaI (NaI(Tl)) and Ce-doped yttrium aluminum perovskite (YAP(Ce)). The most commonly used material for point scintillator detectors is NaI(Tl). It has a good x-ray stopping power, and has a scintillation emission maximum at 415 nm. Its two main disadvantages are (a) its softness and sensitivity to moisture (it is highly hygroscopic) and (b) its relatively long 'dead time' of approximately 250 ns,[9] which therefore limits the maximum reliable counting rate to about 1 MHz.

[9] The dead time of a system is the time after an event, during which the system is not able to record another event if it happens, and can be thought of also as a 'recovery time'.

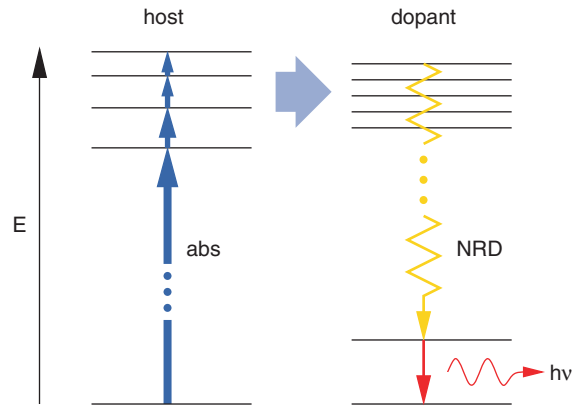

Figure 4.32 *A simplified energy-level scheme for scintillation in inorganic crystals. An x-ray photon is absorbed by the host crystal. The excited energy is transferred to energy states of the dopant material, which rapidly relax via nonradiative decay (NRD) to one or more low-lying levels, which in turn decay via luminescence and the ejection of a visible or near-visible photon, back to the ground state or low vibrational state.*

YAP(Ce) is a modern scintillator material with its peak scintillation peak centred at 350 nm. Although its efficiency is only about 40% of that of NaI(Tl), it is both mechanically robust and non-hygroscopic, and has a dead time of only 25 ns, making counting rates in the range of 10 MHz possible.

Scintillator plates can be used to record x-ray images. For example, a scintillator plate can be placed in front of a two-dimensional array of photosensitive elements, such as a CCD array (see below) in techniques such as x-ray tomography (see Section 7.2). The thicker the scintillator material, the more efficient is the x-ray stopping power of the plate and hence the stronger the signal. So why not make the scintillator layer as thick as possible? The reason why, in imaging applications, is that the spatial resolution is effected by the so-called 'point-spread function' of the device. We explain this briefly now.

4.6.3 The Point-Spread Function

The point spread function (PSF) describes the response of an imaging system to a point source or point object. Imagine, for example, an x-ray beam focused down to a Gaussian spot of 1 μm full-width at half-maximum, imaged by an x-ray area detector consisting of an array of pixels, each of size $100 \times 100\,\mu m^2$. The best possible imaging of the focus is, of course, a single illuminated pixel. In this limit, the PSF is a square the size of a pixel. In many cases, however, 'blooming' can cause the signal to leak across into the neighbouring area of the detector and increase the extent of the PSF still further, in which case the PSF is intensity dependent. The degree of blooming and its causes vary from system to system. For example, blooming can occur in scintillator plates by multiple scattering of x-rays across the depth of the active layer, and also because the converted visible photons are emitted in all directions and are also multiply scattered (see Figure 4.33). The thicker the active layer, the larger this effect. Blooming in charge-coupled devices is discussed below.

Figure 4.33 *A sharp signal (left) can be smeared out due to various processes (right). This less sharp image is the convolution between the original object and the point-spread function (middle).*

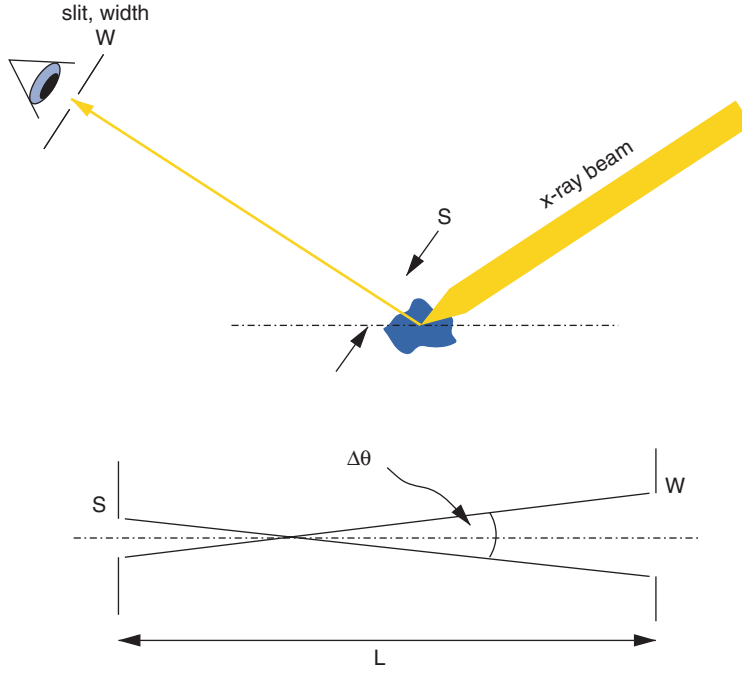

Figure 4.34 *The angular resolution of a point detector is defined by the angles subtended by the irradiated sample volume and by the slit width at the detector.*

4.6.4 Crystal Analysers

Scintillator detectors, in conjunction with slits, are known as point detectors. Consider such a system with the sensitive area defined by two pairs of slits (Figure 4.34).

The resolution in each orthogonal direction $\Delta\theta_{x,y}$ is limited by four factors. The first two originate from the incoming beam which will have a nonzero divergence (typically of the order of 0.1 to 1 mrad) and energy bandwidth (about 10^{-4} at 10 keV for x-rays monochromated by a crystal monochromator). The divergence and the irradiated volume of the sample can be minimized by using sets of slits before the sample at the cost

of signal intensity. Let us assume, for argument's sake, that the beam *is* completely monochromatic and parallel. The signal width is then determined by the size of the irradiated volume (controlled by slits positioned in the incoming beam in front of the sample) and the width of the slit aperture in front of the detector, also at the cost of signal strength. The angular resolution is given by

$$\Delta\theta = \frac{S + W}{L}, \tag{4.26}$$

where S is the irradiated sample size, W is the slit width and L is the sample–detector distance. So, for example, a typical powder diffraction setup, consisting of a fully irradiated capillary sample of diameter $S = 200\,\mu\text{m}$, a sample-to-detector slit distance of 1 m and a detector slit opening of 2 mm, results in an angular resolution of 2.5 mrad, or 0.14°.

An additional problem associated with simple scintillator–slit point detectors is the fact that, to a first approximation, the scintillator does not distinguish between x-ray photons of different energy. Both this and the problem of loss of signal intensity with slit width can be resolved using an adaptation called the crystal-analyser detector.

Consider Figure 4.35. A high-quality crystal such as Si(111) is mounted in the detector arm. The crystal planes parallel to the crystal surface have an interplanar distance d_C. The crystal is rotatable about an axis lying in its front surface. It is rotated to an angle θ_C subtended by the crystal surface and the line connecting the crystal rotation axis and the sample rotation axis such that, from Equation (4.10),

$$\begin{aligned} m\lambda &= 2d_C \sin\theta_C \\ \rightsquigarrow \sin\theta_C &= \frac{mhc}{2Ed_C}, \end{aligned} \tag{4.27}$$

where E is the photon energy. Do not forget that in diffraction techniques, we are dealing with elastic processes, and the photon energy is set by the incoming monochromatic beam – any fluorescent signal from the sample (which has lower photon energies) will therefore be filtered out by the crystal analyser. Only those elastically scattered rays

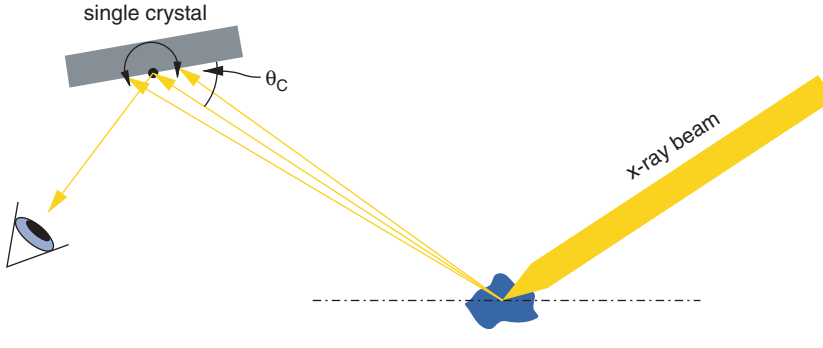

Figure 4.35 *The introduction of a single crystal in the detector system to select the diffracted signal can improve the resolution by precisely defining the angle between the incoming beam and the diffracted signal. In addition, it suppresses lower-energy fluorescence signal.*

impinging on the surface with exactly the angle θ_C to the surface plane will be diffracted and there is hence no need to use mechanical slits – the angles are accurately determined by the strict diffraction condition set by the crystal. Quantitatively, the 'effective' slit width, that is, the width of a slit that would have to be introduced into the detector system if the analyser crystal were absent (and ignoring the sample size), is given by

$$W = L\,\Delta\theta_C, \tag{4.28}$$

where L is the distance between the sample and analyser crystal and $\Delta\theta_C$ is the analyser-crystal rocking curve width, typically of the order of 10 μrad. Hence, the effective slit width of an analyser crystal detector is of the order of 5 μm. The signal intensity, however, is much larger than that achieved with a 5 μm mechanical slit because, in the absence of a slit aperture, the full x-ray spot size on the sample can be exploited.

Systems for data acquisition using a point detector (such as a scintillator, crystal analyser and/or slit system) are limited by their lack of 'peripheral vision' – spectral scans such as rocking curves probe one dimension only. Optimization of the signal by iteratively scanning different directions in the neighbourhood of a signal of an unknown intensity profile can easily lead to the recording of a local maximum, while the real signal of interest may be quite different. By using an area detector, especially one with low background noise, data-acquisition rates can be increased significantly.

4.6.5 Image Plates and Charge-Coupled Devices

Image plates became popular in the mid-1980s for x-ray crystallography. The heart of image plates is a storage phosphor screen. The most common materials are the barium fluorohalides. Wherever x-ray photons are absorbed on the phosphor screen, they produce secondary electrons which are trapped in so-called 'F-centres'.[10] After the end of the x-ray exposure, these metastable centres can then be read by a finely focused laser beam that scans across the image plate surface. The laser light causes the F-centres to release visible photons, which are detected by a PMT.

Nowadays, image plates are less popular as their specifications, in particular regarding efficiency, read-out times, point-spread function and dynamic range, lag behind those of charge-coupled devices (CCDs) [14]. The use of CCDs for synchrotron radiation was first described by Strauss *et al.* in 1988 [15]. The smallest pixel size at the time was approximately 20 μm, while the maximum imaging area was approximately $5 \times 5\,\mathrm{cm}^2$. Since then, the quality, resolution and efficiency of CCDs have improved with developments in semiconductor technology, so that nowadays, pixel sizes of well under 10 μm are readily available, with active areas in excess of $500\,\mathrm{cm}^2$.

A CCD records images by converting photons of light into electrons. The electrons are temporarily stored in individual picture elements (pixels) on a photosensitive semiconductor chip. At the end of an exposure, the accumulated charges are read off the chip and converted into an array of digital numbers, each number being proportional to

[10] The 'F' in F-centre comes from the German word for colour, 'Farbe'. An F-centre is a crystallographic defect in which an anionic (e.g. oxygen) vacancy in a crystal is filled by one or more electrons, depending on the charge of the missing ion. Electrons in such a vacancy tend to absorb light in the visible range of the electromagnetic spectrum, such that a material that is usually transparent becomes coloured.

the accumulated charge at the corresponding pixel. A digital image has been created. To record x-ray images, CCDs are used in conjunction with a overlayer of scintillator material typically a few tens of microns thick. The point-spread function is of the order of 10 to 50 μm.

The images stored in CCDs can be read out in less than one second, almost two orders of magnitude faster than that possible using image plates.

During long exposures needed to record faint objects, any bright features also in the field can exceed the maximum capacity (i.e. the electron-holding capacity) of the pixels on which they are being recorded. When this occurs, the excess charge spills over into adjacent pixels. This spillover causes blooming. Although a metallic so-called 'antiblooming grid' can be used to drain off excess charge from each pixel, this reduces the light-gathering capacity of the CCD and also makes the response nonlinear. Such nonlinearity is unacceptable for accurate scientific measurements, such as spectroscopy, crystallography and photometry, but can be satisfactorily used for qualitative imaging purposes.

4.6.6 Pixel and Microstrip Detectors

A revolutionary new x-ray area detector type has been developed in the first decade of the twenty-first century. The Pilatus detector [16] is a novel type of two-dimensional pixel array detector, which operates in single-photon counting mode, that is, counting registers for each pixel increment by one every time an x-ray photon above a set energy is detected (see Figure 4.36(a)).

The photon counting electronics comprise of a preamplifier, a comparator and a counter. The preamplifier amplifies the charge generated in the sensor by the incoming x-ray; the comparator produces a digital signal if the charge exceeds a predefined threshold. Thus, together with the counter, one obtains a complete digital storage and read-out of the number of detected x-rays per pixel without any readout noise (produced by the readout electronics) or dark current (signal occuring even in total darkness due to the random generation of electron-hole pairs in the photosensitive region of the device), along with being able to suppress lower-energy photons. This is very important when fluorescent signal should be eliminated in diffraction experiments, where only the elastically scattered photons should be recorded (see Figure 4.36(b)).

Pilatus detectors feature several advantages compared to current state-of-the-art CCD and imaging plate detectors (Figure 4.37). The main features include: no readout noise, superior signal-to-noise ratio, a read-out time of 5 ms (allowing a frame rate of up to 200 Hz, an order of magnitude faster than that of most CCDs), a dynamic range of 20 bits (i.e. the counter of any one pixel only overflows once 2^{20} events have been recorded), high detector quantum efficiency and the possibility to suppress fluorescence by an energy threshold that is set individually for each pixel. The short readout and fast framing times allow one to take diffraction data in a continuous mode without opening and closing a shutter for each frame. The next generation of these detectors, called Eiger, is, in 2011, being commissioned and has a 75 μm pixel size (see Figure 4.38) and, because the dead time is only 3 μs, offers a frame rate of over 20 kHz (using a reduced dynamic range). Future generations promise 50 μm technology and frame rates up to 100 kHz.

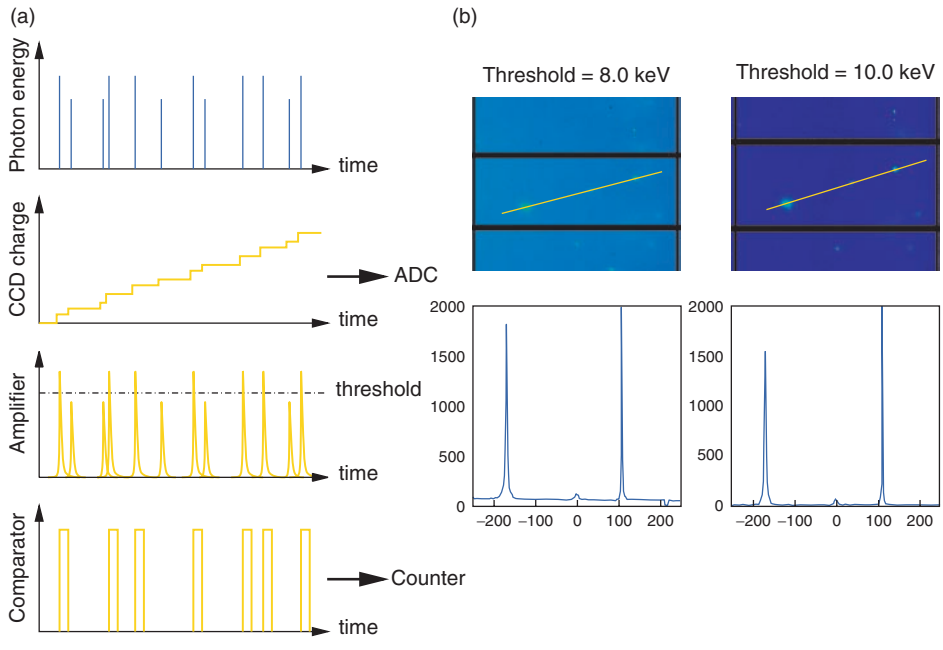

Figure 4.36 *Hybrid pixel detectors. (a) Consider x-rays of different energies impinging on a detector. In the case of integrating detectors, such as CCDs, charge is accumulated and then converted using an analog-to-digital converter (ADC). In the case of single-photon counting detectors, such as Pilatus, the charge generated by individual photons is amplified and converted into a voltage pulse with a height that is proportional to the photon energy. This is then compared to a voltage corresponding to a given energy threshold for each pixel, which sends only those pulses to the counter which are higher than the threshold. (b) By setting a threshold of 10.0 keV, fluorescence emission at 8.1 keV from Cu-atoms in an Al-Cu-Fe quasicrystal can be suppressed, resulting in a much lower background signal and the emergence of fine detail in the elastically scattered signal, as demonstrated by the two line plots (below) through the yellow paths in the images (above). Courtesy Beat Henrich, Paul Scherrer Institut.*

Because of their specifications, hybrid pixel detectors offer important advantages over state-of-the-art CCD and imaging plate detectors for various x-ray detection experiments. Major improvements can be expected for time-resolved experiments, for the study of weak diffraction phenomena (e.g. diffuse scattering) and in protein crystallography.

An analogous one-dimensional detector, or strip detector, has also been developed at the Paul Scherrer Institut [17]. The Mythen detector consists of a line array of modules, each module containing 1280 silicon strips of width 0.05×8 mm (see Figure 4.39). Like Pilatus, it also operates in a photon-counting mode. Its ultimate intrinsic resolution is of the order of 0.004°, given by the size of the individual strips and their distance from the sample. In reality, the effective resolution may be poorer, depending on the volume of the sample that is being illuminated with x-rays. A Mythen detector is being used at the powder diffraction station of the Materials Science beamline at the Swiss Light Source, which spans an angle of 120° (24 modules), and is revolutionizing time-resolved

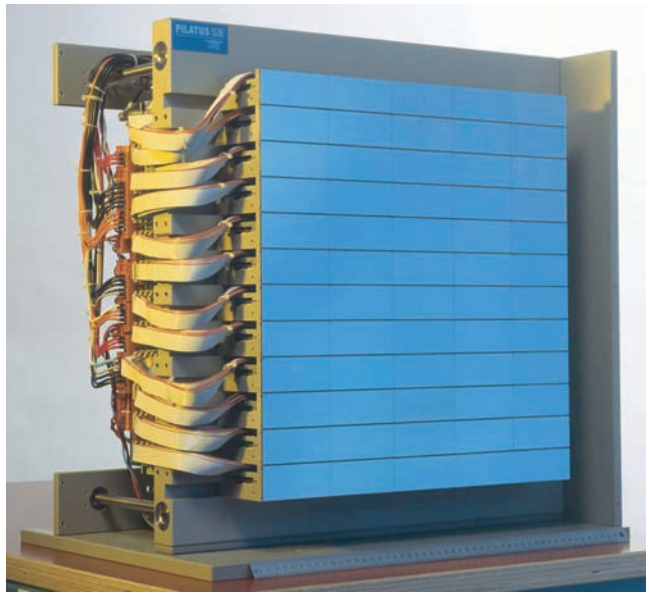

Figure 4.37 *The Pilatus 6M detector. This image, with the outer casing of the detector removed, shows the detector to be composed of 5 × 12 modules, corresponding to 2463 × 2527 = 6 224 001 pixels and a total active area of 424 × 435 mm^2. This detector is used at the X06SA Protein Crystallography beamline at the Swiss Light Source. Thanks to the fast readout times, a novel, time-efficient, method has been developed to record protein crystals, thereby helping to minimize radiation damage (see Section 5.11). Courtesy Christian Brönnimann, Dectris AG.*

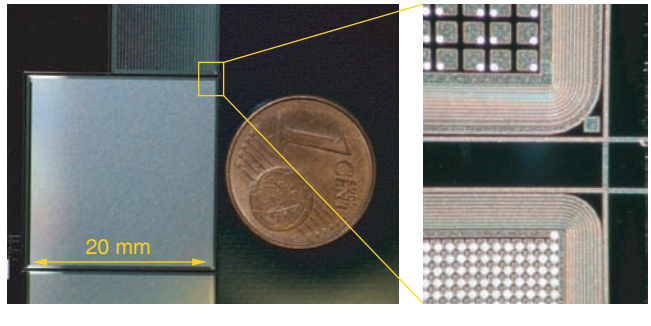

Figure 4.38 *The Eiger single-chip sensor (bottom, left image) is larger than that of the Pilatus (above the Eiger chip), and has pixels with linear dimensions that are only 44% as large. Courtesy Beat Henrich, Paul Scherrer Institut.*

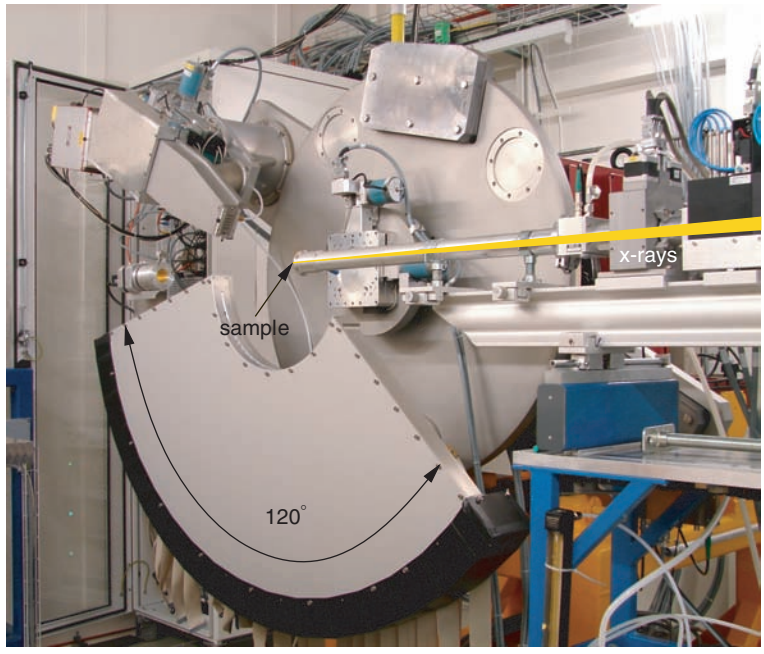

Figure 4.39 *The 120° Mythen detector at the Powder Diffraction station of the Swiss Light Source, containing 30 720 readout elements. Courtesy Fabia Gozzo, Paul Scherrer Institut.*

diffraction studies – the minimum readout time for the entire detector is of the order of a millisecond, which should be compared to several hours scan time when using more conventional point detectors, especially when the latter are used in conjunction with a crystal analyser (see Chapter 5). This sensational increase in data acquisition rates on the one hand opens new vistas of time-resolved powder diffraction studies, while on the other, it allows one to record entire data sets within a fraction of a second, thereby avoiding, or at least minimizing, the deleterious effects of radiation damage on the sample.

An example of the benefits of such detectors is illustrated in Figure 4.40. The pharmaceutical compound bupivacaine hydrochloride can exhibit several polymorphs. Despite their chemical similarity, the biological impact of one polymorph over another can be fundamentally different, hence the need to identify and characterize each polymorph cannot be overstated.

Crystals of pharmaceutical products are particularly prone to radiation damage by x-rays, due to their relatively low bond strengths. A crystalline powder sample of the so-called D-form of S-bupivacaine was therefore investigated using a multicrystal analyser detector in a 'fast' scan over 60 ° lasting 15 minutes (4 °/minute, step size of 0.004 °). A small section of the scan is shown in plot AC1 of Figure 4.40. A second scan of the same sample using identical conditions was recorded immediately afterwards. This scan

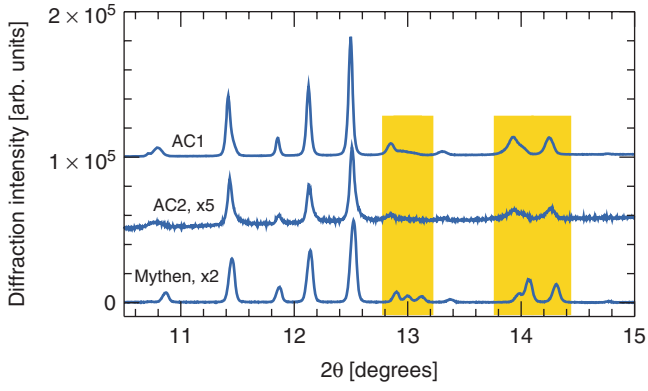

Figure 4.40 *Parallel data acquisition using a microstrip detector can be essential in some compounds in order to avoid radiation damage from the synchrotron beam. Two successive 15-minute analyser-crystal scans (AC1 and AC2) of a powder sample of a polymorph of the pharmaceutical compound bupivacaine, plus a 1 s-scan of a fresh sample using the Mythen microstrip detector. Note the highlighted regions around 13 and 14.1°. Courtesy Fabia Gozzo, Paul Scherrer Institut.*

was significantly weaker (note the 5× scaling of plot AC2 in Figure 4.40) and showed different peak positions and intensities compared to the first run, indicative of radiation damage. A fresh sample was then investigated using the Mythen detector (lowest plot, Figure 4.40) using a 1 s exposure. The difference in the signal shape compared to *both* plots AC1 and AC2 is clearly identifiable in the highlighted regions, demonstrating that the sample suffered significant radiation damage even after recording the first 13° of the first scan, that is, after approximately three minutes exposure! Therefore, obtaining reliable data using conventional detectors for subsequent structural refinement is all but impossible for such sensitive organic crystals, and the use of one-dimensional detectors for parallel acquisition is indispensable.

4.6.7 Energy-Dispersive Detectors

Spectroscopy experiments, in which the energy spectra of photons or electrons originating from the sample are investigated (such as x-ray fluorescence or photoelectron spectroscopy) require energy-dispersive detectors. We discuss photon detectors in the range of 100 eV to several tens of keV, and electron detectors from approximately one meV to a few keV.

4.6.7.1 Photon Detectors

Spectral analysis of an x-ray spectrum, such as in x-ray fluorescence, is achieved in one of two ways. In wavelength dispersive spectrometers (WDX or WDS), the photon energy is selected by diffraction on a single crystal or a grating in much the same way as in crystal or grating monochromators, described above.

In energy dispersive spectrometers (EDX or EDS), a solid-state detector allows the determination of the energy of the photon when it is detected. Such semiconductor detectors are usually fabricated from silicon (sometimes doped with lithium) or germanium. The device is polarized with a high voltage. When an x-ray photon hits it, electron-hole pairs are created, that drift in the applied high electric field. The electric charge is collected in a manner similar to the charging of a capacitor. Crucially, the voltage increment due to the collected charge is proportional to the photon energy, hence it is possible to determine the energy spectrum. The voltage is reset regularly to avoid saturation. Such detectors also capture larger solid angles of signal and therefore have higher sensitivity and are more suited to experiments where signal is weak, such as when trace elements are being investigated.

EDX spectrometers are smaller, cheaper, and measurements are faster. However, the resolution, typically between 100 and 200 eV, is far lower than those of WDX spectrometers (see Figure 4.41(b)).

A subclass of energy-dispersive photon detectors (so-called low-temperature detectors, LTDs) operate at cryogenic temperatures and offer an order of magnitude better resolution than conventional solid-state detectors, though with lower detector efficiency. This is summarized in Figure 4.41(a).

The speed of data acquisition using EDX spectrometers permits the user to rapidly obtain chemically resolved two-dimensional maps with a spatial resolution down to the micron range and sensitivities in the sub-mg kg^{-1} range (see Figure 4.42).

Which detector type one uses depends, needless to say, on the application. Which is more important – sensitivity or spectral resolution? In resonant inelastic x-ray scattering (RIXS, see Section 6.6.3), for example, line separations are typically of the order of a few eV or less, marginally below the present capabilities of modern LTDs. Future improvements in LTDs might therefore potentially open up new vistas of possible

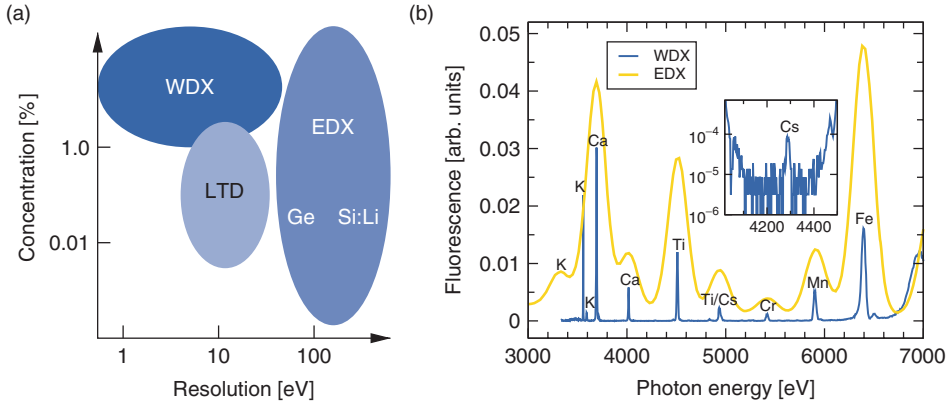

Figure 4.41 *Energy-dispersive photon detectors. (a) A broad classification of different detector types, according to their sensitivity and spectral resolution. (b) The fluorescence spectra of a clay mineral recorded with EDX and WDX, containing approximately 100 parts per million Cs. The Cs Lα line is just detectable using WDX (see inset), but is completely lost using EDX. Courtesy Daniel Grolimund, Paul Scherrer Institut.*

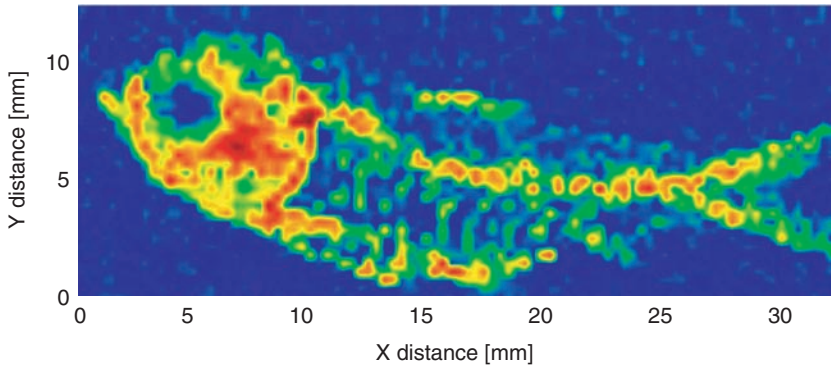

Figure 4.42 *An energy-dispersive x-ray fluorescence map taken at the Th Lα emission line (12 969 eV) of a fossil fish found in the Green River Formation of Fossil Lake, Wyoming. Warm regions depict high concentrations of thorium. The incident x-ray beam was tuned to 17.2 keV. The high levels of thorium are caused by its substitution for calcium in the phosphate mineral that composes the bone, fluorapatite. By studying such paleontological specimens, it is hoped to better understand the chemistry of lanthanides and actinides in the natural environment. Adapted from [18] courtesy of Troy Rasbury, Stony Brook University, with permission of Elsevier.*

experiments in techniques such as RIXS, where signal intensity is often at a premium and the deciding factor in determining the feasibility of a given experiment.

4.6.7.2 Electron-Energy Analysers

Electron-energy analysers are necessary for photoemission spectroscopy experiments. How does one measure the kinetic energies of electrons? There are two common detector types, the cylindrical mirror analyser (CMA) and the concentric hemispherical analyser (CHA). Even though the resolving power of CMAs can be improved using a double-pass configuration, most electron spectroscopy instruments nowadays use CHAs, and it is these that are now described in some more detail.

A schematic of a CHA is shown in Figure 4.43. Two metallic hemispheres of radii R_1 and R_2 ($> R_1$) are positioned concentrically. The inner and outer hemispheres are floated to the negative potentials $-V_1$ and $-V_2$, respectively.

An electron entering the CHA approximately parallel to the equipotential surfaces with a kinetic energy E will be focused at the exit to the detector if

$$e\Delta V = E\left(\frac{R_2}{R_1} - \frac{R_1}{R_2}\right), \tag{4.29}$$

where $\Delta V = V_2 - V_1$. Equation 4.29 can be reexpressed as $E = \kappa e \Delta V$, where κ is known as the spectrometer constant. Electron energy spectra can therefore be recorded by sweeping ΔV.

Note that the electrons can enter the spectrometer at any angle in the $x-z$-plane (see Figure 4.43) and still arrive exactly at the exit to the detector. This is not so for the perpendicular $y-z$-plane. Suppose that the electrons are injected at an angle $\delta\alpha$ to the

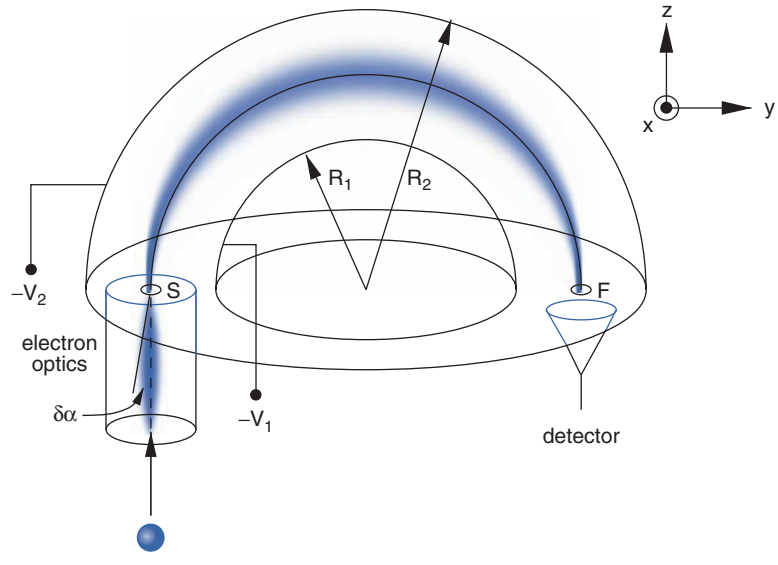

Figure 4.43 *Schematic of a concentric hemispherical analyser. Electrons enter the CHA through the aperture S and are focused at the exit to the detector F.*

entrance-aperture normal in the $y-z$-plane. For entrance and exit apertures S and F of the same widths W, the resolution of the spectrometer is given by

$$\frac{\Delta E}{E} = \frac{2W}{R_1 + R_2} + (\delta\alpha)^2. \tag{4.30}$$

Therefore the resolution improves for small aperture dimensions and small electron divergences, though at the expense of signal intensity.

Spectra of electron energies can be recorded with either a constant $\Delta E/E$ or a constant ΔE. Spectra taken with a constant $\Delta E/E$ therefore show signals with widths that are directly proportional to their electron energies. More commonly, it is desired that the instrumental peak widths remain constant in any given spectrum, and the constant ΔE mode is chosen. This is achieved by accelerating or decelerating the electrons within the electron optics immediately before entering the CHA to a fixed energy value, known as the pass energy, which typically lies at about 50 to 100 eV. Hence, the energy selection is not performed by varying ΔV in the CHA. This is kept constant according to Equation (4.29) for $E = E_{\text{pass}}$. Instead, the voltage used to change the electron energy is scanned. For example, if one chooses a pass energy of 70 eV, and is measuring electrons with kinetic energies of 1000 eV, a negative voltage of −930 V must be applied with respect to the electrons as they enter the electron optics unit.

The efficiency, resolution and sophistication of CHAs has recently increased by the replacement of the point detector at F with a multichannel plate (MCP) and CCD camera. The entrance aperture S is a slit which allows up to a 30° acceptance angle in the $x-z$-plane. This is projected and imaged along one axis of the MCP (three yellow paths

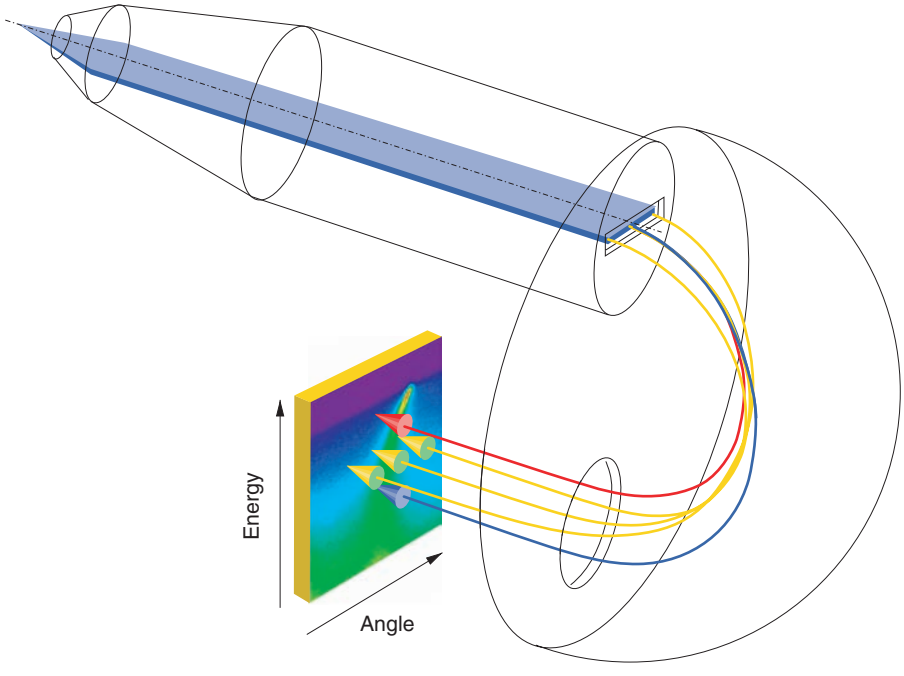

Figure 4.44 *Data spanning different electron kinetic energies and trajectories can be simultaneously recorded using a multichannel plate and CCD camera as the detector system in a CHA.*

in Figure 4.44), while electrons of different energy (the blue and red paths in Figure 4.44) impinge on the MCP at different positions in the perpendicular direction. At the time of writing, such detectors have an energy resolution as small as 1 meV.

4.7 Time-Resolved Experiments

In recent years, more and more studies at third-generation synchrotron facilities involve time-resolved experiments. This might be on the subsecond to millisecond scale, using the hybrid pixel and microstrip detectors described above, or go down to the range of less than 1 ns using fast electronic devices such as avalanche photodiode detector (APD) arrays. Increasingly experiments are designed to exploit the time structure of the storage-ring current, with a timescale measured in a few picoseconds. Indeed, the fourth-generation x-ray free-electron lasers will provide unsurpassed temporal resolution of the order of a few femtoseconds (see also Section 3.8.2).

In third-generation facilities, measurements requiring time resolutions below a few picoseconds require the use of streak cameras, which have been demonstrated to be able to probe timescales down to under 250 fs [19]. Here, we briefly describe the operating principles of APDs and streak cameras.

4.7.1 Avalanche Photodiodes

Avalanche photodiodes (APDs) can be thought of as the solid-state equivalent of photomultiplier tubes, the amplifier stage of the scintillator detectors shown in Figure 4.31, insofar that they exploit the same avalanche production of electrons through collisional ionization. They can operate in two modes, whereby the applied reverse bias voltage is either marginally below that required for breakdown of the semiconductor, or marginally above this, in which case there is a catastrophic drop in resistance and the gain in current is of the order of 10^5 so that single photons incident on the APD can be detected. In this latter case they are said to operate in the photon-counting mode and are called Geiger-mode avalanche photodiodes (GMPDs). Detector electronics ensure that the photodiode itself is not destroyed by this runaway process by limiting the available maximum current and its duration (so-called 'current quenching').

GMPDs can resolve events down to approximately 100 ps. An array of 8×8 GMPDs in conjunction with a 100 μm-thick, patterned scintillator crystal of lutetium oxyorthosilicate (LSO) is shown in Figure 4.45. Here, the temporal resolution is limited by the recovery time of the scintillator, which is of the order of 20 ns, much slower than the capabilities of the GMPD itself [20].

4.7.2 Streak Cameras

Streak cameras are devices which convert a short signal over time into one across space. The signal is recorded on an area detector, whereby one coordinate corresponds to time and the orthogonal axis to space. Consider Figure 4.46. A part of an incident transient x-ray signal is selected using a narrow slit aperture. This impinges on a negatively biased photocathode, resulting in the ejection of a packet of photoelectrons, which are then accelerated to a partially transmitting metal (gold) grounded grid. The electrons now have a well-defined energy, given by the potential difference between the photocathode and the anode and typically of the order of 10 kV. As they begin to drift between a pair of sweep electrodes, a ramp voltage is applied across these. The start of the ramp is

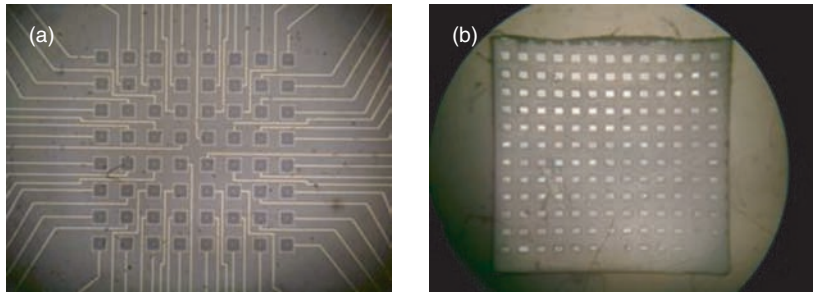

Figure 4.45 *(a) An array of 8×8 GMPD pixels. Each pixel is $100 \times 100\,\mu m^2$. (b) A matching array of LSO crystals is placed on top of the GMPD array in order to provide the visible light necessary for detection. Reprinted from [20] with permission of IUCr.*

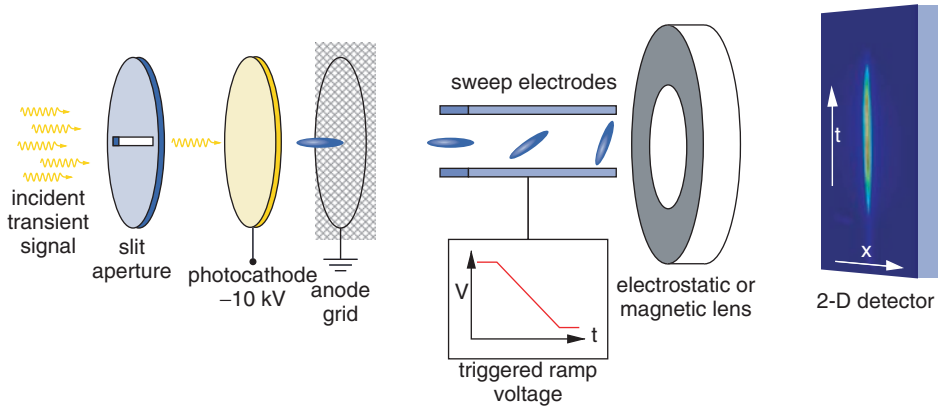

Figure 4.46 *The operating principle of a streak camera.*

triggered by the incident signal such that the leading edge of the electron package sees a different electrical field than the trailing edge and is therefore deflected by a different amount. This effectively skews the package perpendicular to both the slit and the streak camera axis. The skewed signal is then magnified using an electrostatic or magnetic lens and imaged on an area detector. The axis parallel to the slit aperture corresponds to the spatial coordinate, while the orthogonal axis represents time. Indeed, it may be possible to scan the source of the transient signal (in Figure 4.46, this would be in the vertical direction), so that different parts of it are selected by the slit aperture, thereby obtaining information about the second spatial coordinate.

Streak cameras can be used for several applications, from the temporal form of x-ray pulses to transient x-ray emission phenomena from samples. Recently, a time-resolution of less than 250 fs using a streak camera with a particularly large magnetic solenoid lens was demonstrated [19].

4.8 Concluding Remarks

Beamlines can appear to be very bewildering and intimidating to the uninitiated user, and at one level they are complex, insofar that there are usually many components to consider. Each component, however, can be understood by scientists from all disciplines. It is therefore hoped that the above breakdown of commonly found equipment in beamlines given in this chapter will help ameliorate the trepidation felt by most first-time users.

Needless to say, this chapter has not been exhaustive in the listing of beamline equipment, or described all the possible variations that can arise, both of individual components and also how they are combined, which are innumerable – there are seldom two beamlines that are clones of one another. Indeed it is this possibility of ‘mixing and matching’ to the user’s needs which makes synchrotron experiments so diverse and exciting.

References

[1] C. Schulze-Briese, B. Ketterer, C. Pradervand, C. Bronnimann, C. David, R. Horisberger, A. Puig-Molina, and H. Graafsma, "A CVD-diamond based beam profile monitor for undulator radiation," *Nucl. Instrum. Methods A*, 467, pp. 230–234, 2001.

[2] P. Kirkpatrick and A. V. Baez, "Formation of optical images by X-rays," *J. Opt. Soc. Am*., 38, pp. 766–774, 1948.

[3] R. L. Johnson, "Grazing-incidence monochromators for synchrotron radiation – a review," *Nucl. Instrum. Methods A*, 246, pp. 303–309, 1986.

[4] H. Petersen, C. Jung, C. Hellwig, W. B. Peatman, and W. Gudat, "Review of plane grating focusing for soft x-ray monochromators," *Rev. Sci. Instrum*., 66, pp. 1–14, 1995.

[5] P. Melpignano, S. DiFonzo, A. Bianco, and W. Jark, "Optimization of spherical grating monochromators operating with variable included angle for different applications," *Rev. Sci. Instrum*., 66, pp. 2125–2128, 1995.

[6] L. E. Berman, J. B. Hastings, D. P. Siddons, M. Koike, V. Stojanoff, and M. Hart, "Diamond crystal x-ray optics for high-power-density synchrotron radiation beams," *Nucl. Instrum. Methods A*, 329, pp. 555–563, 1993.

[7] Y. V. Shvyd'ko, S. Stoupin, A. Cunsolo, A. H. Said, and X. R. Huang, "High-reflectivity, high-resolution x-ray crystal optics with diamonds," *Nature Physics*, 6, pp. 196–199, 2010.

[8] J. W. M. DuMond, "Theory of the use of more than two successive x-ray crystal reflections to obtain increased resolving power," *Phys. Rev*., 52, pp. 872–883, 1937.

[9] V. Honkimaki, H. Reichert, J. S. Okasinski, and H. Dosch, "X-ray optics for liquid surface/interface spectrometers," *J. Synch. Rad.* 13, pp. 426–431, 2006.

[10] C. G. Schroer, O. Kurapova, J. Patommel, P. Boye, J. Feldkamp, B. Lengeler, M. Burghammer, C. Riekel, L. Vincze, A. van der Hart, and M. Kuchler, "Hard x-ray nanoprobe based on refractive x-ray lenses," *Appl. Phys. Lett*., 87, p. 124103, 2005.

[11] K. Evans-Lutterodt, J. M. Ablett, A. Stein, C.-C. Kao, D. M. Tennant, F. Klemens, A. Taylor, C. Jacobsen, P. L. Gammel, H. Huggins, S. Ustin, G. Bogart, and L. Ocola, "Single-element elliptical hard x-ray micro-optics," *Optics Express*, 11, pp. 919–926, 2003.

[12] D. H. Bilderback, S. A. Hoffman, and D. J. Thiel, "Nanometer spatial-resolution achieved in hard x-ray imaging and Laue diffraction experiments," *Science*, 263, pp. 201–203, 1994.

[13] K. Jefimovs, J. Vila-Comamala, T. Pilvi, J. Raabe, M. Ritala, and C. David, "Zone-doubling technique to produce ultrahigh-resolution x-ray optics," *Phys. Rev. Lett*., 99, p. 264801, 2007.

[14] G. E. Smith, "Nobel Lecture: The invention and early history of the CCD," *Rev. Mod. Phys*., 82, pp. 2307–2312, 2010.

[15] M. G. Strauss, I. Naday, I. S. Sherman, M. R. Kraimer, E. M. Westbrook, and N. J. Zaluzec, "CCD sensors in synchrotron x-ray detectors," *Nucl. Instrum. Methods A*, 266, pp. 563–577, 1988.

[16] P. Kraft, A. Bergamaschi, C. Brönnimann, R. Dinapoli, E. F. Eikenberry, B. Henrich, I. Johnson, A. Mozzanica, C. M. Schlepütz, P. R. Willmott, and B. Schmitt, "Performance of single-photon-counting PILATUS detector modules," *J. Synchrotron Rad*., 16, pp. 368–375, 2009.

[17] B. Schmitt, C. Brönnimann, E. F. Eikenberry, F. Gozzo, C. Hormann, R. Horisberger, and B. Patterson, "Mythen detector system," *Nucl. Instrum. Methods A*, 501, pp. 267–272, 2003.

[18] J. M. Cole, J. Nienstedt, G. Spataro, E. T. Rasbury, A. Lanzirotti, A. J. Celestian, M. Nilsson, and G. N. Hanson, "Phosphor imaging as a tool for in situ mapping of ppm levels of uranium and thorium in rocks and minerals," *Chem. Geology*, 193, pp. 127–136, 2003.

[19] J. Feng, H. J. Shin, J. R. Nasiatka, W. Wan, A. T. Young, G. Huang, A. Comin, J. Byrd, and H. A. Padmore, "An x-ray streak camera with high spatio-temporal resolution," *Appl. Phys. Lett*., 91, p. 134102, 2007.

[20] I. Johnson, Z. Sadygov, O. Bunk, A. Menzel, F. Pfeiffer, and D. Renker, "A Geiger-mode avalanche photodiode array for X-ray photon correlation spectroscopy," *J. Synchrotron Rad*., 16, pp. 105–109, 2009.

5

Scattering Techniques

5.1 Introduction

Diffraction is one of the most established and venerable synchrotron methods – virtually no modern synchrotron facility is without one or more protein crystallography beamlines, for example. Insights from diffraction data taken at synchrotrons continue to surprise and fascinate the scientific community, as illustrated by the recent spate of Nobel prizes in Chemistry in 2003, 2006, and 2009 for research into fundamental biological processes using macromolecular crystallography.

In the multidisciplinary investigations that are now so common in synchrotron research, many users who would gain invaluable information from diffraction studies are not hardcore crystallographers, and indeed may have little or no background in solid-state physics or materials science. Such scientists do not necessarily feel the need or motivation to plough through any of the many excellent crystallography monographs that are available [1–3] in order to gain the perceived necessary level of understanding to conduct a successful diffraction experiment, and can thus sometimes be deterred from pursuing this line of research. In this chapter, I will attempt to convey the most important aspects of diffraction, provide useful insights that assist in obtaining a physical understanding from diffraction data, and explain experimental and technical aspects of synchrotron-based diffraction methods in a manner such that the reader does not require a deep knowledge of crystallography.

X-ray diffraction (XRD) is a technique that enables one to determine the atomic structure of crystals with unsurpassed precision. The word ‘diffraction’ has its etymological roots in the past participle of the Latin verb ‘diffringere’, which means ‘to break up into pieces’. The ‘pieces’ that are observed in diffraction are those parts of an incoming beam that constructively interfere after being scattered by the individual particles (in x-ray diffraction, atoms, or more precisely, electron clouds around the atoms) within the scattering body. XRD has many important applications in science and technology, for example, in determining the structure and functionality of proteins; in identification of changes in

An Introduction to Synchrotron Radiation: Techniques and Applications, First Edition. Philip Willmott.

crystal structure due to external influences or crystallographic defects, which can induce property changes and are of great importance in electronics and sensor technology; in tracking the dynamics of chemical transitions; and in many other technological fields.

When x-rays impinge on a crystal, they are scattered by the electrons within that crystal. The electron density has the same periodicity as that of the crystal lattice (see below), and hence it acts like a diffraction grating in just the same manner as does an optical grating for visible light. In certain directions, the scattered x-rays add up constructively to produce diffraction peaks. The principal aim in XRD is to record this diffraction pattern and then convert it back into the periodic electron density distribution that produced it.

This goal is hampered by the fact that x-ray detectors do not measure the electric field amplitudes and phases of the x-radiation,[1] but rather the *intensity*, that is, the absolute square of the amplitude, in which the phase information is lost. For relatively simple crystalline systems, containing less than a few hundreds of atoms per unit cell, the brute calculating power brought to bear by modern computers normally allows one overcome this so-called 'phase problem' and to reconstruct the atomic structure, using reasonable physical and chemical constraints. This becomes quite impractical, however, when attempting to determine more complicated structures, especially of zeolites and proteins, which may contain several tens of thousand atoms per unit cell. In this case, additional experimental information is required. We discuss the phase problem in a semiquantitative manner in Section 5.6.

In this chapter, we will study the diffraction of x-rays by regular (three-dimensional and two-dimensional) periodic structures, and also scattering from noncrystalline objects. We will only concern ourselves with elastic (Thomson) scattering of x-rays by electrons.

As is usual in all diffraction phenomena, the wavelength of the x-rays used for diffraction studies is of the same order of magnitude as the distance between the scattering centres used to produce the pattern, and therefore lies in the broad region between approximately 0.1 and 2 Å, corresponding to typical photon energies between 100 keV and 5 keV. The exception to this is resonant soft x-ray scattering, which uses photon energies somewhat below 1 keV.

5.2 Diffraction at Synchrotron Sources

The tunability and high brilliance of x-ray radiation at diffraction facilities at third-generation synchrotron sources offer the user greater possibilities compared to laboratory-based sources. What problems might motivate a potential user to consider using synchrotron radiation, and what factors determine the choice of the wavelength to be used? The answer for any one experiment may be manifold and include subtle aspects too specialized to be generalized. However, one can usually say that one or more of the following play a role:

[1] This should come as no surprise, as typical frequencies of electromagnetic radiation in the x-ray region lie around $10^{18}-10^{19}$ Hz. The fastest electronic equipment has a bandwidth of approximately 10^{12} Hz, corresponding to the far-infrared region.

- Many experiments require samples to be probed deeply – for example, stress measurements and studies of granular structure, strain, and microcracks in welds in mechanical structures such as automotive or aeronautical components. In such cases, very high photon energies are used, where the absorption coefficient is low, and the properties of the sample even some centimetres below the surface can be probed. The powder diffraction station at the undulator beamline ID15a at the ESRF has been designed with such experiments in mind. Its source of x-rays allows one to reach photon energies as high as 500 keV.
- The very low beam-divergence required, especially for structure determination, may only be achievable at laboratory-based diffractometers after compromising the flux so much that the experiment becomes impractical.
- Far from any absorption edge of the contributing atoms, the ratios of the intensities of the diffraction spots do not significantly differ. This is not true, however, as one crosses an absorption edge, for which there can be large changes between the relative scattering intensities, which provide added information in the battle to solve the phase problem. This phenomenon of anomalous scattering can only be exploited at a synchrotron facility, where the photon energy can be tuned to lie anywhere around an absorption edge. These techniques are indispensable and are now standard tools in the determination of the structures of complex protein molecules in protein crystallography, covered later in this chapter.
- The intense flux provided by synchrotron sources enables one to measure small volumes of material. This may occur, for example, if (a) individual micronsized crystallites are to be investigated; (b) when an ultra-thin film must be measured; (c) in the investigation of nanomaterials; or (d) when the surface region of a sample, or interface region between two or more heterogeneous crystalline materials, differs subtlely from the bulk structure.
- Some experiments require special and bulky sample environments, such as in-situ, nonambient setups including high-pressure anvils, high-temperature furnaces, film-deposition chambers, cryogenic coolers or large 1- or 2-D detectors. Most of these cannot be mounted on a conventional laboratory-based diffractometer and require the larger experimental 'elbow room' available at most synchrotron-based diffractometers.
- Time-resolved studies with characteristic timescales in the millisecond regime or below, where there is a premium on high flux [4] require the high beam intensities offered by synchrotron radiation. Laue and powder diffraction are especially suited to such studies, as both are capable of recording large volumes of diffraction data in parallel.

Here, we will present, as much as is possible, concepts of x-ray diffraction in qualitative and pictorial terms, restricting the use of mathematical descriptions to a minimum. For the more mathematically-inclined reader, however, we describe XRD using the powerful concept of the convolution theorem at the end of the theoretical section. We then move on to practical applications. No attempt is made to cover all the menagerie of experimental diffraction techniques available, but instead we concentrate on those exploiting the advantages of synchrotron radiation. Before all this, however, we begin with a brief recap of periodic crystal structures.

5.3 Description of Crystals

An ideal crystal consists of an infinite array of identical units separated evenly from each other in three directions in space. Clearly, there exists no ideal crystal in reality, as all crystals are finite in size. In addition to this, all real crystals contain crystallographic defects of some sort. More often than not these define the limits of periodicity. Let us, however, begin with the ideal crystal model as a mathematical construct for understanding diffraction effects.

5.3.1 Lattices and Bases

An ideal crystal can be described in terms of an infinite regular lattice of points in space, the so-called 'Bravais lattice', which maps out the crystal's periodicity, and a group of atoms called the 'basis' that is attached to each and every Bravais-lattice point (Figure 5.1). Together, the lattice and basis form the crystal structure. As we will see, Bragg's law depends only on the lattice, but cannot predict the intensity of the diffraction peaks. For that, the basis must also be known.

One can move from one Bravais-lattice point to another by a translation vector **T** such that

$$\mathbf{T} = u_1\mathbf{a} + u_2\mathbf{b} + u_3\mathbf{c}, \tag{5.1}$$

where u_1, u_2, and u_3 are integers and **a**, **b**, and **c** are the three primitive translation vectors which give the directions and minimum distances required to translate the crystal so that it is indistinguishable from its original position. These vectors form three edges of a parallelepiped of volume

$$V_c = |\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}|. \tag{5.2}$$

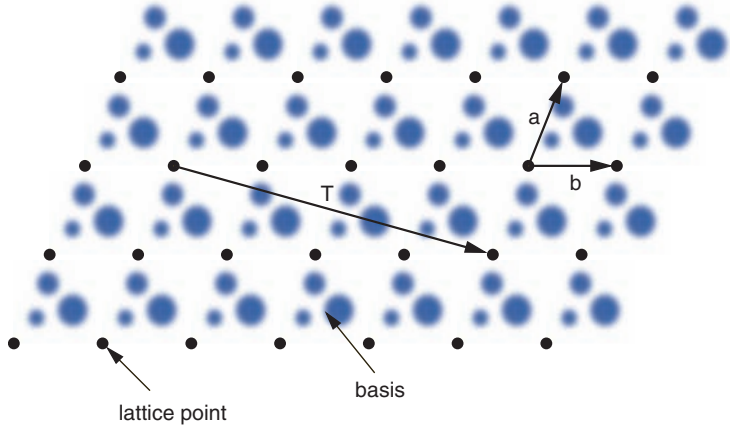

Figure 5.1 *A portion of a two-dimensional crystal structure. All lattice points are separated from one another by a crystal translation vector* **T**, *which is itself composed of an integral number of the primitive translation vectors* **a** *and* **b** *(in this example,* $\mathbf{T} = -\mathbf{a} + 4\mathbf{b}$*). Within the two-dimensional 'volume'* $|\mathbf{a} \times \mathbf{b}|$ *lies the basis, consisting of one or more atoms.*

They describe a primitive cell for which there are lattice points only at the cell's corners. All crystals can be mapped out using such primitive cells. The associated primitive basis contains the minimum number of atoms that can be used to describe the crystal structure and one lattice point each. An important type of primitive cell, called the Wigner–Seitz cell, is constructed by the volume enclosed by planes that perpendicularly bisect lines between pairs of lattice points, shown in Figure 5.2.

Often, however, non-primitive cells and bases are chosen, for reasons of convenience and high symmetry. So, while the simple cubic cell is primitive, the 'conventional' body-centred cubic (bcc) cell is nonprimitive and has twice the volume of the associated primitive cell for a bcc crystal (see Figure 5.2(b)).

The basis associated with the crystal structure consists of N atoms within the volume defined by **a**, **b**, and **c**, such that

$$\text{basis} = \left\{ x_j \mathbf{a} + y_j \mathbf{b} + z_j \mathbf{c}, \quad j = 1 \cdots N \right\} \tag{5.3}$$

where $0 \leq x_j, y_j, z_j \leq 1$. The diamond lattice, shown in Figure 5.3, is face-centred cubic with a basis of two identical atoms at (0,0,0) and (1/4,1/4,1/4).

Note that the lattice parameters **a**, **b**, and **c** are vectors, and therefore their relative orientations are implicitly given. Another way of representing a unit cell, be it primitive or non-primitive, is by using the parameters $a = |\mathbf{a}|$, $b = |\mathbf{b}|$, and $c = |\mathbf{c}|$ and α, β, and γ, as shown in Figure 5.4.

5.3.2 Crystal Planes

A plane can be defined by any three points in a volume, provided they are not collinear. The orientation of a crystal plane defined by three lattice points is specified by the so-called 'Miller indices' (hkl) which are themselves determined in the following manner.

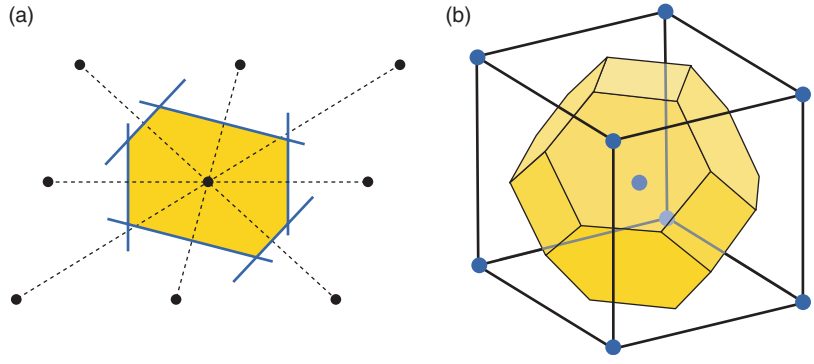

Figure 5.2 *Wigner–Seitz cells. (a) Construction of a two-dimensional primitive Wigner–Seitz cell. Planes which bisect lines joining neighbouring atoms define the boundaries of the Wigner–Seitz cell. (b) The Wigner–Seitz cell (shown in yellow) of a body-centered cubic conventional unit cell (black lines) is a truncated octahedron. The square facets derive from the planes bisecting lines between the body-centre lattice points, while the hexagons are from the planes bisecting corner lattice points to the body-centre lattice point. The volume of the bcc Wigner–Seitz cell is half that of the conventional bcc cell.*

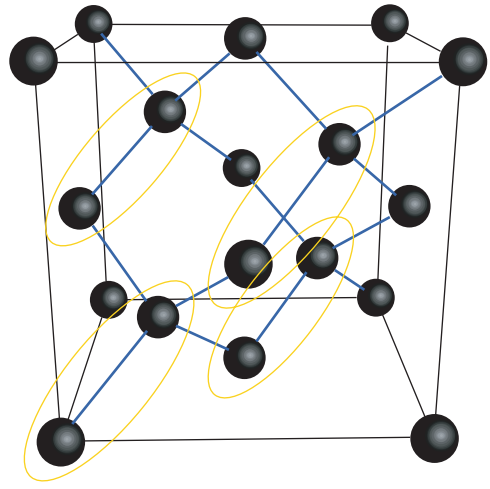

Figure 5.3 *The crystal structure of diamond. Each atom is bonded tetrahedrally to four nearest neighbours. The face-centred unit cell contains four lattice points, each associated with a pair of carbon atoms, highlighted by the ellipses.*

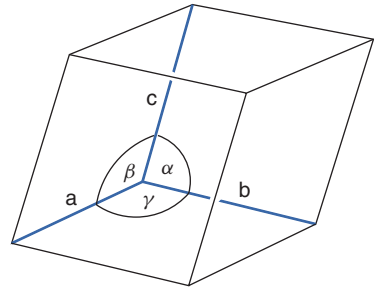

Figure 5.4 *A unit cell of a crystal can be defined by the parameters a, b, c, α, β, and γ.*

- Find the intercepts of the plane on the crystal axes in units of their respective lattice constants a, b, and c.
- Take the reciprocals of these numbers and then reduce these to the smallest three integers that have the same ratio. The result (hkl) is called the index of the plane. Some examples are given for a cubic crystal in Figure 5.5.

An important quantity to determine is the spacing d_{hkl} between (hkl) planes, as shown schematically for a two-dimensional crystal in Figure 5.6.

The interplanar spacing d_{hkl} is in general given by

$$d_{hkl} = \frac{X}{Y}, \tag{5.4}$$

whereby

$$X = \left[1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\,\cos\beta\,\cos\gamma\right]^{1/2} \tag{5.5}$$

and

$$Y = \left[\left(\frac{h}{a}\right)^2 \sin^2\alpha + \left(\frac{k}{b}\right)^2 \sin^2\beta + \left(\frac{l}{c}\right)^2 \sin^2\gamma - \frac{2kl}{bc}(\cos\alpha - \cos\beta\ \cos\gamma) - \frac{2lh}{ca}(\cos\beta - \cos\gamma\ \cos\alpha) - \frac{2hk}{ab}(\cos\gamma - \cos\alpha\ \cos\beta) \right]^{1/2}. \tag{5.6}$$

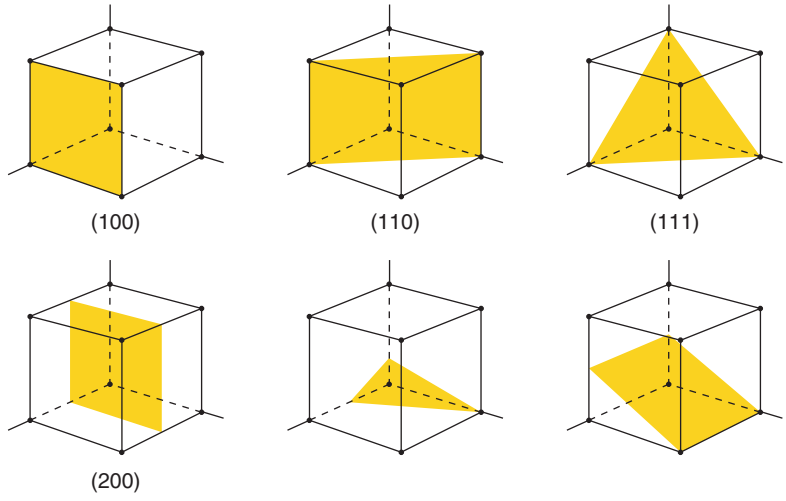

Figure 5.5 *Miller indices of crystal planes in a cubic lattice. The indices of the last two examples are left to the reader to determine.*

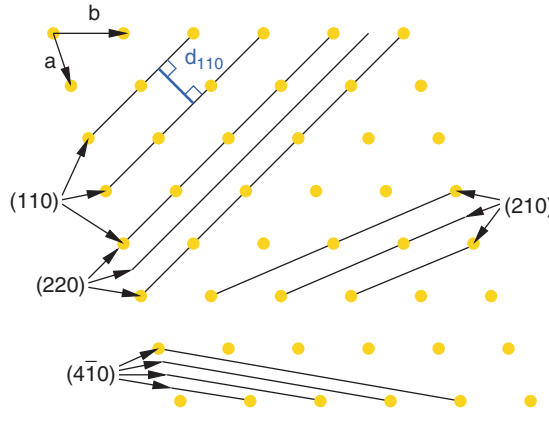

Figure 5.6 *For any given unit cell, the spacing between planes defined by the Miller indices (hkl) can be determined using Equations (5.4), (5.5) and (5.6).*

In general, the interplanar spacing decreases as the Miller indices increase, and the density of lattice points in a plane – that is, the number of lattice points per unit area of the plane – decreases as the Miller indices increase. Note that Equation (5.4) becomes significantly simpler for high-symmetry systems such as orthorhombic, tetragonal, and cubic unit cells, for which $\alpha = \beta = \gamma = 90^\circ$.

5.3.3 Labelling Crystallographic Planes and Axes

A specific crystal axis, normal to the crystal plane, is labelled $[uvw]$. Sets of equivalent axes, due to lattice symmetry, are labelled $\langle uvw \rangle$. Hence, in a cubic crystal, the set of axes $\langle 100 \rangle$ include [100], [010], [001], $[\bar{1}00]$, $[0\bar{1}0]$, and $[00\bar{1}]$, whereas in an orthorhombic cell $\langle 100 \rangle$ only includes [100] and $[\bar{1}00]$.

As described above, specific crystal planes are described by the Miller indices (hkl). Equivalent planes are denoted by $\{hkl\}$.

5.4 Basic Tenets of X-ray Diffraction

5.4.1 Introduction

Perhaps the earliest example of diffraction that a student encounters is Young's double-slit interference experiment. This involves two identical sources and results in the familiar double-slit fringe pattern. What is the difference between interference and diffraction? Essentially, there is none, though one tends to talk of interference effects when a small number of scatterers are involved and diffraction when there is a large array of scatterers.[2]

Diffraction patterns are in fact nothing more than the square of the Fourier transforms of the scattering arrays. To the uninitiated, this may appear to be an unhelpful statement (it isn't!) and immediately begs the question, what *is* a Fourier transform? It lies beyond the scope of this book to give a rigorous description of Fourier theory, for which there are innumerable sources, not least nowadays to be found on the web. Here, I shall attempt to convey the quintessential aspects of Fourier analysis with a simple example.

Essentially, a Fourier representation of an object (that is, its Fourier transform) describes it in term of the frequencies and their relative phase differences that go to make the object up. Hence, periodic things, such as the electromagnetic field of monochromatic light (such as that produced by lasers), or crystals, are particularly simple when viewed from the perspective of their Fourier transforms. Consider the pulse-train periodic structure shown in Figure 5.7(a). For periodic structures in general, their Fourier transforms simplify to 'Fourier series' consisting of discrete components with frequencies which are integer multiples of the the fundamental frequency of the original function. In the case of the structure in Figure 5.7(a), its Fourier series is given by

$$F(x) = Ad + \sum_{n=1}^{\infty} \frac{2A}{n\pi} \sin(n\pi d) \cos(2n\pi x/L), \tag{5.7}$$

whereby A is the amplitude of the structure, l is the pulse width, L is the periodicity, and the duty cycle $d = l/L$. In Figure 5.7, $d = 0.25$. One can reconstruct the original

[2] The verbs are slightly more distinct: objects diffract; waves interfere.

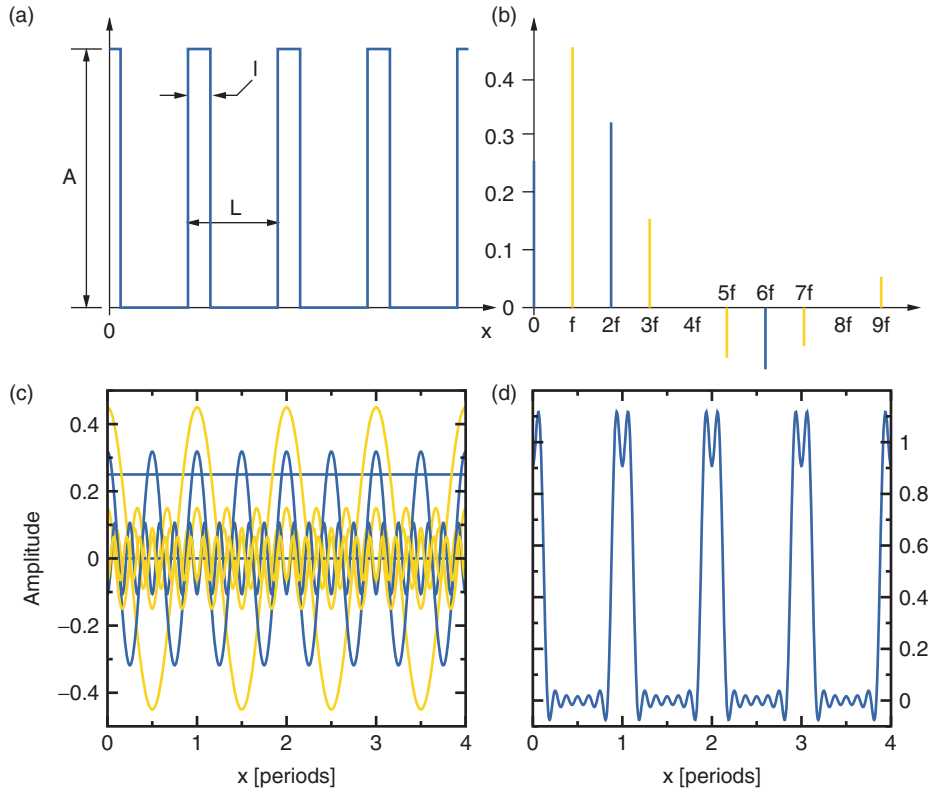

Figure 5.7 *The Fourier series of a pulse train. (a) The periodic structure in space is given by abrupt spike features of width l and height A, separated by a period L, with $d = l/L = 0.25$. (b) The amplitudes of the lowest ten spatial frequency components in the Fourier transform describing this structure, as given in Equation (5.7). The amplitudes of the even components are shown in blue, and the odd components in yellow. (c) The first eight Fourier components (including the zero-frequency component) shown individually, which when added together, (d), already closely approximate the original function shown in (a).*

periodic structure by summing waves having frequencies equal to integer multiples of the structure's spatial frequency $f = 1/L$, each component weighted by the amplitude $(2A/n\pi)\sin(n\pi d)$. The summation is to infinity to reconstruct the original structure perfectly. However, even the summation of the first few components yields a fair approximation (Figure 5.7(d)).

There is an interesting limiting case to this structure, one in which l is infinitely narrow, which in three dimensions describes the array of points of the Bravais lattice of crystals, discussed in Section 5.3. As the limit of $l/L = 0$ is approached, $\sin(n\pi d) \approx n\pi d$, and the amplitudes of the Fourier components are all equal to $2Ad$. In other words, the Fourier transform (or the diffraction pattern) of an infinitely long set of spikes equally separated by L and of the same height A (known as a 'comb function') is another set of spikes of height $2Ad$ and separation $1/L$.

This simple one-dimensional example illustrates that a highly ordered periodic scatterer, such as a crystal, will scatter x-rays in very specific directions to produce a scattering pattern with sharp maxima, the diffraction peaks. The diffraction peaks together make up the (square of the) Fourier transform of the electron-density distribution within the crystal's unit cell. Each peak, a so-called Fourier component, represents a sinusoidal wave of electron density with a certain frequency and direction determined by the peak's position (its (*hkl*)-values) within the pattern. Note that, as each wave component of this pattern must have a spatial frequency equal to a multiple integer of the structure's periodicity, the diffraction maxima (variously also referred to as the 'Fourier components' or 'structure factors') are distributed in a regular, equally spaced array in so-called 'Fourier space'.[3] Once the phase relationship between all these Fourier components is known, the sinusoidal waves they represent, when superimposed upon one another, reconstruct the electron density within the unit cell.

For nonperiodic structures, the Fourier transform is no longer a series of evenly separated and precisely defined components, but contains a continuous distribution of spatial frequencies, although it is no less deterministic in nature. We will discuss the analysis of noncrystalline structures by scattering at the end of this chapter in Section 5.16.

Once Fourier transforms are understood, they help considerably in interpreting the meaning of the profiles of diffraction peaks, the overall envelope of intensity change in a diffraction pattern, and the general shapes of the crystallites contributing to a given diffraction pattern.

To begin, however, we will avoid the formal use of Fourier transforms – we will be using them, but for the innocent bystander this will not be obvious. Instead, we consider the conditions which lead to constructive interference. We will discover that although the periodic spacings in the crystal can be derived from the positions of the peaks in the diffraction pattern using simple geometrical considerations (the Bragg law), the peak intensities are determined by the positions and types of the scattering centres (i.e. the atoms) within the unit cell of the crystal and that it is this information that enables us (sometimes only with considerable effort) to regenerate the atomic basis.

In Figure 5.8, basic examples of scattering from arrays of scattering centres plus their resulting diffraction patterns are given. What general features do diffraction patterns have in common? A few of the most important points are listed here.

- The scattering vector $\mathbf{Q}$, that is, the vector joining up the incoming beam $\mathbf{k}_{\text{in}}$ and the diffracted beam $\mathbf{k}_{\text{out}}$, always lies perpendicular to the scattering planes (see below).
- The sharpness of the diffraction signal is proportional to the number of scattering planes that are involved.
- The separation between the peaks in the diffraction pattern is inversely proportional to the separation of the diffracting planes in real space.
- The maximum number of accessible reflections N in a diffraction pattern is directly proportional to the unit cell volume V and the cube of the photon energy. Precisely,

$$N = 33.5\frac{V}{\lambda^3}. \tag{5.8}$$

[3] The Fourier space is the three-dimensional equivalent of the abscissa in Figure 5.7(b) and is also called 'reciprocal space' (as the axes have dimensions of inverse length), 'k-space' or 'momentum space'.

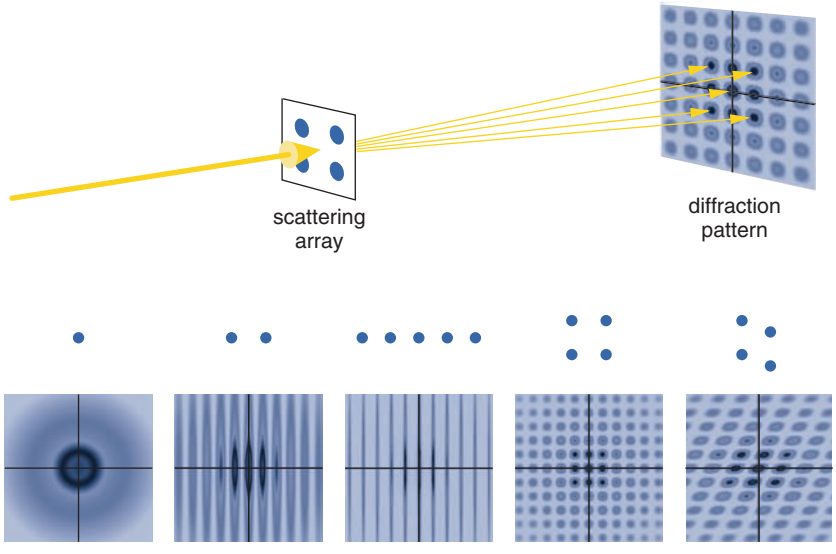

Figure 5.8 *Diffraction from regular one- and two-dimensional arrays of scatterers. The diffraction maxima become sharper with an increasing number of scatterers. Two other important aspects are highlighted in the array on the far right: an increase in distance between the scatterers causes the diffraction spots to move closer in an inversely proportional manner, while an acute angle θ in the monoclinic real-space arrangement results in a diffraction pattern with an obtuse angle of $\pi - \theta$.*

The maximum potential volume of data for protein crystals, which have unit cell dimensions of the order of 100 Å can therefore easily exceed ten million structure factors!

5.4.2 The Bragg Law and the Reciprocal Lattice

The peaks of an x-ray diffraction pattern were explained by W.L. Bragg and his father W.H. Bragg in 1913 by the famous equation named after them

$$m\lambda = 2d \sin\theta. \tag{5.9}$$

Here, λ is the wavelength of the x-ray light, d is the interplanar spacing of the (hkl) planes, θ is the angle of incidence above the plane surface, and m is an integer (Figure 5.9). The arguments in its derivation are the same as those used in obtaining the diffraction angles from a grating in the optical region.

We can re-express the Bragg law in terms of the photon energy by substituting λ with hc/E. This yields

$$\sin\theta = \frac{6.1998}{d[\text{Å}]\,E[\text{keV}]}, \tag{5.10}$$

where d is the spacing of the (hkl)-scattering plane.

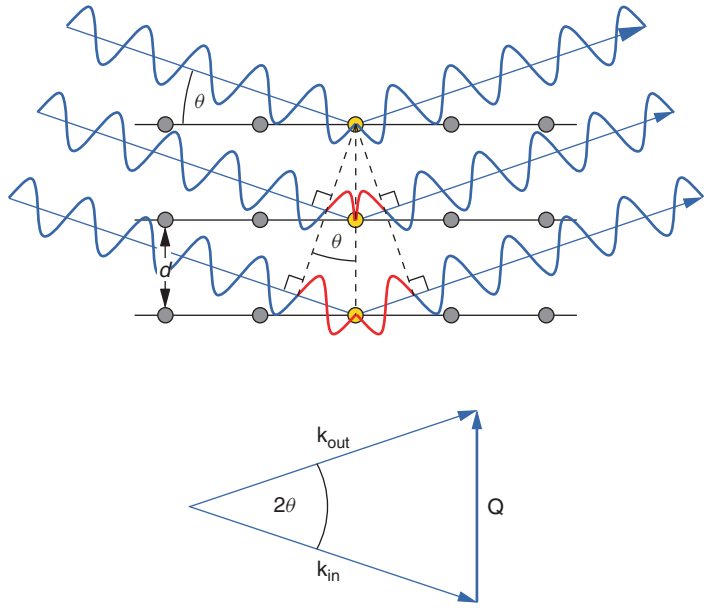

Figure 5.9 *The Bragg law describes interference between rays elastically scattered off successive atomic planes, separated from one another by a distance d. When the optical path difference between adjacent rays is an integer multiple m of the wavelength of the x-ray light that is used, interference is constructive, and a diffraction peak will be seen at that angle. The scattering vector* **Q** *is the vectorial difference between the incoming wavevector* $\mathbf{k}_{\text{in}}$ *and the outgoing wavevector* $\mathbf{k}_{\text{out}}$*. The magnitudes of* $\mathbf{k}_{\text{in}}$ *and* $\mathbf{k}_{\text{out}}$ *are equal to* $2\pi/\lambda$*.*

One of the most important conclusions of Bragg's law is that when Equation (5.9) is satisfied, *the scattering vector* **Q** *always lies perpendicular to the scattering planes, or in other words, the angle subtended by* $\mathbf{k}_{\text{in}}$ *(or* $\mathbf{k}_{\text{out}}$*) and the scattering planes is* θ.

In addition, when the condition for detecting a diffraction maximum is met, the scattering vector **Q** always connects the (000) diffraction spot (in other words, the spot produced by the direct beam) to another maximum in the diffraction pattern.

How are the diffraction maxima positioned relative to one another? We have already argued above that they should be regularly spaced in a three-dimensional periodic array with the (000) direct spot at the centre. The three periodicities describing this array, known as the 'reciprocal lattice', are called the 'reciprocal lattice vectors' and are related to the lattice vectors in real space by

$$\mathbf{a}^* = 2\pi \frac{\mathbf{b} \times \mathbf{c}}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})};$$

$$\mathbf{b}^* = 2\pi \frac{\mathbf{c} \times \mathbf{a}}{\mathbf{b} \cdot (\mathbf{c} \times \mathbf{a})};$$

$$\mathbf{c}^* = 2\pi \frac{\mathbf{a} \times \mathbf{b}}{\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})}. \tag{5.11}$$

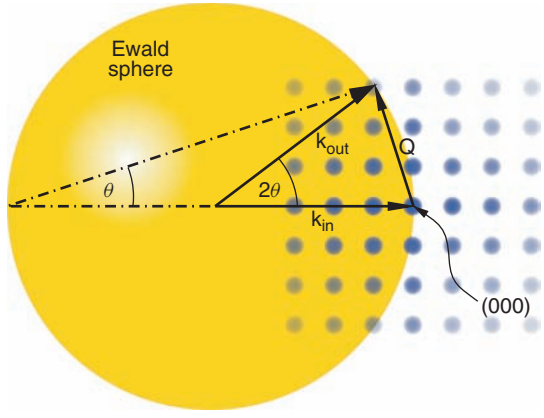

Figure 5.10 *The Ewald construction.*

The denominators in the above three equations are all equal to the unit cell volume V, a scalar quantity without a direction. The numerators, on the other hand, are the cross-products of two of the lattice vectors, and therefore have their direction perpendicular to the planes defined by those two vectors. So, for example, the reciprocal lattice vector $\mathbf{a}^*$ is perpendicular to the plane containing the two vectors $\mathbf{b}$ and $\mathbf{c}$.

The above rules can be very simply and intuitively represented by the so-called 'Ewald construction' in reciprocal space, shown in Figure 5.10: the incident wavevector $\mathbf{k}_{in}$ must end, and the scattering vector $\mathbf{\Delta k} = \mathbf{Q}$, must begin at the (000) diffraction spot of the direct beam, while $\mathbf{Q}$ and $\mathbf{k}_{out}$ end at another diffraction maximum (a 'reciprocal lattice point'). As x-ray diffraction is an elastic process, this means that these two points must lie on the surface of a sphere (the 'Ewald sphere') of radius $|\mathbf{k}|$ whose centre lies at the base of the $\mathbf{k}_{\text{in}}$ and $\mathbf{k}_{\text{out}}$ vectors. This therefore geometrically defines the value of θ (and 2θ) by the magnitude of $\mathbf{k}$. If one rotates the crystal in real space, the array of diffraction maxima is rotated by the same amount in reciprocal space around the (000) point.

Diffraction patterns are all too often plotted out as a function of the scattering angle 2θ. In order to determine the scattering plane separations, however, one must also know the photon energy (or wavelength). It is more convenient, therefore, to plot the pattern out as a function of $\mathbf{Q}$. It is clear from Figure 5.9 that

$$\mathbf{Q} = 2|k|\,\sin\theta = \frac{4\pi}{\lambda}\sin\theta. \tag{5.12}$$

Substituting for $\sin\theta/\lambda$ using Equation (5.9), we obtain

$$|Q| = \frac{2\pi}{d_{hkl}}, \tag{5.13}$$

which, importantly, is independent of the photon energy used. We are able to drop the order m from the equation, as this is implicitly given by the Miller indices (hkl).

The Bragg law reflects the periodicity of the lattice, but tells us nothing of the basis within the unit cell. It is the composition of the basis, however, which determines the relative intensities of the diffraction peaks. We now turn to this problem.

5.4.3 The Influence of the Basis

The earliest structures, including rock salt and diamond, to be investigated by x-ray diffraction by the Braggs could be solved using symmetry arguments and the Bragg law, without having to resort to knowing in detail the scattering strength of each individual atom in the unit cell. Unravelling the structure of more complex crystals would not be so trivial, and it was the son W. L. Bragg who developed the necessary methods to tackle these more complex systems, first shown for the mineral diopside $CaMg(SiO_3)_2$ with Warren in 1928 [5]. A beautiful review of this seminal work and the early history of x-ray analysis is given by one of Bragg's most prestigious students, Max Perutz [6].

Consider Figure 5.11. An incoming x-ray beam is weakly scattered by the electron clouds around the three atoms a1, a2, and a3 of a unit cell. We want to find out how these scattered waves sw1, sw2, and sw3, add up in the direction which satisfies Bragg's law. The amplitude of each scattered wave is proportional to the sum of the scattering strengths of all the electrons attached to the scattering atom, that is, to its atomic scattering factor f, while the phases between the scattered waves depend on the relative positions of the atoms in the unit cell. This is conveniently represented in Figure 5.11(b) in an Argand diagram. The total scattering amplitude, or so-called 'structure factor' $\mathbf{F}_{hkl}$ is the vector sum of the individual atomic scattering factors $\mathbf{f_1} + \mathbf{f_2} + \mathbf{f_3}$, whereby the phases determine their relative orientations. The scattered intensity I_{hkl} is the absolute square of $\mathbf{F}_{hkl}$.

There may arise a situation whereby the magnitude F_{hkl} is zero, that is, the scattering vectors from the individual atoms within the unit cell cancel one another out, in which case no Bragg peak is observed. This is referred to as a 'systematic absence'. For example, it can be easily demonstrated that the structure factor of the bcc lattice is zero

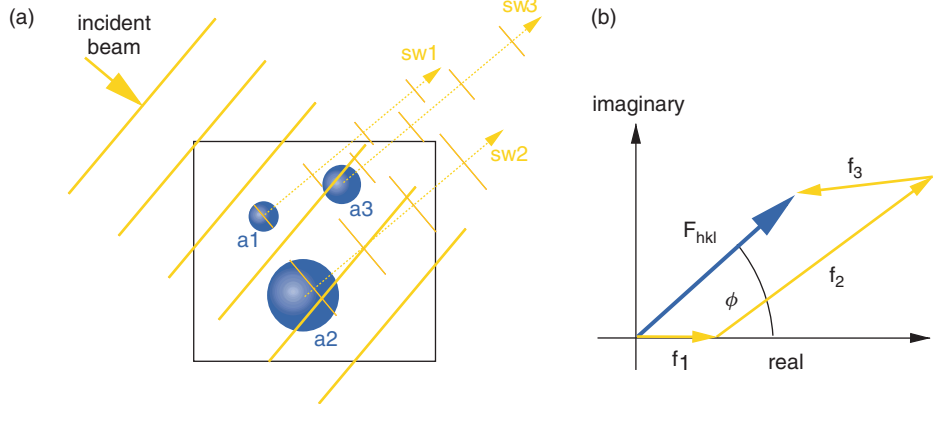

Figure 5.11 *(a) The intensity of a Bragg peak depends on the positions and atomic scattering factors (or 'form factors') of the atoms making up the basis. In this schematic, there are three atoms. The scattered waves sw1, sw2, and sw3 are, of course, spherical, but here we draw them as plane waves in the direction of 2θ for reasons of clarity. (b) The structure factor F_{hkl}, of the wave scattered at the Bragg angle is the vector sum of the contributions of each atom (i.e. f_1, f_2 and f_3), taking into account their phases relative to the incident beam.*

if $h + k + l$ is an odd integer. Hence, reflections such as the (100), (111), (300) are not seen in bcc diffraction patterns. This helps crystallographers identify the structure and symmetry of crystals.

The phenomenon of systematic absences highlights the shortcomings of the simple Bragg law. The reason systematic absences occur, and why some diffraction peaks are strong while others are weak, is because we must consider the positions and scattering strengths of the atoms between the crystal planes, and not just the planes' separation.

In general, in order to determine the relative intensities of Bragg peaks, we must know the atomic form factors f_j of the j atoms making up the basis as a function of scattering angle, as first mentioned in Chapter 2. It is beyond the scope of this text to give a detailed description of their dependence here. As a rule of thumb, however, the atomic form factor increases approximately linearly with the atomic number Z (i.e. the number of electrons) and decreases quasimonotonically with the scattering vector $\mathbf{Q}$ (see also Figure 2.8). Values for f can be found in the *International Tables for X-ray Crystallography*, Volume 3 [7].

5.4.4 Kinematical and Dynamical Diffraction

Until now, it has been tacitly assumed that the interaction between the x-rays and the crystal planes has been weak, and hence that the intensity of the incoming beam at each successive plane is the same. No consideration was made for a drop in intensity of the incident beam due to elastic scattering or absorption, or multiple scattering effects. This simplifies analysis considerably and is known as the *kinematical approximation*. When multiple scattering becomes important in macroscopic crystals, the change in intensity of the incident beam as it penetrates the material can no longer be ignored, in which case one talks of *dynamical diffraction*. One of the consequences of this is that the number of crystal planes required to completely scatter the entire incoming beam (known as the extinction depth) can be significantly less than the absorption depth – this is often the case close to the Bragg condition. As the number of scattering planes is finite, there is therefore a finite width of angles over which interference between the scattered signals from all the involved planes is still constructive, resulting in a top-hat profile for the diffraction maximum. This flat region has a width called the 'Darwin width'.

Also, condensed matter in general does not have the same refractive index as air (albeit only by a small amount, see Equation (2.13)) and thus the wavelength of the x-rays in the medium is marginally larger than in air. This results in an increase in the Bragg angle compared to the value one would expect if refraction effects were ignored.

It lies outside the scope of this course to discuss these effects in detail, although the reader should be aware of them.

5.5 Diffraction and the Convolution Theorem

Before we move on to practical examples of diffraction experiments performed at synchrotron sources, we briefly introduce the more mathematically inclined student to the interpretation of diffraction patterns in terms of the convolution theorem, which can provides insights that can be very useful.

5.5.1 The Convolution Theorem

What is a convolution? Convolutions arise in many fields of science and engineering, such as in image and signal processing, but are also very important in understanding diffraction theory. Firstly, a convolution involves at least two functions to produce a third function. Essentially, it describes the action of measuring the weighted mean of a physical quantity over a certain range. The measuring device is described by the first function, while the measured quantity is the second function. Still confused? Consider the following example. A photodiode having a detection width Δx is scanned (the coordinate u) across an x-ray beam with a perfect 'top hat' profile of width D in which a partially opaque block is inserted (see Figure 5.12).

The recorded profile from the output of the photodiode will be the convolution of the beam profile and the spatial response function of the photodiode. It should be clear to the reader that if $\Delta x \ll D$ and d, then the recorded profile is an accurate approximation of the true profile. In general, the closer the range of the response function of a detecting instrument is to what it is trying to detect, the more convolved will be the output.

Mathematically, a convolution is defined as the integral over all space of one function at x multiplied by another function at $u - x$. The integration is taken over the variable x (which may be a 1-D, 2-D or 3-D variable), typically from minus infinity to plus infinity over all the involved dimensions. Hence, the convolution $C(u)$ between two functions $f(x)$ and $g(x)$ is given by

$$C(u) = f(x) \otimes g(x) = \int f(x)g(u-x)dx, \tag{5.14}$$

where $a \otimes b$ means 'convolved with', and the variable u represents the 'scanning' parameter.

Importantly, the convolution theorem states that if one of the original functions is known, Fourier analysis of the convoluted (measured) signal can be applied to obtain

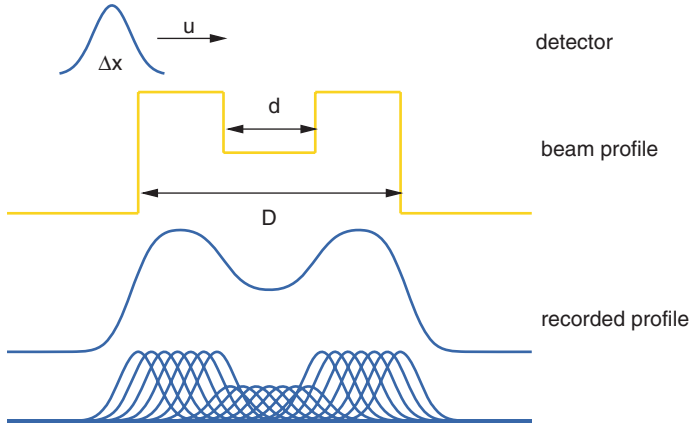

Figure 5.12 *Sketch of a signal resulting from the convolution of a top hat x-ray beam profile of width D illuminating a partially transparent homogeneous block of width d and the spatial response of a photodiode of resolution Δx that has been scanned across the beam along the coordinate u.*

the form of the other function. Precisely, the convolution theorem states that *the Fourier transform of a convolution of two functions is equal to the product of their individual Fourier transforms*. This can also be read backwards as *the Fourier transform of the product of two functions is the convolution of their individual Fourier transforms*. Mathematically, this is written as

$$f(x) \otimes g(x) \leftrightarrow F(k)G(k), \tag{5.15}$$

where $F(k)$ and $G(k)$ are the FTs of $f(x)$ and $g(x)$, respectively. Interchanging f and g has no effect on the convolution (i.e. they commute).

5.5.2 Understanding the Structure Factor

How can we apply the convolution theorem in diffraction theory? The first thing to keep in mind is that a perfect crystal consists of an infinite array of lattice points (the Bravais lattice). Being lattice *points*, each one is of zero extent, that is, a delta function. Each point of the Bravais lattice 'anchors' the electron density within the volume of the unit cell that represents the basis of the crystal, as we have already described at the beginning of this chapter. The effect of convolving any function g with a delta function δ is merely to shift g by an amount equal to the distance of δ from the origin (if you are having problems seeing this, consider again Figure 5.12 and think of the photodiode as having pointlike resolution, that is, $\Delta x \ll D$). One can therefore think of the electronic distribution of the crystal as the convolution of the Bravais lattice with the electronic distribution of the basis, as shown schematically in Figure 5.13(a).

Let us now look at these two aspects, the lattice array and basis, one at a time. The lattice array is a three-dimensional comb function, which we have already met in Section 5.4, that is, a set of equally spaced, infinitely narrow spikes. The Fourier transform (FT) of a comb array is another comb array with separations inversely proportional to the separations in the real-space comb array. Using the convolution theorem, the diffraction pattern of an infinite crystal (assuming the kinematical limit) is the product of the comb function FT of the Bravais lattice and the FT of the electronic distribution within a unit cell. This latter FT is the 'structure factor' and therefore determines the intensities of the peaks in the diffraction pattern (Figure 5.13(b)).

Lastly, it is briefly mentioned that we will return to the convolution theorem in Section 5.13 in helping us understand features of surface diffraction. It is therefore recommended that all but the most extreme mathematics-phobe should take five (or ten) minutes to understand some of this subsection.

5.6 The Phase Problem and Anomalous Diffraction

5.6.1 Introduction

Given an electron density in real space within a unit cell of a crystal, the calculation of the resulting diffraction pattern is routine. However, the bread-and-butter work of a crystallographer consists of the reverse process – determining the electron density in real space from the recorded diffraction pattern.

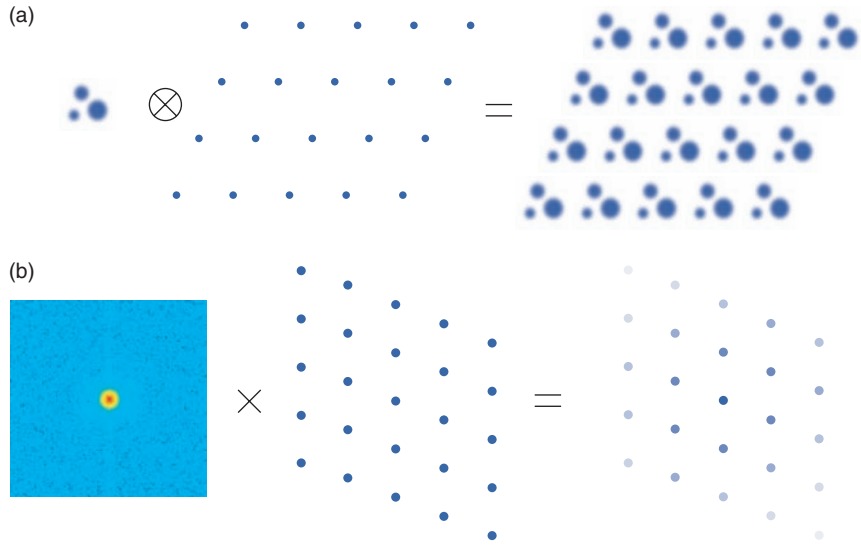

Figure 5.13 *Diffraction pattern intensities explained using the convolution theorem. (a) An infinite crystal can be thought of as the convolution of the electronic distribution within the unit cell with the Bravais lattice. (b) Using the convolution theorem, it becomes apparent that the diffraction pattern is the product of the FT of the electronic distribution in the unit cell (called the 'structure factor') and the FT of the Bravais lattice, which, like the Bravais lattice itself, is another infinite 3-D comb function (see text).*

It has already been mentioned in the introduction to this chapter that, because we measure the *intensity* of x-ray diffraction signals ($|F_{hkl}|^2$, a scalar quantity), and not the amplitude (i.e. the *complex* structure factor, which has both a magnitude *and* a phase), we lose half the information about the system we are measuring, namely the phase differences between the elastically scattered waves from the different atoms within the unit cell.

Referring back to Figure 5.11, we know that the structure factor that results in the measured diffracted intensity $I_{hkl} = F_{hkl}F^*_{hkl}$ must lie in the Argand diagram on a circle of radius $\sqrt{I_{hkl}}$, as shown in Figure 5.14.

The information we have at our disposal includes the integrated intensities of the Bragg peaks of our diffraction pattern, I_{hkl}, the atomic scattering factors of the constituent atoms as a function of $\sin\theta/\lambda$, and the angles of the crystallographic planes relative to the incoming beam and Bragg spots. For a given set of atomic positions within a unit cell, the phase relationship between the atomic scattering factors (given by the angles between the red, green, and yellow arrows in Figure 5.14) changes in a predetermined way from Bragg peak to Bragg peak. Our task is to find the atomic positions that satisfy the Bragg intensities for all recorded diffraction spots. The more Bragg peaks are recorded, the tighter becomes the set of conditions which must be satisfied to achieve a self-consistent model. The more diffraction data one records, the more difficult it is to obtain self-consistency, but once this is achieved, the more confident one is that this represents the real structure.

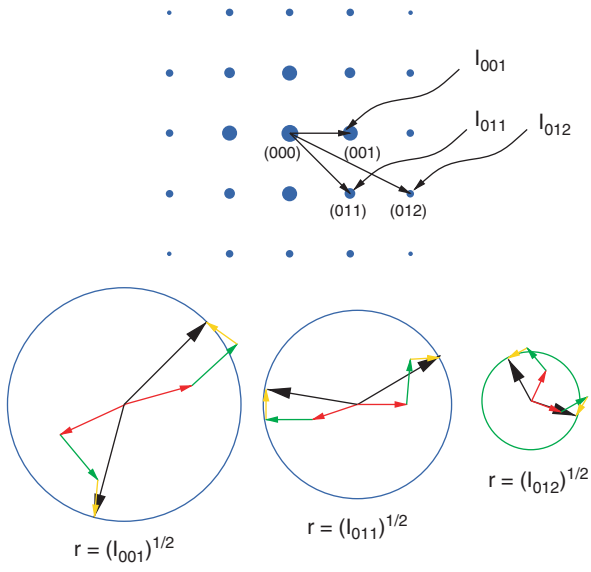

Figure 5.14 *The phase problem. Different phase relationships between the constituent atoms of a unit cell can lead to the same total scattering amplitude* $r = |F_{hkl}| = \sqrt{I_{hkl}}$*, although the direction may be different.*

For unit cells containing only a few atoms, a trial-and-error procedure using the brute force of modern computer power can rapidly deliver the unit-cell structure (often in conjunction with some reasonable physical constraints, such as, for example, that atom centres cannot be closer to one another than 1 Å – one of the smallest atomic separations in condensed matter is approximately 1.25 Å, between adjacent ions in LiF). This approach becomes more impractical as the complexity of the unit cell increases. In the case of macromolecular unit cells, such as protein crystals, other supplementary information must be provided. This is discussed in more detail in Section 5.11.

5.6.2 The Patterson Map

However, all is not lost. We can obtain valuable information about the unit cell even without knowing the phases. It turns out that the inverse Fourier transform (IFT) of the diffraction pattern intensities yields the *relative* distances and angles between the atoms in a unit cell, referred to as the Patterson map (PM) or Patterson function, as shown schematically in Figure 5.15.[4] In fact, one can easily demonstrate that the Patterson map is in fact the convolution of the unit cell with its inverse.

In addition to providing the relative distances and angles of atoms within a unit cell, the PM exhibits two other important features. Firstly, the number of maxima, including the central maximum, in the PM of a unit cell containing N atoms is equal to $1 + N(N - 1)$, which for large N is approximately equal to N^2 – a PM can soon become a crowded

[4] This is not a treatise on crystallography, and the reader is referred to the many excellent textbooks on crystallography available in the literature to understand the derivation of the Patterson function.

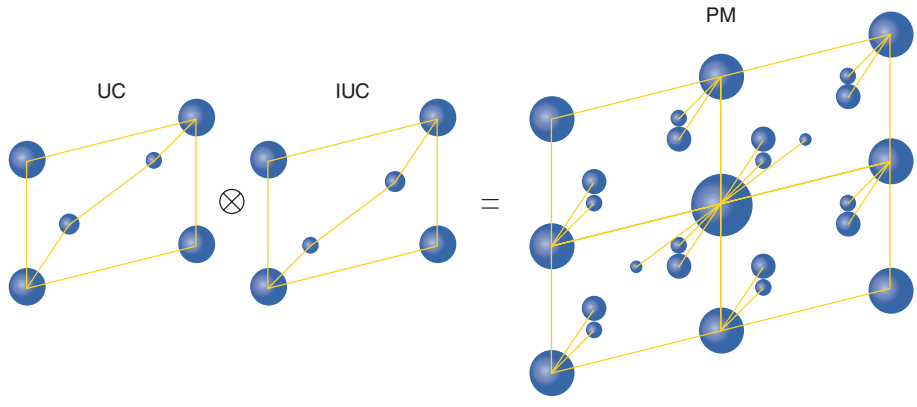

Figure 5.15 *A Patterson map (PM) is the convolution of the unit cell (UC) with its own inverse (IUC).*

affair. Note also that because the vector connecting a given atom A to another atom B is the inverse of that connecting atom B to atom A, a Patterson map has centrosymmetry around its origin.

Importantly for the technique of multiple isomorphous replacement, (MIR), a technique used extensively in protein crystallography to help overcome the phase problem, and discussed in more detail in Section 5.11, the intensity of the maximum in a PM associated with two atoms A and B is proportional to $Z_A Z_B$, the atomic numbers of those atoms. Thus, peaks corresponding to the distances and angles between heavy atoms dominate PMs.

5.6.3 Friedel's Law and Bijvoet Mates

George Friedel, a French crystallographer in the nineteenth and twentieth centuries, discovered that the Fourier transform of a *real* object (i.e. one without imaginary components) has Fourier components $F(k)$ and phases $\phi(k)$ with the following properties:

$$\begin{aligned} F(k) &= F^*(-k), \\ |F(k)|^2 &= |F(-k)|^2, \\ \phi(k) &= -\phi(-k), \end{aligned} \tag{5.16}$$

whereby F^* is the complex conjugate of F. This is Friedel's law. Far from absorption edges, the electron cloud within the unit cell of a crystal is indeed a real quantity, hence diffraction patterns exhibit centrosymmetry, that is

$$I(hkl) = I(\overline{hkl}). \tag{5.17}$$

$I(hkl)$ and $I(\overline{hkl})$ are referred to as Friedel pairs, or Friedel mates. This can also be understood by considering Figure 5.16.

The relative phases of the scattered x-rays from the component atoms of the unit cell (ignoring for the time being absorption phenomena) are directly proportional to the differences in their distances away from the Bragg planes. Hence, the phase diagram for

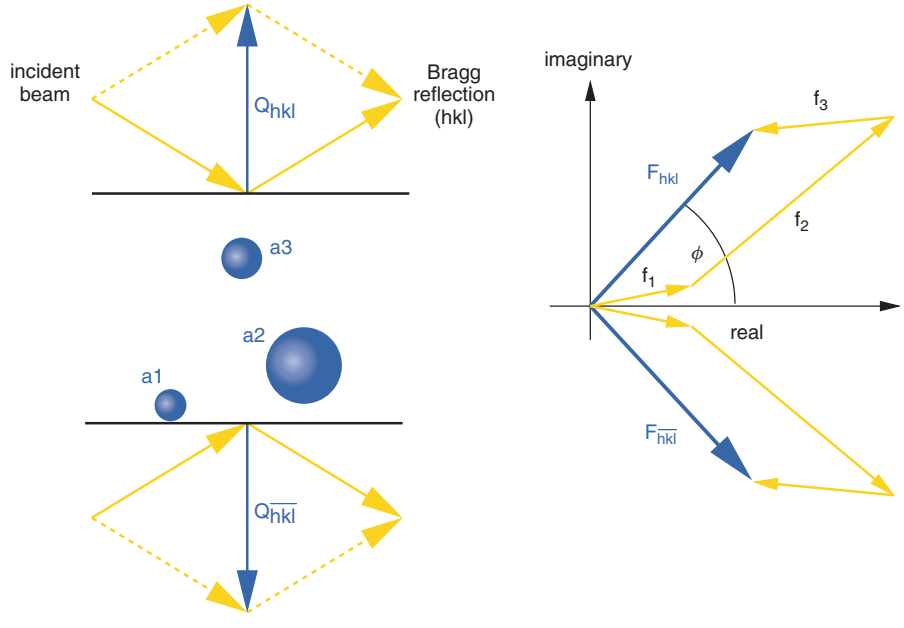

Figure 5.16 *Friedel's law. A Bragg reflection (hkl) is shown for a fictitious set of three atoms a1, a2, and a3, between Bragg planes. The same atoms will produce the Friedel mate ($\overline{hkl}$) for the same diffraction condition, but this time for a beam incident from below.*

a Bragg peak from an incident beam from below the unit cell in Figure 5.16 will mirror that of a Bragg reflection from an incident beam approaching from above. The scattering vectors are equal in magnitude but opposite in direction. If we label the Bragg reflection on the upper side as (hkl), the one on the lower side is $(\overline{hkl})$.

For unit cells with internal symmetry, there will be additional reflections with the same intensity, such as the (hkl) and $(hk\bar{l})$ reflections in cubic, tetragonal, orthorhombic, and hexagonal systems. These equivalent reflections are referred to as Bijvoet mates.

As one approaches the energy of an absorption edge, however, Friedel's law breaks down. We now discuss this in more detail.

5.6.4 Anomalous Diffraction

Here, we will look more closely at the change in the structure factors of crystalline materials in the region of an absorption edge. We begin by studying the illustrative example of the compound semiconductor, GaAs. Next, we show how the breakdown of Friedel's law in this region can provide valuable extra information about the atomic structure. This will prepare us for methods used for structure determination in protein crystallography described in Section 5.11.

5.6.4.1 Tuning the Structure Factors

The intensity of a Bragg peak depends both on the relative positions of the atoms in the unit cell, and on the complex form factor, $f_1 + if_2$, of each atom (see Figure 5.11).

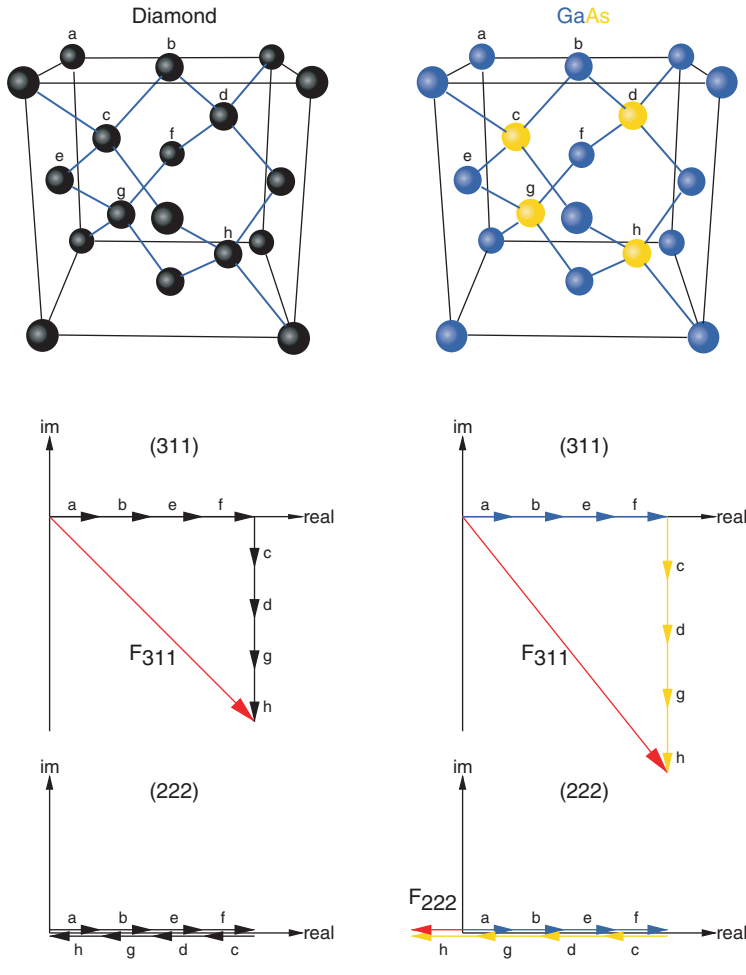

Figure 5.17 *The face-centred cubic unit cells of diamond and GaAs. In the case of diamond, elastic scattering from the atoms a, b, e, and f interfere destructively in the (222) direction with that from atoms c, d, g, and h and there is thus no Bragg peak here. Away from absorption edges, the magnitude of the structure factors of Ga and As are different (see Figure 5.18), and hence the (222) reflection is nonzero. The (311) structure factors are nonzero for both crystal types.*

Because the atomic form factors change with photon energy, especially in the neighbourhood of an absorption edge, the intensities of the Bragg peaks will change.

It can be easily demonstrated that the structure factor of diamond is equal to $4f_C(1-i)$ for the (311) reflection (where f_C is the atomic form factor of carbon), but is equal to zero for the (222) reflection, that is, it has a systematic absence.[5] We show this in the Argand diagram of Figure 5.17 – the structure factor F_{311} for the (311) reflection

[5] More generally, it can be (relatively) simply shown that crystals having the diamond structure have allowed Bragg reflections either when $h+k+l=4n$, whereby n is an integer and h, k, and l are all even, or when h, k, and l are all odd.

is always nonzero, while, in the case of the (222) reflection of diamond, half of the scattering contributions from the eight carbon atoms in the unit cell cancel out the other half, resulting in a null structure factor $F_{222} = 0$.

GaAs has a zincblende structure, which differs from that of diamond in that half the atoms (c, d, g, and h in Figure 5.17) differ from the other half. As shown in Figure 5.17, the structure factors for the (311) and (222) reflections are simply

$$F_{311} = 4(f_{Ga} - if_{As}) \tag{5.18}$$

and

$$F_{222} = 4(f_{Ga} - f_{As}). \tag{5.19}$$

The atomic numbers Ga and As only differ by 2, hence they appear to be similar from the viewpoint of elastic scattering of an x-ray photon (i.e. their atomic form factors are in most energy regions fairly similar). At an absorption edge, the difference in f_1 (the real part of the form factor, and in general significantly larger in magnitude than f_2, the imaginary part) can vary significantly. As can be seen from Figure 5.18(a), the difference is largest at 10.367 keV, the absorption edge of Ga. At 11.5 keV, however, the two curves for $f_1^{(Ga)}$ and $f_1^{(As)}$ cross one another, and therefore here, the GaAs (222)-reflection intensity is very close to zero, as indeed observed in Figure 5.18(c). Indeed, if this structure factor were recorded from 11.5 keV to approximately 12.30 keV, where $f_1^{(Ga)}$ and $f_1^{(As)}$ again cross, one would observe that it grows from being negligible at 11.5 keV to exhibiting a maximum at 11.867 keV, before dropping again to nearly zero at 12.30 keV.

5.6.4.2 The Complex Electron Density – Breaking Friedel's Law

In reasoning that the intensities of Friedel pairs are equal, we were careful to include the caveat that the x-ray photon energy should lie far from any absorption edge of the atoms within the unit cell. The reason why the individual atomic form factors f_1, f_2, and f_3 in the Argand diagram of Figure 5.16 have imaginary components and different angles is entirely due to the positions of the respective atoms between the Bragg planes, which results in interference between the scattered waves. It is not due to absorption. So what happens when we *do* introduce absorption?

Consider the atom selenium exactly at its K-edge at 12.658 keV. As we have seen for the example of silicon (Figure 2.10), the real part of the atomic form factor f_1 becomes smaller at the absorption edge, while the imaginary component f_2 increases sharply. In the case of Se, f_1 decreases from approximately 34 (the number of electrons in a Se-atom) to less than 22, while f_2 increases from approximately 0.3 to 3.75. In other words, far above the absorption edge, $f_2/f_1 \approx 0.01$, while at the K-edge, it is 0.17. Crucially, the absorption term of the atomic form factor f_2 always lies at 90° to f_1, and, in contrast to the phase differences caused by the atoms' positions within the unit cell, the sign of this phase remains invariant for both Friedel mates. The result is that the magnitudes of the total structure factors F_{hkl} and $F_{\overline{hkl}}$ are no longer equal and Friedel's law breaks down. This is shown schematically in Figure 5.19 for a Se-atom in combination with three other low-Z elements.

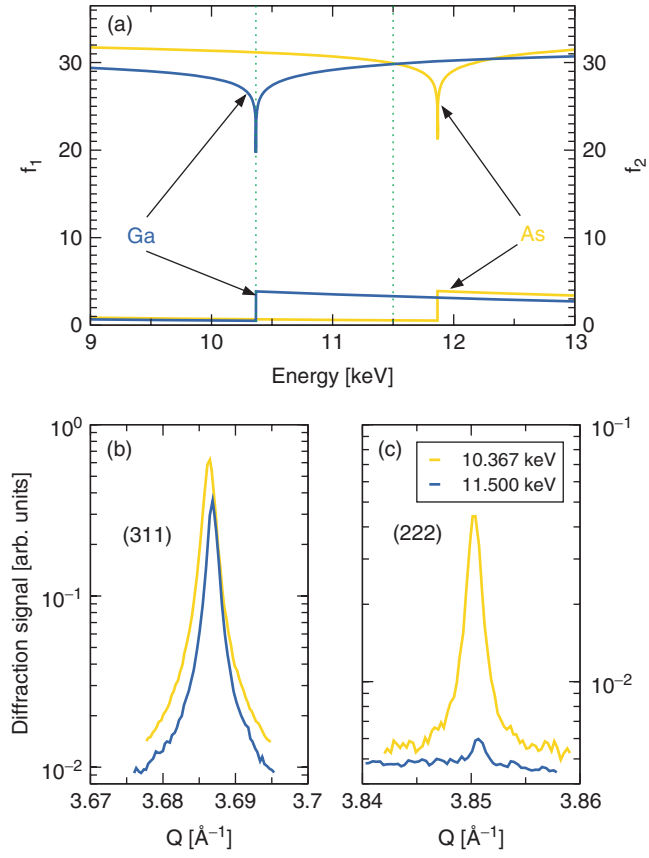

Figure 5.18 *(a) Variation of the atomic scattering factors f_1 and f_2 for Ga and As near their K-absorption edges. The dotted lines represent the two energies used to record the GaAs(311) and (222) reflections, shown in (b) and (c). Courtesy Bruce Patterson, Paul Scherrer Institut.*

The derivation of Friedel's law is based on the assumption that the electron-density distribution is real. Its breakdown therefore implies the introduction of an imaginary component to the electron density, which is merely a manifestation of absorption (remember from Equations (2.12) and (2.14) that the refractive index has both a real and imaginary part). This breakdown increases the number of inequivalent Bragg reflections and hence also the amount of information available to solve the crystal structure.

Note that only the tunability of synchrotron radiation allows one to perform these anomalous scattering experiments. Exactly how one extracts this added information from data sets recorded at different photon energies across an absorption edge is discussed in more detail in Section 5.11.

5.6.5 Direct Methods

In the last part of this section, we briefly discuss so-called 'direct methods' for retrieving the phase information. There are many variants to direct methods, but they all share

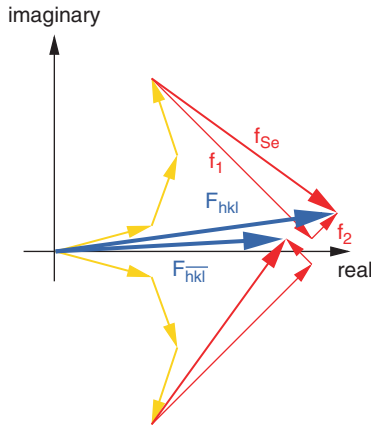

Figure 5.19 *The effect of anomalous scattering. Consider a unit cell consisting of a single Se-atom and three low-Z atoms. The atomic form factor of f_{Se} (shown in red) dominates the phase diagram. At the K-absorption edge of Se, the imaginary component of f_{Se}, f_2 becomes significant. This is included in the phase diagram as a vector at 90° from the real component f_1. Importantly, this phase of f_2 with respect to f_1 remains the same in the Friedel pair reflection, while all other phases are reversed, resulting in the total structure factors F_{hkl} and $F_{\overline{hkl}}$ having different magnitudes and directions that are not mirrored across the real axis.*

the common property that they are iterative algorithms that search for self-consistency between a priori available information (so-called 'constraints') in real and in reciprocal space. It is briefly noted that at the present level of sophistication, direct methods in crystallography are limited to unit cells containing at most a few hundred atoms, and, despite some successes with small protein structures, are not usually capable of solving macromolecular structures.

Traditionally, structure determination from diffraction data follows model-fitting methods, such as Rietveld refinement [8]. Model fitting suffers from several limitations, which become increasingly severe with the complexity of the unit cell under investigation. Perhaps the most serious problem lies in the fact that these methods normally apply a 'goodness of fit minimization' algorithm, which means that they can become trapped in local minima within the fitting-parameter space being used. There are a few 'clever' search techniques which allow one to escape local minima and find the global solution, such as so-called 'simulated annealing', or Monte Carlo sampling. This comes, however, at considerable cost in computing speed.

Another problem with model fitting is that the starting model is, to an extent, a subjective quantity. Based perhaps on previous chemical or physical knowledge of similar systems or information gleaned from complementary techniques such as nuclear magnetic resonance, infrared spectroscopy, or Raman spectroscopy, an eminently reasonable starting guess may be chosen that may in fact lie far from the true solution. The likelihood of this happening increases with the complexity of the system. It would therefore be advantageous to use structural-solution methods which are model-independent.

Iterative phase-retrieval methods (a subclass of direct methods) provide such a possibility [9]. Although they vary in detail, they all rely on alternating between real- and

reciprocal space, whereby 'constraints' are imposed on both the diffraction data and the real-space electron density until self-consistency is achieved [10–12]. Starting with *any* random model composed of the known amplitudes of the structure factors $F_{hkl} = \sqrt{I_{hkl}}$ plus a random set of phases $\{\phi\}$, the basic concept of these algorithms follows four simple steps (see also Figure 5.20):

1. Inverse Fourier transform the data (including the phases) to yield a real-space model.
2. Apply real-space constraints to this model. Examples of such constraints might include replacing any negative electron densities (which are obviously unphysical) that the IFT has generated with positive values, or that any electron density outside the known unit-cell volume is removed. In more sophisticated algorithms, such as the so-called 'hybrid input-output' algorithm first proposed by Fienup [11], the input to the real-space domain need not comply strictly with the constraints (e.g. electron-density positivity or the maximum allowed size and shape of the real-space object), but can instead be adapted by using information about the known shortcomings of the input used in the previous iteration. In this manner, it is possible to avoid being trapped in local solution minima.
3. Fourier transform this modified real-space model.
4. Keep the resulting phases, but replace the calculated moduli of the FT by the square-root of the measured diffraction intensities (this is the constraint in reciprocal space).

These steps constitute a single iteration of the algorithm, and are repeated until a state of self-consistency is reached between the real-space and reciprocal-space domains.

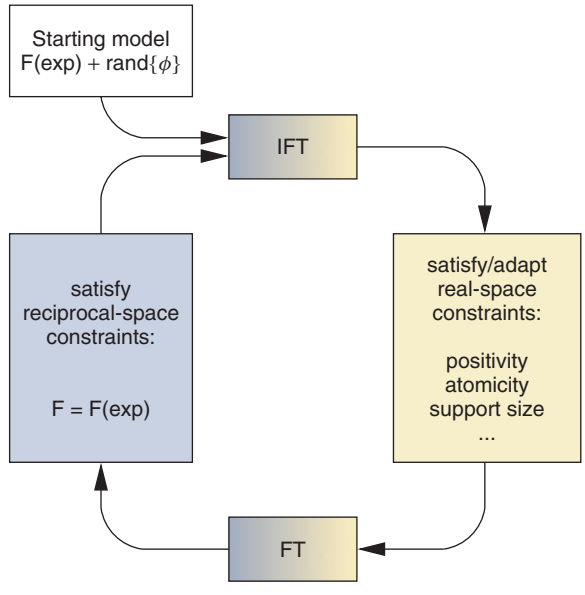

Figure 5.20 *Iteration scheme for iterative direct methods algorithms. Note that in more sophisticated algorithms, the output from the inverse Fourier transform (IFT) must not strictly comply with the real-space constraints, but can be adapted in an appropriate manner based on the shortcomings of the previous iteration.*

5.7 Types of Crystalline Samples

Before discussing different types of synchrotron-based diffraction experiments, it is worth first pausing and considering the myriad types of samples that can be investigated using XRD, as an understanding of these helps in deciding the type of experiment to undertake (see Figure 5.21).

We can divide sample types into three broad classifications, namely (and in increasing degree of disorder) single-crystal, textured, and powder samples. Within each class, however, there are strikingly different manifestations. Take, for example, single crystals. Even the highest-quality silicon single crystal still contains crystallographic defects (including vacancies, interstitials, dislocations, and mosaic domains), hence in reality there is no such thing as a perfect single crystal. At what level of imperfection does one begin to regard a system as being no longer a single crystal? Of course, there is no well-defined boundary, and depending on experimental methods, preconceptions, and 'what you're used to', the degree of perfection may be very different. So high-quality large crystals such as diamonds, micron-sized protein crystals (which, incidentally, size-for-size, are far more valuable than diamond!) or heteroepitaxial thin films grown on single-crystal substrates and fabricated for microelectronic applications are all, in their own fields, regarded as being single crystals. This fuzziness in boundaries extends all the way to nanocrystalline powder samples and semi-amorphous structures.

A well-known phenomenon in single crystals is that of twinning, in which two crystals share crystal lattice points at a boundary in a symmetric manner. The resulting diffraction pattern of the twins (there can be more than two twins, depending on the crystallographic system) becomes more complicated. For example, each diffraction spot of an untwinned system may be split into two or more closely lying features. One can therefore consider twinning as the first step towards disorder and eventually to powder samples.

Textured samples exhibit an intermediate degree of order, with preferred directions for certain crystallographic planes. This may be due, for example, to a 'brushing' effect in the fabrication process of the sample (such as in polymer fibre pulling), to the synthetic process, such as in natural fibres such as flax, or because of energetic reasons.

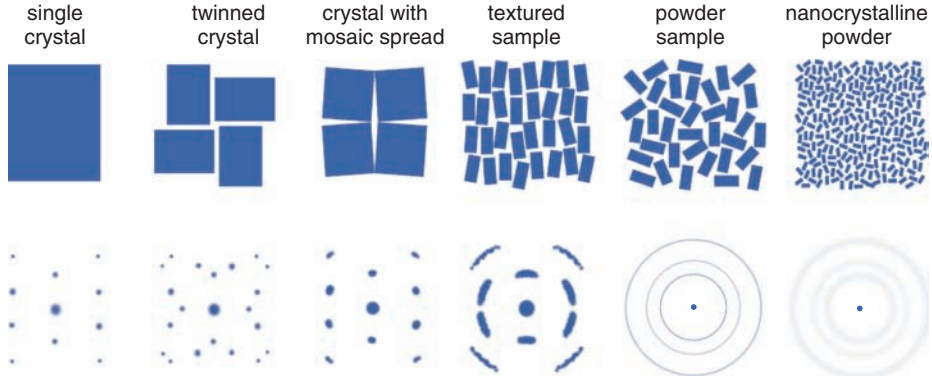

Figure 5.21 *Schematic of different crystalline sample types and their diffraction patterns.*

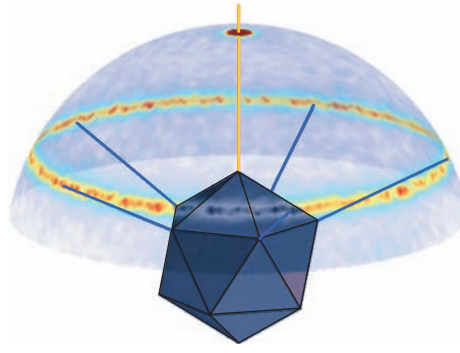

Figure 5.22 *Pole figure of a QC-film mapped on a spherical surface. An icosahedron is shown at the centre as an aid to understanding the pattern: the fivefold symmetry axis pointing perpendicularly out of the film is shown in yellow, while the five other fivefold axes, shown in blue, produce the ring feature, as the icosahedron is rotated around the yellow axis. Courtesy Christian Schlepütz, Paul Scherrer Institut.*

Figure 5.22 shows texturing in a so-called 'quasicrystalline' thin film of a Ti-Ni-Zr alloy exhibiting icosahedral symmetry. The angular distribution of the Bragg peak associated with the fivefold crystalline-axis was probed in a so-called 'pole figure'. In pole figures, the detector is fixed to a certain 2θ-value of interest. The sample is then rotated azimuthally by as much as 360°, and the polar axis is tilted by up to 90°. In this manner the orientations of the crystal planes associated with the 2θ-value are probed over as much as a hemisphere (2π-steradians). The pole figure of the quasicrystalline film showed a strong signal when the scattering vector was parallel to the film normal. Because this quasicrystal film has icosahedral symmetry, one should therefore expect additional signal for this Bragg angle somewhere in the circle at a polar angle of 63.435° associated with the subtended angle of the five adjacent vertices of the icosahedron (shown as blue axes in Figure 5.22) with the normal vertex (shown in yellow). In fact, at this polar angle, signal was found continuously for all azimuthal angles, indicative that although there is a preferred out-of-plane orientation, there is no in-plane texturing. The reason for the film having out-of-plane texture is that the fivefold-symmetric plane contains the highest density of atoms and therefore has the lowest surface energy. Because the film was grown on sapphire (with a hexagonal sixfold surface symmetry), there was no natural in-plane orientation, as demonstrated by the homogeneous ring feature.

A true powder sample consists of crystalline grains that have no preferred orientation in space, and are significantly smaller than the illuminated volume. For some samples, this latter condition may be only poorly met, in which case the sample can be spun, in order to record a spatially averaged signal. Note that for undulator radiation, the focused beam cross-section might be only of the order of a square micron, and that the sampled volume may be less than $100\,\mu m^3$.

When the characteristic size of the individual grains is reduced to only a few unit cells, such as in nanocrystals, the widths of the diffraction peaks broaden accordingly, producing diffuse ring patterns. This lower limit in crystallinity often best represents the grim reality of 'real' systems, and as such is a burgeoning area of research in twenty-first-century crystallography.

5.8 Single Crystal Diffraction

5.8.1 Laue Diffraction

Although the majority of x-ray diffraction methods use (quasi-)monochromatic x-rays and a sample which one has to orient in space in order to satisfy the diffraction condition, it is also possible to record diffraction patterns of stationary single crystals using a broad spectrum of x-rays, a technique known as the Laue method. Indeed, the Laue method is historically the oldest and was used in the original discovery of x-ray diffraction and is named after its founder, Max von Laue (see Figure 1.7). After the development of monochromatic sources using characteristic radiation from x-ray tubes, Laue diffraction fell somewhat into disuse, except to orient single crystals and to determine their crystal quality. However, with the advent of synchrotron radiation (especially from third generation facilities), and enormous improvements in computing power and detector technology, Laue diffraction has enjoyed a renaissance as a technique which provides a fast and efficient means to record diffraction data, and has been used to great effect in dynamical studies of transient crystalline states, not least in protein crystallography [13], as discussed in more detail in Section 5.11.6.

Consider a static single crystal illuminated by a broad and continuous spectrum of x-rays (Figure 5.23). Each set of crystal planes (hkl) is at a well-defined angle θ_{hkl} relative to the incident beam. From Bragg's law, Equation (5.9), it is evident that, given a sufficiently broad spectrum of x-rays, there will be a specific wavelength λ_{hkl} that satisfies the Bragg condition, such that

$$\lambda_{hkl} = 2d_{hkl} \sin\theta_{hkl}. \tag{5.20}$$

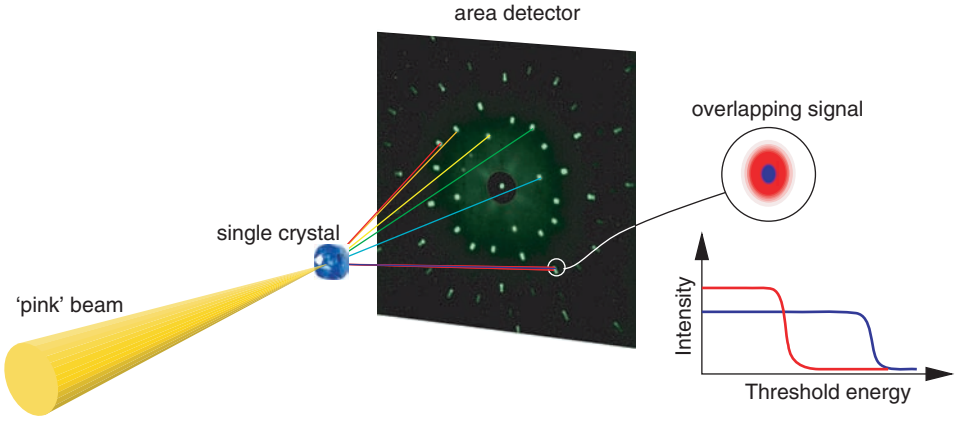

Figure 5.23 *A polychromatic 'pink' beam is focused on to a stationary single crystal sample. Because there is a continuum of wavelengths in the beam, there will always exist certain wavelengths which satisfy the Bragg condition for any given set of crystal planes and orientation. Higher-order reflections can overlap, whereby the photon energies are a multiple of that of the first-order diffraction spot. These overlapping signals can be separated by scanning the lower threshold energy in modern pixel detectors (see Section 4.6.6).*

Hence, each diffraction peak has a different 'colour'. Importantly for time-resolved studies, all possible Bragg reflections can in principle be simultaneously recorded.

In modern synchrotron-based Laue-diffraction studies, the mode of operation is almost always that of transmission (in contrast to the 'back-reflection' mode, for which $2\theta > 90\,^{\circ}$). The region of reciprocal space that can be accessed for a given setup is defined by the wavelength range of the polychromatic beam (i.e. λ_{max} and λ_{min}) and the half-angle $2\theta_{max}$ subtended by the area detector. This is shown succinctly in the Laue reconstruction of Figure 5.24.

Laue diffraction therefore provides a lot of structural information in a very short time. However, it is not as well-suited as monochromatic scattering for determining the full atomic structure of a crystal, due on the one hand to the often complex and unknown intensity distribution of the 'pink' incident x-ray beam.[6]

In addition, families of lattice planes that are parallel to one another, for example the (111), (222), (333) ... planes, have Laue diffraction maxima overlapping at the same position, resulting in a loss of information. This is called the 'energy overlapping problem'.

Recently, however, this positional degeneracy has been lifted by the availability of energy-dispersive pixel detectors, as described in Section 4.6.6 – by tuning the threshold energy, reflections below a chosen value are suppressed. A new recording mode thus presents itself. The threshold energy of the area detector is scanned and successive images

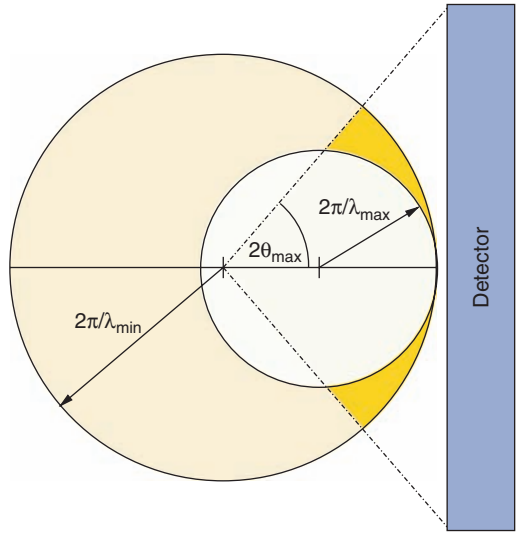

Figure 5.24 *The volume of reciprocal space that can be simultaneously accessed in Laue diffraction, shown here in bright yellow, depends on the range of photon energies of the polychromatic beam and the maximum angle that can be subtended by the area detector.*

[6] The polychromatic beam is referred to as being 'pink' (rather than 'white') because it has been reflected at low angles from mirrors and/or focusing optics, which filter out the more energetic photons with their lower critical angles for total external reflection.

are recorded. The difference in intensity between patterns with different thresholds therefore yields the contributions to the pattern from within the energy range bracketed by the two thresholds (Figure 5.23).

5.8.1.1 *Worked Example – Real-Time Observations of the Control of Flow of Ions Across Cell Membranes*

Ion channels are pore-forming proteins that induce and control a small voltage gradient across the plasma membrane of cells, and are used to regulate chemical processes by switching on and off the permeation of ions in and out of the cell, a process referred to as 'gating'. The manner and timescale in which this electrochemical control is achieved is still far from being fully understood.

The functionality of a particular potassium-ion channel was investigated using Laue diffraction, or, more precisely, so-called 'diffracted x-ray tracking' (DXT) [14]. The experimental setup is shown in Figure 5.25. The ion-channel molecules (10 nm long) were deposited on a thin quartz plate coated with a surface modifier, which ensured that the transmembrane (TM) domain half of the channels was attached to the plate, while the

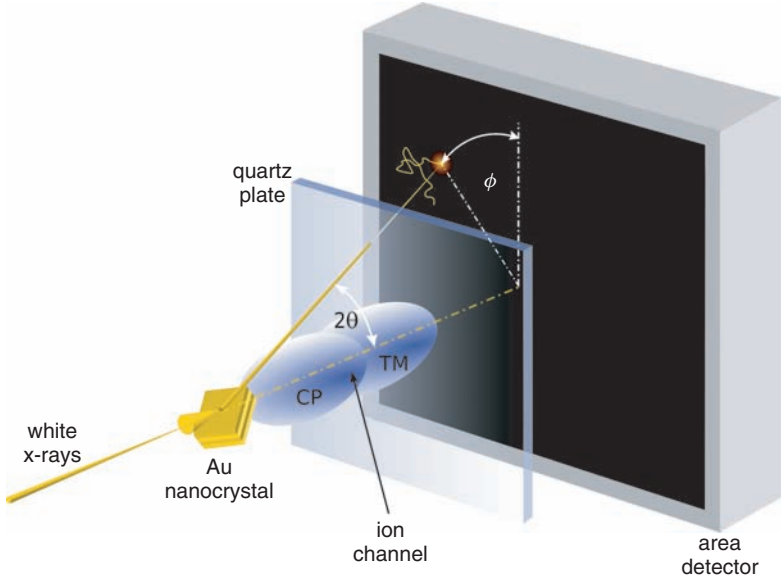

Figure 5.25 *Diffracted x-ray tracing (DXT) of ion-channel movement. An ensemble of potassium-containing ion channels was deposited on a coated quartz plate so that the cytoplasmic (CP) parts were all pointing away from the quartz surface. On to these were attached gold nanoparticles exhibiting one-dimensional periodicity. Each nanoparticle therefore acted like a tiny mirror, reflecting only that wavelength from the polychromatic x-ray spectrum which satisfied the Bragg condition for the angle θ it subtended with the pink beam. Movements of the 'reflected' Bragg spots on the area detector (dotted line) could be directly translated into angular movements of the ion channel. Adapted from [14] with permission of Elsevier.*

cytoplasmic (CP) part was pointing outwards.[7] The ion-channels were then 'tagged' with gold nanocrystals. These were platelike structures, with typical thicknesses of less than 10 nm and lateral dimensions of approximately 25 nm. They diffract x-rays as if they were tiny mirrors, due to the fact that they have the curious property of being one-dimensional periodic structures, the periodicity being perpendicular to the plate surface. Hence, any rotation of a given nanocrystal, defined by the 'polar' angle θ and 'azimuthal' angle ϕ (see Figure 5.25) are directly translated on the area detector to angular shifts of the diffraction spot of 2θ and ϕ, respectively.

The ion channel remained 'closed' under neutral (pH 7.5) conditions. Several diffraction spots from the nanocrystals were monitored and showed only low-amplitude oscillations, resulting in traces on the area detector showing predominantly radial motion, that is, small variations in θ (see Figure 5.26(a)), caused by Brownian fluctuations and a resultant 'flexing' or 'bending' of the ion channels. Upon changing the local environment to a pH-value of 4.0 (acidic), the trajectories of the diffraction spots changed dramatically (see Figure 5.26(b) and 5.26(c)). The azimuthal motions (changes in ϕ), which were previously insignificant, dominated the traces, indicating a twisting of the ion channel. Both clockwise and anticlockwise motions were observed, reflecting random opening and closing motions. Although the twisting motion now dominated, larger radial shifts than those observed in the quiescent (neutral) state were also recorded, indicating that gating is associated with both a twisting and concomitant bending motion.

In this manner, a mechanical interplay between the cytoplasmic and transmembrane domains was suggested, which has provided important insights into the mechanistic features of conformational changes during ion-channel gating. A word of caution is required, however: the mass of the gold nanocrystals is significantly larger than the mass of the ion channels to which they are attached, which will affect the mechanics and dynamics of the system in an unknown but probably significant manner.

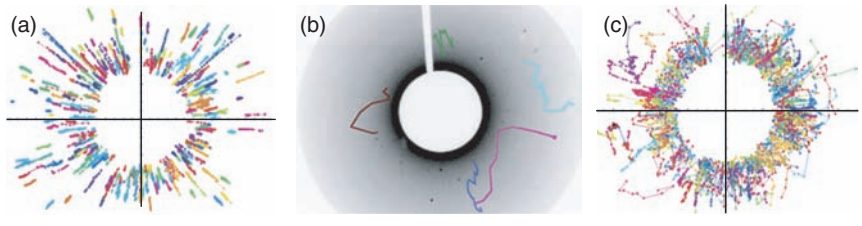

Figure 5.26 *DXT results of ion channel movement. (a) In the neutral state, the ion channels flex backward and forward due to Brownian fluctuations, resulting in radial traces on the area detector. (b) and (c) Once the environment becomes acidic, the traces of the diffraction spots on the detector become significantly larger and involve both circumferential (azimuthal) and radial motions, implying a twisting motion, coupled with enhanced bending. Adapted from [14] with permission of Elsevier.*

[7] The cytoplasmic part of an ion channel resides within the cell membrane, while the transmembrane part extends outside the cell walls.

5.8.2 Single Crystal Diffraction With Monochromatic X-rays

5.8.2.1 The Rotation Method

In the large majority of single-crystal diffraction experiments, monochromatic radiation is used and the so-called 'rotation' or 'oscillation' method is applied. This is shown schematically in Figure 5.27.

Data collation runs as follows. Firstly the crystal symmetry, the unit cell parameters, the crystal orientation and the resolution limit must be ascertained. Armed with this information, a data collection strategy which will maximize both the resolution and

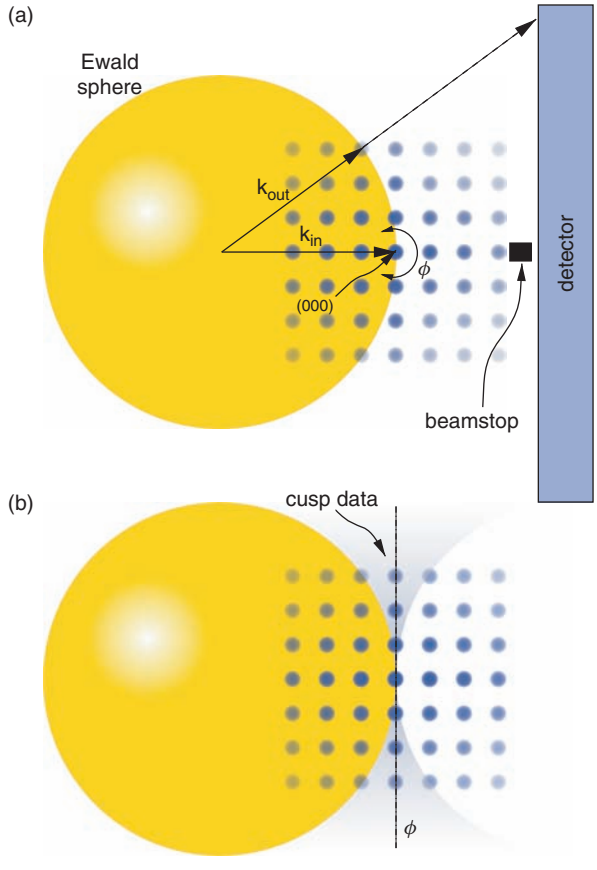

Figure 5.27 *Schematic figure showing data collection in the oscillation technique. (a) By rotating the crystal around an axis perpendicular to the incident beam (ϕ), diffraction maxima pass through the surface of the Ewald sphere and are registered on a 2-D x-ray detector. Refer also to Figure 5.10. The highly intense (000) reflection is prevented from damaging the area detector by the use of a beam stop. (b) When viewed from above the plane containing the ϕ-axis, one sees that for a given crystal orientation relative to the axis, some data cannot be accessed (known as 'cusp' data and shaded blue here) as it never passes through the surface of the Ewald sphere. However, by rotating the crystal (typically by 90°), this data can also be recorded.*

completeness of the data set is derived. The method used is to rotate the crystal through a small angle, typically 1° for protein crystallography, while simultaneously recording the x-ray diffraction pattern.

There are several factors which go towards determining the angular range $\Delta\phi$ over which a detector image is integrated. If the diffraction pattern is very crowded, then the rotation angle should be reduced so that each spot can be resolved on the image (within $\Delta\phi$ it is not known when a given diffraction maximum passed through the Ewald sphere). This process must be repeated until the crystal has moved through at least 30° and sometimes as much as 180°, depending on the crystal symmetry. The lower the symmetry, the more data are required.

Because traditional detectors such as image plates or CCDs require significant readout times, data acquisition is 'stop-and-go' – it consists of ramping up the rotation velocity of ϕ, opening an x-ray shutter, gating the detector over $\Delta\phi$, closing the shutter again, then ramping down the ϕ-motor while reading out the image (hence the term 'oscillation method'). The angular position of the motor must then be reset so that the next 'slice' of the ϕ-range can be recorded (see Figure 5.28(a)). Recording too large a range $\Delta\phi$ not only causes overlapping, but also results in larger background x-ray signal, which will degrade the signal-to-noise ratio. On the other hand, making $\Delta\phi$ too small increases the total 'dead time' resulting from reading out the images. Note also that the position of the maximum of a peak can lie anywhere within the integrating range.

Recent advances in single-photon detector technology have made possible a novel operational mode referred to as 'fine phi-slicing' (Figure 5.28(b)). In this mode, the ϕ-motor is allowed to operate continuously and no x-ray shutter is used. Images are continuously recorded over small $\Delta\phi$ ranges. This is made possible by the very short readout times, typically measured in a few milliseconds, compared to exposure times of the order of 0.1 to 1 s, corresponding to $\Delta\phi$-values of the order of 0.1°.

This mode is essentially excluded from systems using CCDs for two main reasons. First, CCDs generate noise during readout, hence short exposure times (small $\Delta\phi$-values) become unacceptably noisy. Also, because the readout time of CCDs is long, an x-ray shutter must be used. For short shutter opening times, 'jitter' in the opening and closing mechanism introduces additional inaccuracies.

Importantly, in the fine phi-slicing mode not only are the integrated Bragg-peak intensities recorded, but also their profiles. This not only more accurately determines the position of the diffraction peak, but also yields information on properties such as strain and domain size.

In many cases, it is necessary to record a complete data set in order to determine the crystal structure. However, from Figure 5.27, it is apparent that those reciprocal lattice points (i.e. Bragg peaks) that lie very close to the ϕ rotation axis never cross the Ewald sphere and therefore are never recorded. To acquire these so-called 'cusp', or 'blind' data, two methods can be applied. For crystals with low symmetry, for example triclinic systems in which the only symmetry partners (i.e. equivalent reflections with the same intensities) are the Friedel pairs (hkl) and $(\overline{hkl})$, if one Bragg peak lies outside the cusp volume so must its partner. In this case, a second set of data must be recorded for which the crystal is rotated around a second axis, normally orthogonal to the first. For systems exhibiting higher symmetry, one can offset the ϕ rotation axis from a high-symmetry axis, so that although there will be reflections that still lie permanently in the cusp

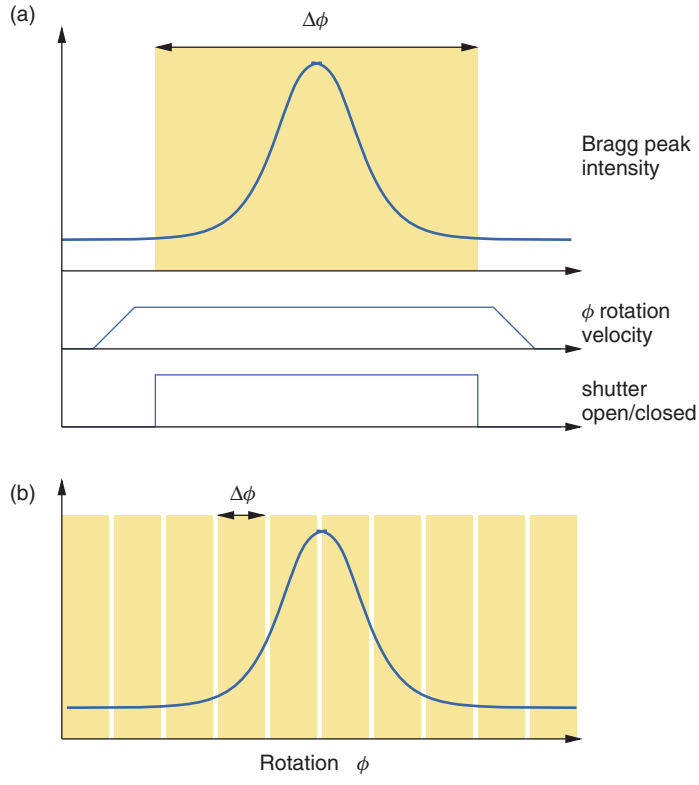

Figure 5.28 *Recording single-crystal diffraction data. (a) In the 'oscillation' method, signal is integrated over an angle $\Delta\phi$, chosen to be large enough to include entire Bragg peaks, but not so large so as to make the recorded image overcrowded. After the exposure, the image is read out, which typically takes a second or more. This process is repeated until the entire required ϕ-range is covered. (b) In the fine phi-slicing mode using modern photon-counting detectors, the readout time is insignificant compared to the typical times the ϕ-motor needs to move over a small $\Delta\phi$, in which case no x-ray shutter is needed, the ϕ-motor runs constantly, and data acquisition is quasi-continuous, with only ms-dead times between exposures.*

volume, there will be one or more symmetry (Bijvoet) partners that will pass through the Ewald sphere.

5.8.2.2 The Selected Bragg Peak Method

In relatively rare cases, one is interested in recording only a subset of the diffraction data. An example might be orbital ordering of electronic states which increases the crystal periodicity and thereby introduces additional (and usually very weak) Bragg peaks at nominally non-integral positions in reciprocal space. In such cases, the signal of interest might be swamped by signal from the more intense 'main' Bragg peaks, and one therefore records across individual Bragg peaks by the use of a crystal goniometer (such as a so-called 'kappa goniometer stage') and a point detector (or equivalently, a small-area detector), as shown schematically in Figure 5.29.

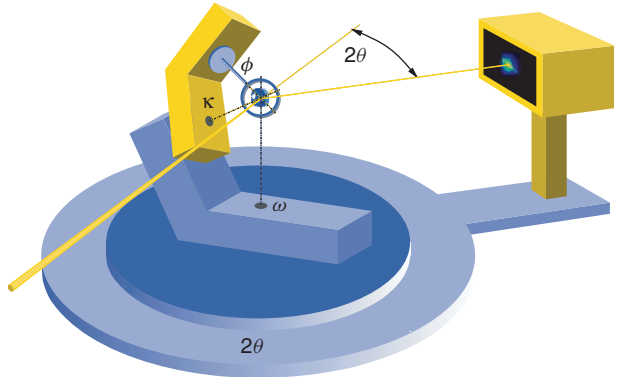

Figure 5.29 *A kappa goniometer stage. The detector angle 2θ moves only in one plane (here, the horizontal plane), hence the scattering vector associated with a given Bragg peak hkl must also lie in this plane. To achieve this, the diffracting planes (hkl) must be vertical (by rotating the sample around κ and ϕ) and subtend an angle θ with the incident beam, by adjusting ω.*

5.9 Textured Samples

As we have already discussed, crystalline samples may on the one hand be polycrystalline, with small crystallites with random orientations, or on the other hand be single crystals with a well-defined set of crystal axes in the laboratory frame of reference. Textured samples are a half-way house between these two extremes – the crystallites do not all point exactly in the same direction, although there is a preferred direction (to a greater or lesser degree). There are many possible reasons why a preferred direction arises, for example due to strain, 'brushing' of crystallites by an external force or energy minimization of exposed surfaces, to name only three.

5.9.1 Worked Example – Microdiffraction of Ancient Textiles

In 1947, an Arab shepherd boy was throwing stones through the opening of a cave at Qumran, near the shores of the Dead Sea in Jordan. He heard something break and went to investigate. He discovered a large collection of parchment and papyrus scrolls inside earthenware jars, with Hebrew script, covering almost the entire contents of the Old Testament (see Figure 5.30). Carbon dating identified these now-famous Dead-Sea Scrolls as being the oldest discovered record of the Old Testament. They were written around or even before the time of Jesus Christ, about 2100 years ago.

It is believed that the scrolls were written by a religious sect called the Essenes, of which very little is known. Any related archaeological information is therefore considered to be important, as it might help piece together a more complete image about their lifestyle. In addition to the scrolls, several fragments of woven cloth were also found in the caves. Identification of the type and origin of the fibres should help provide information on the Essenes.

The traditional method to identify textiles is via optical or electron microscopy. This readily reveals the handedness (sense of rotation) of spun yarns, while the shape of

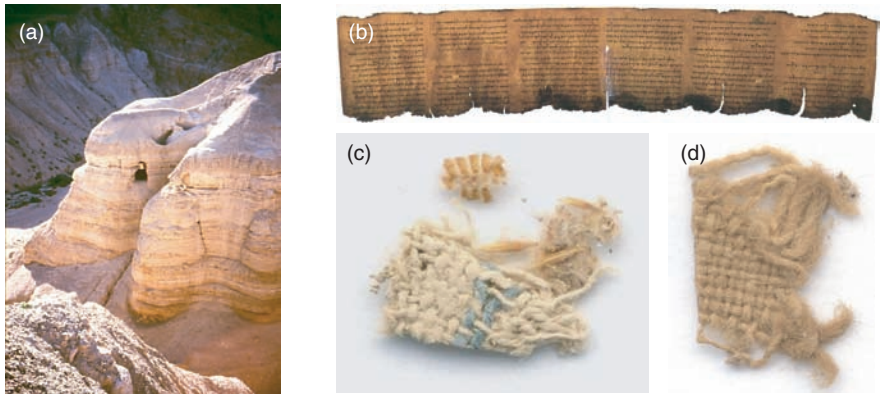

Figure 5.30 *The Dead Sea scrolls. (a) The Qumran caves on the shores of the Dead sea, (b) an example of the scrolls found in them, (c) and (d) linen and cotton samples taken from woven cloth found at the same site. Reprinted from [15] with permission of Elsevier.*

individual strands allows one to discriminate between animal and plant fibres. Although much information can be gleaned using these imaging techniques, open questions still remain, particularly concerning the plant fibres. X-ray diffraction can in principle provide the necessary complementary information. Early experiments on small fragments of the cloth using an x-ray beam focused down to 200 µm, however, yielded diffraction patterns which were dominated by signal originating from fine mineral particles of soil adhering to the fibres. These essentially swamped the weak and more diffuse signal generated by the cellulose nanocrystallites (so-called 'microfibrils') which have low scattering strength and typical diameters of only 4 to 7 nm, that go to make up individual fibres, which themselves have diameters ranging from some 10 to 50 µm and lengths of a few millimetres. What was therefore required was studies of individual fibres and a technique with a spatial resolution of a few microns or better.

The ID13 microbeam diffraction beamline at the ESRF provides tightly focused beams with diameters of a few microns and a brilliance of the order of 10^{18} photons $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ at a photon energy of 12.7 keV and a monochromator bandwidth of 2×10^{-4}. The tight focus was achieved by prefocusing with an ellipsoidal mirror, followed by secondary focusing using tapered capillaries. Single fibres were extracted from the textile fragments and mounted in an open frame so the fibre axis was perpendicular to the incoming beam (see Figure 5.31). Two-dimensional diffraction patterns were recorded on a CCD area detector (active area of approximately $132 \times 132\,\mathrm{mm}^2$) placed 58 mm away from the fibre [15]. This allowed a maximum 2θ-angle of approximately 49° and a maximum accessible scattering vector $Q_{\max} = 5.3\,\text{Å}^{-1}$.

Because the crystalline component of most of the fibres of interest here consists of cellulose microfibrils, azimuthally integrated diffraction patterns only differ subtly from one another. These small discrepancies are mainly in the width of the diffraction peaks, indicative of differences in the size (translating to the number of unit cells) of the cellulose microfibrils. However, the inherent texture of a fibre, given by how much the cellulose molecules align themselves to the longitudinal fibre axis, can be

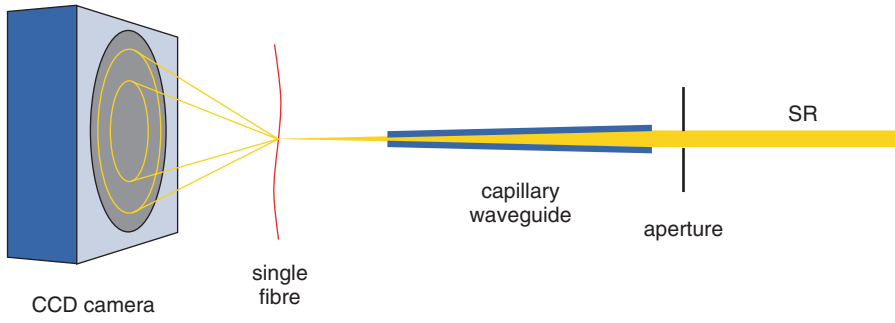

Figure 5.31 *A schematic diagram of the experimental setup used for the fibre diffraction experiments on the microbeam diffraction beamline at ID13, ESRF.*

used as a signature of the fibre type – in the case of bast fibres such as flax, hemp, and jute, the cellulose microfibrils are highly oriented along the fibre axis, due to the nutrient-conducting functionality of the sieve tubes common to these structures. This results in strong (200) diffraction maxima in the plane perpendicular to the fibre axis. In contrast, cotton is characterized by having a helical microfibril structure in the secondary wall of the cotton fibres, which is manifested by a 2-D diffraction pattern, in which four azimuthal maxima in a crossed configuration on the (200) Debye-Scherrer cone can be seen at about $\pm 30^\circ$. Hence, for example, two ancient textile samples could be identified as being bast (possibly linen from flax, or ramie) and cotton (see Figure 5.32) by comparison with diffraction patterns of modern versions of the same fibre types.

Because cotton was not introduced to this area of the world until much later than when the scrolls were written, dating studies were performed on these samples, which identified them as originating from the thirteenth or fourteenth century A.D. and therefore being of no significance to the history of the Essenes and the provenance of the scrolls.

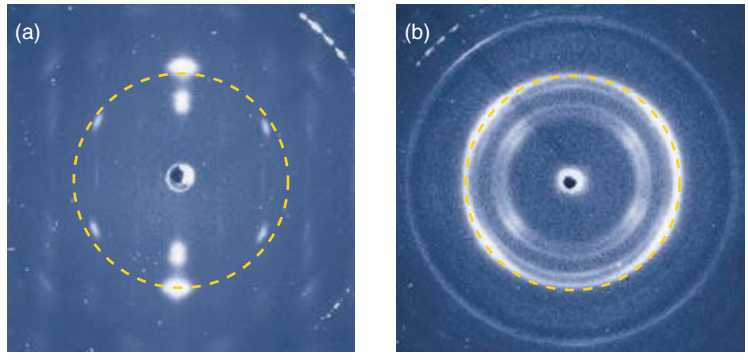

Figure 5.32 *Two-dimensional x-ray diffraction images of (a) an ancient bast fibre from the time of the Essenes and (b) a thirteenth- to fourteenth-century cotton fibre found at the same site. Note the difference in the azimuthal distribution of intensity, most marked around the (200) signal, indicated by dashed circles. Reprinted from [15] with permission of Elsevier.*

Other samples, perhaps only some 50 years or so younger than the bast textiles, were clearly made of wool – this is made from collagen and exhibits a much diffuser diffraction pattern, typical for nearly amorphous material. The textiles are far more sophisticated than those linen samples directly associated with the Dead Sea Scrolls – they are dyed in rich red and blue colours, and were left by Jews hiding in the caves from the besieging Romans [16].

In conclusion, x-ray microbeam investigations of single archaeological textile fibres have provided an unambiguous identification of the fibre type, where other methods failed, primarily because of the high degree of soil-encrustation and textile degradation.

5.10 Powder Diffraction

5.10.1 Introduction

For some materials, it may prove to be very difficult or indeed impossible to grow macroscopic crystals of sufficient crystallographic quality to be investigated using single-crystal diffraction. In such instances, powder diffraction (also called the 'Debye–Scherrer' method) is an invaluable technique, from which it is often possible to obtain an unambiguous unit-cell structure. Powder diffraction suffers from the fact that because the microcrystallites making up the powder sample are randomly oriented, the angularly separated information one obtains in single-crystal diffraction is projected on to a single coordinate, and distinct reflections can overlap.[8]

It should not be thought, however, that powder diffraction is a 'last resort' technique only to be used when large, high-quality, single crystals are unavailable. Powder diffraction not only provides a rapid and nondestructive means to identify the composite parts in multicomponent mixtures or complex systems in fields as disparate as metallurgy, archaeology, pharmaceutical sciences, mineralogy, and condensed-matter physics, but is also indispensable in extreme environmental studies, where phase changes are studied as a function of temperature and/or pressure. In such cases, information can often only be extracted from powder samples which, due to the microcrystallites' large surface area-to-volume ratio, can accommodate phase changes under these severe conditions by relaxation and propagation of crystallographic faults far better than can larger single-crystal samples, which most often degrade unacceptably and in an anisotropic manner. In addition, powder diffraction lends itself to time-resolved studies, whereby the projection of the crystallographic information in to one dimension can be considered in this case to be positively beneficial.

The time-resolution of such studies has been extended into the submillisecond regime since the advent of novel detector types such as the Mythen microstrip detector (see Section 4.6.6). For example, the dynamics of self-propagating exothermic reactions in nanoscale metallic multilayers were followed from powder diffraction patterns over a 120° range with 0.004° resolution and with a complete-pattern frame-rate of 8000 Hz [17].

[8] As a trivial example, the signals for the {430} and {500} reflections in a simple cubic system overlap, as they both have interplanar spacings of $a/5$. See Equations (5.4) to (5.6).

Lastly, it should be mentioned that the most widespread use of powder diffraction is in the identification of crystalline samples, whose patterns can be used as 'fingerprints' and compared with the enormous reserve of data of known materials to be found in the 'International Centre for Diffraction Data' (ICDD) database [18]. The use of synchrotron light for such routine experimental checks is, however, frowned upon.

5.10.2 Basics of Powder Diffraction

Consider Figure 5.33. A sample is irradiated with quasimonochromatic x-rays of energy $E = \hbar ck$. A detector is placed at an angle 2θ to the incoming x-ray beam. It will detect a strong diffracted signal only if the diffraction condition is met, that is, there must be crystal planes with Miller indices (hkl) at an angle $\theta = \arcsin(\lambda/2d_{hkl})$ to the incoming beam.

Imagine now that the sample consists of a large number of randomly oriented crystallites with an interplanar spacing d_{hkl}. The detector is positioned at an angle 2θ that satisfies the Bragg law for the x-ray wavelength λ and interplanar spacing d_{hkl}. The subset of the crystallites with their (hkl) planes at an angle $\theta \pm \delta\theta$ to the incoming beam and the normal to these planes in the plane defined by the incident beam and the detector will satisfy the diffraction condition (see Figure 5.33). The magnitude of $\delta\theta$ depends on the one hand on the beam parallelism and monochromacity, and on the other on the size and crystallographic perfection of the crystallites.

Due to the cylindrical symmetry of the setup about the incident beam axis (i.e. rotation about χ), cones of diffracted signal, called Debye–Scherrer cones, are produced (Figure 5.34). A scan is recorded by sweeping a detector in a plane that bisects these Debye–Scherrer cones about an axis situated at the sample. Hence, in powder diffraction, in contrast to single-crystal diffraction, the information is all contained in one degree of freedom. Lastly, it should be clear from Figure 5.33 that no valuable information is obtained by performing rocking curves on powder samples, whereby the detector angle 2θ is held constant and the sample is rotated about its axis (out of the paper in Figure 5.33), apart from establishing that the crystallites are indeed isotropically oriented in space. On the other hand, it is highly advantageous to spin powder samples that have

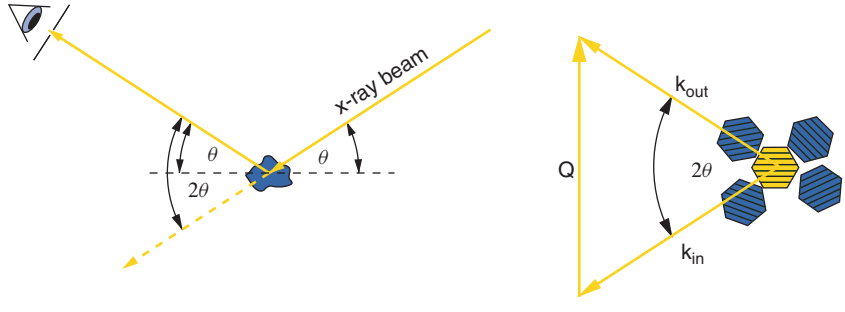

Figure 5.33 *Conditions for diffraction in a powder sample. A detector will only see a diffracted signal if the d_{hkl} spacing, the orientation of the crystallite, and the angle of the detector 2θ to the incident x-ray beam lead to the diffraction condition being satisfied. This is fulfilled by the yellow-highlighted crystallite.*

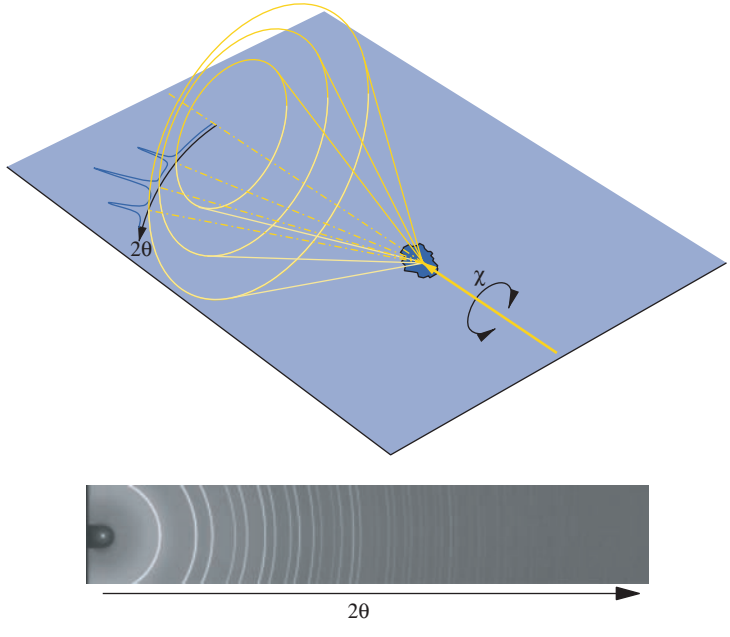

Figure 5.34 *A schematic of a powder diffraction experiment. Those crystallites with crystal planes (hkl) at an angle θ given by the Bragg law to the incoming beam will strongly diffract. The cylindrical symmetry of the experimental setup about the incident beam axis χ means cones of diffracted signal are produced. A diffraction pattern is obtained by scanning radially out from the beam axis with a detector in a plane that contains that axis. In the early days of powder diffraction, the signal was recorded on photographic paper, shown at the bottom of the figure.*

relatively few crystallites, in order to spatially average out any resulting 'graininess' in the signal. The shorter the acquisition time, the faster should be the spinning rate.

5.10.3 Worked Example – Structural Solutions Made Easy

The so-called pentacil zeolite structure ZSM-5 was patented in 1975 by the Mobil Oil Company. It is widely used in the petroleum industry, especially as a catalyst for hydrocarbon isomerization (the conversion of one hydrocarbon to another with the same chemical formula but different physical structure). The structure of ZSM-5 contains open channels of approximately 5.5 Å diameter, which can act as sensitive size-selectors. Hence, one isomer is able to diffuse through the structure much more rapidly than another, depending on the isomer shapes. The unit cell of ZSM-5 has a volume of approximately 5400 Å, and contains 96 silicon or aluminium atoms and 192 oxygen atoms.

ZSM-5 was used as a testbed, on the one hand for the efficiency of the phase-retrieval algorithm called 'charge flipping' [19], and on the other for the suitability of the Mythen microstrip detector (See Section 4.6.6) for structural refinement of complex crystalline structures.

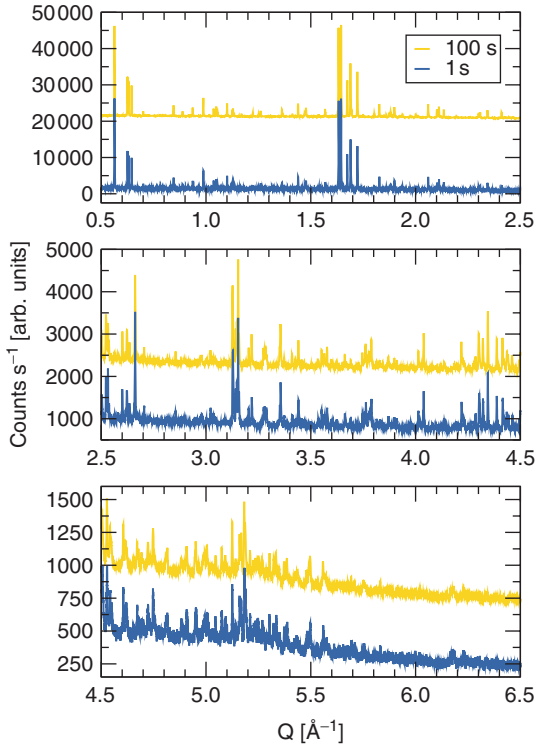

Figure 5.35 *Powder patterns of the zeolite structure recorded with the Mythen microstrip detector with 1 s and 100 s exposures. The 100 s exposure has been shifted vertically for the sake of clarity. The largest scattering vector of 6.5 Å^{-1} corresponds to a 2θ angle of approximately 80°. Courtesy Philip Pattison and Marc Schiltz, Ecole Polytechnique Fédérale de Lausanne, Switzerland.*

A powder sample in a capillary was illuminated with 1.24 Å synchrotron radiation for 1 and 100 s (see Figure 5.35). The maximum scattering vector of 6.5 Å^{-1} provided a spatial resolution of a little under 1 Å. After sensitivity and geometrical corrections were applied, the data was then fed into the charge-flipping algorithm, summarized in Figure 5.36(a). The essence of charge flipping is in the corrections made in the real-space part of each iterative cycle of this direct method. First, a threshold level for the electron density (δ in Figure 5.36) is selected. This should be somewhat lower than the lowest electron density expected to emerge from the structure (in the case of ZSM-5, this would be the electron density of an oxygen atom). The real-space output after performing an inverse Fourier transform (IFT) using the set of phases from the last iteration and the experimentally determined structure factor magnitudes (see Section 5.6.5) is then modified so that any electron density that is below δ has its sign reversed (or 'flipped'). Note that both small positive *and* negative electron densities undergo this change, not only the originally negative values. The iterative algorithm is then repeated until self-consistency is achieved in reciprocal and real space.

Using charge flipping, an electron density emerged, which corresponded precisely to the known unit-cell structure of ZSM-5, shown in Figure 5.36(b). It was thereby

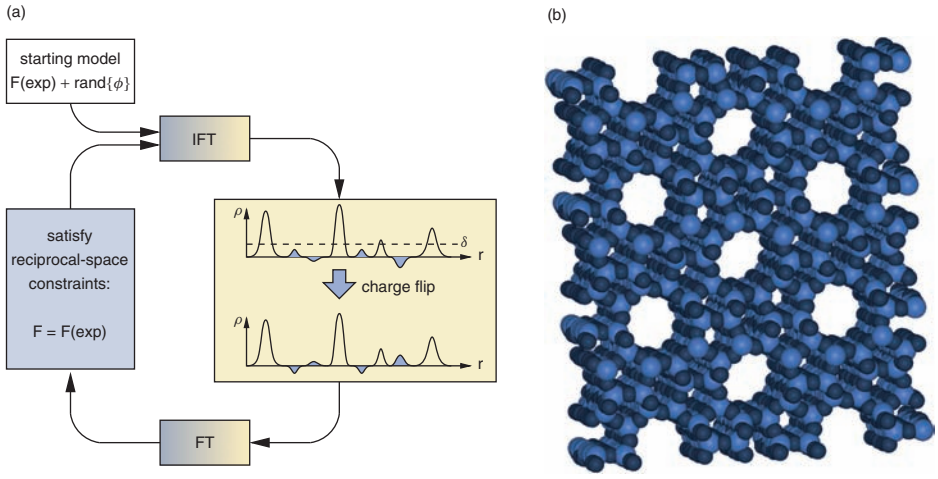

Figure 5.36 *Structural determination of ZSM-5. (a) The charge-flipping phase-retrieval algorithm used to analyse the powder-diffraction data. (b) The unit cell of ZSM-5, which was successfully generated using the charge-flipping algorithm on the powder pattern after the 1 s exposure shown in Figure 5.35.*

demonstrated that the atomic structure of a unit cell containing nearly 300 atoms could be solved using no a priori knowledge and with a one-second exposure. A decade ago, this achievement was unthinkable.

5.10.4 The Pair-Distribution Function

Important structural information can be directly extracted from powder diffraction data under certain conditions. As we have argued above, a one-dimensional powder pattern is essentially the orientation average of the 3-D pattern of a single crystal. In Section 5.6.2, it was argued that the direct Fourier transform of the diffraction intensity distribution of a single crystal delivers a 3-D map of the relative vectorial separations of the atoms within the unit cell. Similarly, through a specialized Fourier transform that takes into account the orientation averaging in powder samples, we can directly Fourier transform a powder pattern to obtain a density profile of the interatomic distances – this is the so-called 'pair-distribution function', or PDF [20]. In order to obtain useful information from a PDF, the resolution in real space must be good enough to clearly distinguish different atomic distances as separate maxima. This requires accurate measurement of the powder pattern up to as high scattering-vector value $Q = 4\pi \sin\theta/\lambda$ as possible and is therefore facilitated by using small wavelengths (0.05 to 0.5 Å) and sufficiently high counting rates at large Q-values, where the x-ray scattering efficiency is strongly suppressed. Generating a high-quality PDF is thus very photon hungry, making it almost exclusively a synchrotron-based technique.

An example of a PDF of silicon powder is shown in Figure 5.37. The maximum 2θ-value was 120° and the pattern was recorded with 25.3 keV photons ($\lambda = 0.49$ Å). The maximum scattering vector was therefore 22.2 $Å^{-1}$, which provides a spatial resolution of 0.283 Å.

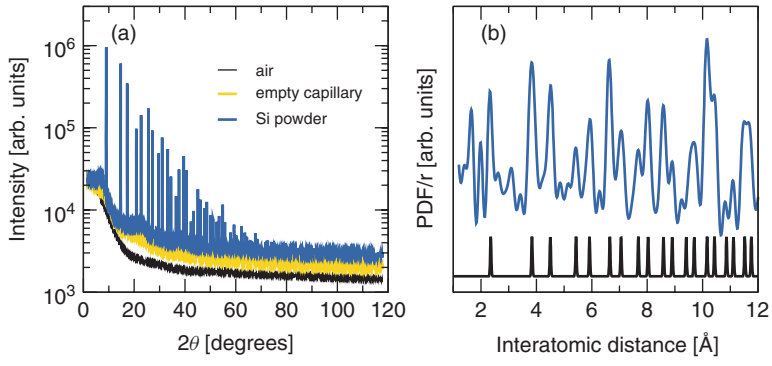

Figure 5.37 *The pair-distribution function of silicon powder. (a) A powder pattern of silicon, plus the patterns formed by air and an empty capillary of the sort used to hold the Si-powder, recorded at 25.3 keV. After subtraction of the two background signals, the PDF was calculated and is shown in (b), weighted by 1/r, whereby r is the interatomic distance. This can be compared to the theoretical interatomic distances for the known structure of silicon. Note the excellent positional agreement. Courtesy Antonio Cervellino, Paul Scherrer Institut.*

5.11 Protein Crystallography

5.11.1 Introduction

Protein crystallography (PX) is a form of single-crystal diffraction using the oscillation method. What, therefore, is the justification for discussing this separately from Section 5.8?

Firstly, the economic drive to synthesize new and increasingly effective drugs means that protein-crystallography beamlines are the most ubiquitous in third-generation synchrotron facilities. Despite this and the burgeoning number of facilities being built, there remains a high overbooking factor for PX experiments. Although complementary methods such as Fourier-transform infrared spectroscopy and nuclear magnetic resonance can provide indispensable information, particularly regarding the presence and local environment of functional groups of low-Z atoms, only PX can provide a comprehensive overall picture [21]. The importance of resolving the molecular structure in the understanding of a protein's function is self-evident. Protein crystallography is the foremost method of choice when studying protein or enzyme structures – indeed over 85% of all the presently known protein structures have been solved using PX.

The crystallization of proteins and protein-like substances has been a science (some call it an art!) for over a century. Van Deen (1864) believed that all naturally occurring organic substances can be crystallized when manipulated effectively. Schimper reported on the first active crystallization attempt by Maschke in 1859, who evaporated the solution from a preparation of Brazil nuts in order to obtain crystals.

The pace of discovery in molecular biology has been exceptionally rapid in the last few decades and has greatly enriched our understanding of the molecular basis of life. The sequencing of the human genome; the construction, substitution, and cloning of

new combinations of genes; the elucidation of DNA transcription; and the unravelling of the structure and functionality of the ribosome are just some of the highlights of recent research.

This progress has been made possible by the explosion in computer technology, improvements in x-ray equipment, and an increasing expertize in obtaining pure protein crystals in macroscopic quantities using recombinant DNA techniques.[9] As a result, the structures of many biologically significant proteins have been determined. The main stumbling block, as ever, is the phase problem, covered earlier in this chapter, although in PX this problem is inordinately larger than for small unit-cell structures and additional methods must be brought to bear.

Another problem is that preparation of protein crystals is a complex procedure. Samples of sufficient crystalline quality may be very small, containing only micrograms of material or less. Most crystals contain between 10^{10} and 10^{15} molecules, each unit cell containing one or more molecules, with lattice constants typically between 50 and 500 Å. Linear crystal dimensions span less than a micron to several tens of microns – we are dealing with exceedingly small amounts of material!

Organic crystals tend to be considerably less robust than inorganic substances – the single C–C bond and the C–OH bond have dissociation energies of approximately 300 and 380 kJ mol^{-1}, respectively, which should be compared, for example, to the Ti-O bond in crystalline anatase of over 670 kJ mol^{-1}. There is therefore the danger that the protein crystal under investigation will suffer radiation damage as a result of exposure to the x-ray beam – radicals are produced by photoabsorption and rapid dissociation, which can then migrate in the crystal lattice and wreak havoc to the local structure.

There are several ways to minimize this problem. Firstly, the crystal is normally cooled to about 100 K using a flow of nitrogen from a so-called 'cryojet', thereby kinetically hindering the diffusion of the photoinduced radicals. Secondly, it is possible to scan the x-ray beam across the crystal so that fresh parts are constantly being exposed. The rate of scanning depends on the sensitivity of the crystal to radiation damage. Such scanning procedures obviously require that the x-ray beam is considerably narrower than the sample size. This sets a premium not only on growing as large as possible high-quality crystals, but also on focusing the x-ray beam to micron-sized dimensions (with the attendant increase in flux and rate of radiation damage), and recording diffraction data as quickly and efficiently as possible. This is best achieved at undulator beamlines, exploiting their naturally low beam divergence.

5.11.2 Geometry and Resolution

The highest resolution $a_{\min}$ with which the electron-density distribution of a crystal unit cell can be mapped is determined by the largest scattering vector $Q_{\max}$ that can be recorded. Remember from Equation (5.12)

$$Q = \frac{4\pi}{\lambda} \sin\theta.$$

[9] A *recombinant* system is one which has been artificially engineered by combining DNA or RNA sequences that do not naturally occur.

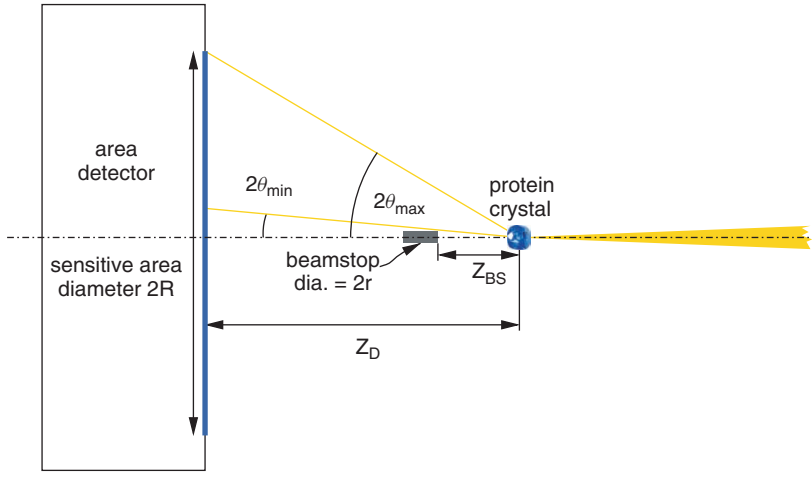

Figure 5.38 *The maximum subtended angle on a circular area detector is $2\theta_{max}$, while a beamstop used to protect the detector from the direct beam occludes diffraction maxima at angles less than $2\theta_{min}$.*

Consider the geometry shown in Figure 5.38. Here, therefore

$$Q_{\max} = \frac{2\pi}{a_{\min}} = \frac{4\pi}{\lambda} \sin\theta_{\max}, \tag{5.21}$$

whereby

$$\theta_{\max} = \frac{\arctan(Z_D/R)}{2}. \tag{5.22}$$

The smallest features which can be resolved therefore have a periodicity, or characteristic length, of

$$a_{\min} = \frac{\lambda}{2\sin\theta_{\max}}. \tag{5.23}$$

So, for example, data from an experimental setup in which a detector with a radius of 200 mm is at a distance of 300 mm from the crystal, which is illuminated with x-rays of energy 12.66 keV, can provide electron-density maps with a resolution of 1.69 Å. Features of the electron density with characteristic lengths smaller than this cannot be resolved.

Similarly, the *largest* features which can be resolved, $a_{\max}$, are given by the low-resolution limit, determined by the smallest scattering vector $Q_{\min}$, which is limited by the shadow of the beamstop in the detector plane (Figure 5.38). Using the same arguments as above for the high-resolution limit, we can calculate that the largest features that can be resolved for a beamstop of diameter 200 µm, placed 100 mm downstream from the sample, will be approximately 1000 Å, larger than all but the biggest unit-cell linear dimensions of protein crystals.

Because the lattice constants in protein crystals can be up to several hundred Angstroms, the angular separation between diffraction spots can be very small, of

the order of a few mrad. This sets stringent upper limits on the beam divergence at PX beamlines in order to avoid signal overlap and is another reason why undulator radiation, with its intrinsic high brilliance and low divergence, is indispensable.

5.11.3 Solving the Phase Problem in PX

The first step in obtaining the unit-cell structure in all crystallography is indexing the diffraction pattern. In indexing, the dimensions and symmetry of the unit cell are identified. There are 243 possible combinations of unit-cell type and internal symmetry. Fortunately, the very large majority of protein crystals are chiral[10] which excludes all but 65 of these.

In order to visualize the protein structure we need to solve the phase problem. We can achieve this using one or more of several different approaches.

If the coordinates of a similar protein are already known, we can try to solve the structure using a process called molecular replacement, which involves taking this model and rotating and translating it into the new crystal system until a good match to the experimental data is obtained. If this is successful, then we can calculate the amplitudes and phases from this solution, which can then be combined with our data to produce an electron-density map.

However, modern phase-solving techniques in protein crystallography are normally based on changing the structure factors of specific atoms within the unit cell. According to how the structure factor is perturbed, deductions can be made about the values of the phases. Changes in the structure factor can be induced by so-called multiple isomorphous replacement (MIR), or by multiwavelength anomalous dispersion (MAD), which are now described briefly.

5.11.3.1 Multiple Isomorphous Replacement

Isomorphous replacement was first used for the class of minerals known as alums as long ago as 1927 by J. M. Cork, a postdoctoral fellow of W. L. Bragg [22]. Its first successful use in solving the structure of macromolecular structures was famously performed by Max Perutz in his pioneering work in determining the structure of haemoglobin [23]. In phase-perturbing techniques such as MIR and MAD, one takes advantage of the enhanced scattering of heavy atoms, as they tend to contribute disproportionately to the overall scattering intensity. The reason for this should be clear – the phases of the individual scattering vectors of the lighter atoms tend statistically to cancel each other out, as their separations are of the order of or larger than the x-ray wavelength. However, scattering from the electrons of a heavy atom all scatter more or less in phase, as they are concentrated around the atom nucleus. The scattering intensity is proportional to the square of the number of electrons, hence the contribution from heavy atoms will 'stick out'.

In Section 5.6.2, we noted that the intensity of a given maximum in a Patterson map is proportional to $Z_A Z_B$, the atomic numbers of the two atoms A and B responsible for the peak. Hence, one can relatively easily establish the position of heavy atoms in a unit

[10] A chiral structure is one which cannot be superposed on its mirror image. Examples include hands (indeed, the Greek for 'hand' is the etymological root for the term chirality), and the molecule C H Cl Br I (chloro-bromo-iodomethane).

cell from a Patterson map, and thereby their contribution to the total structure factor. This is especially true for unit cells which contain more than one heavy atom.

Multiple isomorphous replacement (MIR) exploits the relative importance to the scattering intensities of heavy atoms. In MIR, derivative forms of the protein crystal of interest are grown in which heavier atoms are additionally inserted, or replace one or more atoms within the unit cell. In order for the technique to succeed, however, the addition of the heavy atom(s) or their substitution for lighter atoms should not significantly perturb the structural form (hence the term 'isomorphous', which means 'same shape'). The success of this technique is rarely guaranteed a priori and is usually not known until more x-ray data have been collected.

Commonly used heavy atoms in MIR are Hg^{2+} ions, which bind to thiol (SH) groups; uranyl salts, which bind between carboxyl groups in aspartic acid and glutamic acid; Pb, which binds to cysteine; $PtCl_4^{2-}$, which binds to histidine; and Se, which replaces sulfur in methionine to produce selenomethionine.

One records a data set of the native protein crystal and one or more of its heavy-atom derivatives. Added scattering from the heavy atom(s) causes the intensities of the diffraction maxima (structure factors) to change. Let us consider two experiments, whereby a given diffraction maximum yields values $I_N = F_N^2$ and $I_{D1} = F_{D1}^2$ for the native and heavy-atom derivative, respectively. We therefore know the magnitude of the structure factors (F_N and F_{D1}) but, importantly, we also know from Patterson-map analysis the magnitude *and* phase of the contribution of the structure factors due to the heavy atom alone. We have labelled these F_{H1} and ϕ_{H1} in the so-called 'Harker construction', shown in Figure 5.39(a).

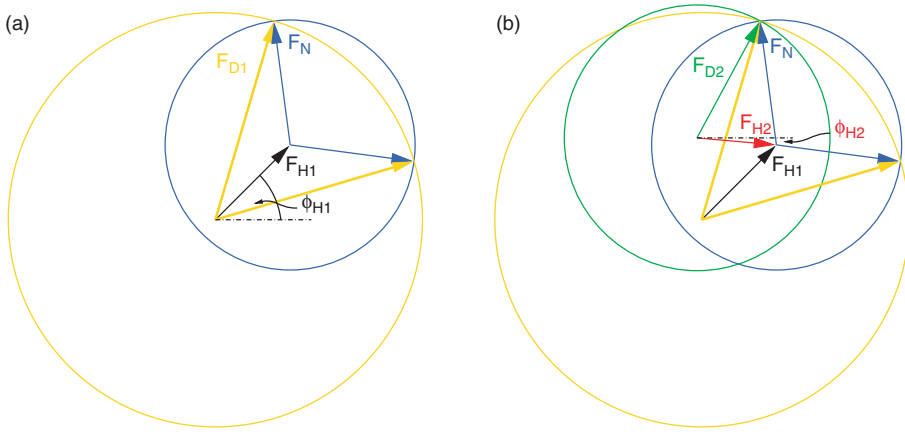

Figure 5.39 *The Harker construction. By substituting with or adding a heavy atom into a native protein crystal in a manner that does not significantly perturb the atomic structure, the phase problem can be resolved. (a) The magnitudes (though not the phases) of structure factors of the native crystal and a derivative, F_N and F_{D1}, respectively, can be experimentally determined. However, F_{D1} is the vector summation of F_N and F_{H1}. We can determine F_{H1} and ϕ_{H1} from the Patterson map of the diffraction pattern of the derivative crystal. Simple geometric considerations therefore limit the phase of F_N relative to F_{D1} to two possible values, given by the intersection of the two circles. (b) Repeating the experiment with a second derivative crystal D2 unambiguously determines the phase.*

As can be seen, these three experimentally determined facts mean that only the pairs of vectors meeting at the two points of intersection between the two circles can possibly be correct. One circle has a radius equal to F_N and the other a radius of F_{D1}. Their origins are separated by the vector representing F_{H1}. This ambiguity of two possible solutions is lifted by recording a third diffraction set from a second derivative with a heavy-atom structure factor F_{H2} and phase angle ϕ_{H2}, which will satisfy only one of the two solutions (Figure 5.39(b)).

5.11.3.2 Multiwavelength Anomalous Dispersion

Multiwavelength anomalous dispersion (MAD) is an elegant and often very effective method that relies entirely on the measurement of the differences produced by one or more anomalously scattering atoms in the crystal.

As we have discussed in Section 5.6.4, the form factor of an atom included in the crystal basis changes significantly near an absorption edge. This was shown in the Argand diagram of Figure 5.19. When the contribution of such an atom to the structure factor is large, this effect is significant, and can be modelled. To obtain detailed information of the electron-density distribution in the unit cell, spectra at different energies around the absorption edge are recorded. Just below the absorption edge, f_2 (also called f'') is very small and usually ignorable. As one crosses the edge of the anomalous scatterer, the absorption term f_2 increases rapidly, while f_1 dips as the scattering strength decreases around the edge. This causes the structure factor to 'jerk' to a new phase and intensity. It then settles more gradually again towards the nonanomalous value as one moves to still higher energies. Hence, because the anomalous scatterer has different atomic scattering factors at different wavelengths across the absorption edge, one can think of MAD as being the 'perfect' MIR experiment, with guaranteed perfect isomorphism.

In practice, several consecutive data sets are recorded from the same crystal at different wavelengths around the x-ray absorption edge of the anomalous scatterer. Typically, at least three measurements are made: at the inflection point of the absorption spectrum, where the dispersion is largest; at the maximum of the absorption peak; and again at energies a few hundred eV below and/or above the absorption edge (see Figure 5.40).

As it requires a tunable x-ray source, MAD can only be performed at a synchrotron. The resulting phase information can often produce very high-quality electron-density maps, thereby simplifying the subsequent interpretation. Selenium is a particularly good anomalous scatterer (its K-absorption edge is at 12.658 keV) and it can be incorporated into proteins by over-expressing them in strains of *E. coli* that are auxotrophic[11] for methionine. The host cells are then grown on minimal media supplemented with amino acids using selenomethionine instead of methionine.

Finally, it is mentioned that both MIR and MAD can be brought to bear on a given problem. In addition, two other methods, single-wavelength anomalous dispersion (SAD) and single isomorphous replacement with anomalous scattering (SIRAS) are being increasingly used. In SIRAS, one exploits the breakdown of Friedel's law close to an absorption edge of a single isomorphous derivative. In addition to the structure factors

[11] Auxotrophism is the need for external nutrients of an organism, such as bacteria, that has lost the ability to synthesize that substance required for its growth and metabolism as a result of mutational changes.

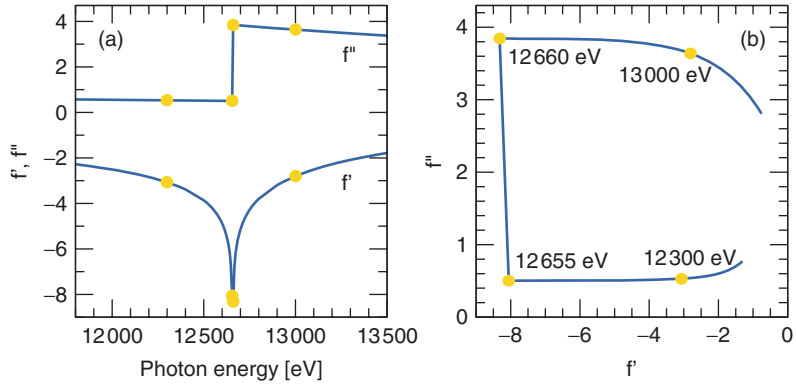

Figure 5.40 *Choice of energies in MAD. (a) The changes in f′ and f″ across the K-absorption edge of selenium, showing typical positions across the edge where MAD data are recorded. (b) The choice of energies is best understood when considering the plot of f″ as a function of f′. At the marked positions, the differences in the magnitudes of f′ and f″ are maximized, thereby providing the largest change in the structure factors.*

from the native crystal, the information obtained by recording the structure factors at $\mathbf{Q}$ and $-\mathbf{Q}$ of the derivative are enough to unambiguously resolve the phase.

5.11.4 Worked Example – Cracking the Rabies Virus Protection Shield

Louis Pasteur developed the first effective vaccine against rabies more than 100 years ago, by growing the virus in rabbits, and then weakening it by drying the affected nerve tissue. Nowadays, there are very efficient, though expensive, vaccines. However, in poorer countries where public health care is far less available, purchase and cooled storage of these vaccines is often unaffordable. Rabies in the Third World is still common among humans, accounting for over 50 000 deaths each year.

Rabies, like some other pernicious forms of viruses, including ebola and measles, is unusual in that in order for it to be able to infect a host cell, its ribonucleic acid (RNA) must first be transcribed into *messenger*-RNA, which then codes for the viral protein, allowing the virus to reproduce. For this first transcription step, the virus's RNA must be cocooned in a large protein-based closed shell, known as a *nucleocapsid*, in order to provide the appropriate chemical environment.

In a joint effort, groups from the Université Joseph Fourier-CNRS and the European Molecular Biology Laboratory in Grenoble were able to produce a recombinant system of the viral rabies RNA and its protein shell. The synthesized RNA molecule is very short, consisting of only 99 nucleotides (i.e. the G, C, U, and A RNA building blocks). The ends of this molecule attract one another chemically so that a circular structure is formed. This relatively compact molecule could then be successfully crystallized, and the orthorhombic structure ($a = 270$ Å, $b = 281$ Å, $c = 237$ Å) was recorded down to a resolution of 3.5 Å (in other words, the largest scattering vector had a magnitude $Q = 2\pi/3.5 = 1.795$ Å^{-1}). By isomorphous replacement, 2 rings of 11 gold atoms each were inserted into each complex, and a second set of x-ray diffraction data was recorded down to a resolution of 6 Å at the Au L-III edge at $\lambda = 1.0372$ Å.

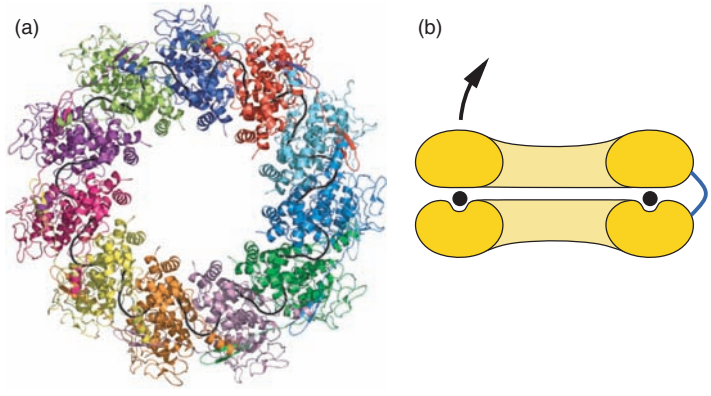

Figure 5.41 *(a) The structure of the nucleoprotein RNA complex of a rabies virus. The 11 tightly coiled protomers are highlighted in colour, while the RNA itself runs between these and is shown as a loosely coiled black strand. Adapted from [24] with permission of AAAS. (b) The nucleocapsid shell can be thought of as a protective cocoon for the RNA, which only opens under conditions favourable for replication.*

The overall structure of the complex was solved [24] and is shown in Figure 5.41(a). It was seen that the shell consists of 11 groups known as 'protomers', each group binding to 9 nucleotides. Crucially, the nucleocapsid forms two main domains above and below the RNA ring, which are connected by two threadlike structures which were seen to protrude outwards and act like hinges. The nucleoprotein shell therefore acts as a set of 'jaws' that totally close round the RNA molecule (see Figure 5.41(b)), protecting the latter from outside chemicals such as enzymes. Although the shell acts as a protective shield, it is able to distinguish between different types of enzymes trying to gain access to the RNA coil inside – those enzymes from the immune system of the host cell which attempt to degrade and destroy the foreign RNA are blocked, while in an environment where the RNA virus can replicate the jaw mechanism opens.

By identifying this protective mechanism, two possible approaches suggest themselves for neutralizing the rabies virus. The first would be to chemically lock the hinge of the jaws, so that replication is always inhibited. The other would be to force the jaw open at a time when the internal RNA can be attacked by the host's immune-system enzymes. Armed with this new information, targeted research into developing cheap, stable, and effective drugs which exploit these findings can begin.

5.11.5 Protein Powder Diffraction

Until recently, protein crystal structures were considered to be too large and complex for meaningful structural information to be extracted from powder samples, in which the myriad signals in reciprocal space (Equation (5.8)) are projected onto a single coordinate (2θ). On the other hand, many macromolecular structures have defied being produced as macroscopic crystals with linear dimensions of a few microns or more. Many of these do crystallize, however, but only on the micron- or submicron scale, and therefore might instead lend themselves to analysis by powder diffraction. As long ago as 1999, Von Dreele demonstrated that, at least for the modest 1261-atom metmyoglobin protein,

Rietveld refinement of the powder data in combination with physically reasonable stereochemical constraints produced a structure of comparable quality to that obtained using large single crystals [25, 26].

In conventional powder diffraction, a single data set might cover an angular range of 20° or more, with step sizes of less than 0.01°. Such scans can take several hours to complete – in the pioneering work by Von Dreele at the NSLS in Brookhaven, the scan from 1.5 to 20° took more than six hours – and the problem of radiation damage rears its ugly head. This can be reduced by a combination of flash-freezing using a cryojet and moving the capillary powder holder laterally in order to expose fresh parts of the sample (assuming there is enough material available). With the advent of one-dimensional microstrip detectors (Section 4.6.6), one can record entire data sets in such a short time that flash-freezing becomes unnecessary. Indeed, this may prove to be a positive advantage, as flash freezing can often destroy otherwise high-quality crystals due to the formation of ice particles on the nanoscale. An example of powder patterns recorded using conventional techniques and with the Mythen detector for the enzyme D-xylose isomerase (with a molecular weight of more than 300 kDa) is shown in Figure 5.42. For a given signal-to-noise ratio, recording entire data sets in parallel using Mythen is approximately four orders of magnitude faster than using a crystal analyser, with only a moderate reduction in resolution.

5.11.6 Time-Resolved Studies

Although the structures of proteins derived from their crystals provide invaluable information about their functionality, the purpose of proteins in nature is to induce chemical and structural changes. The mechanisms by which this is achieved are of great interest. Importantly, and in contrast to crystals of small inorganic and strongly bound systems,

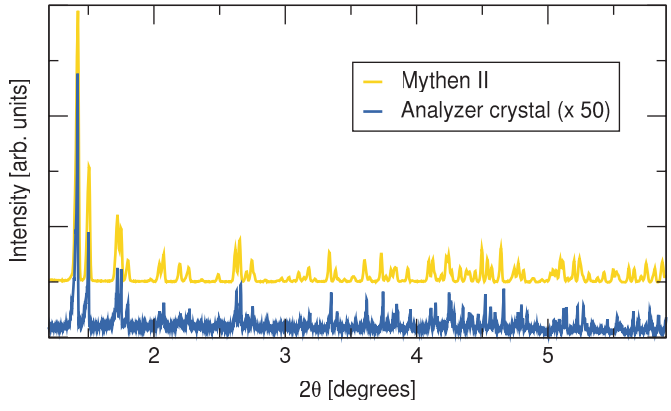

Figure 5.42 *Powder diffraction patterns of the enzyme D-xylose isomerase. At the cost of a slight loss in resolution, the data-acquisition time is reduced from several hours using a crystal analyser detector at the beamline ID31, ESRF, to only 30 s, when using the Mythen microstrip detector at the beamline X04SA at the SLS. Note also the large improvement in the signal-to-noise ratio. Courtesy Philip Pattison, Ecole Polytechnique Fédérale de Lausanne, Switzerland.*

macromolecular crystals contain significant amounts of bulk solvent, which allows them to reconfigure without forcing a significant change in the shape and size of the crystal lattice, which would otherwise exert such enormous stress on the crystal that it would be destroyed. In other words, the relative 'softness' of the structure within the unit cell of a protein crystal allows it to rearrange in a chemical or photochemical reaction without seriously affecting the degree of crystal perfection. This therefore makes time-resolved protein crystallography, in principle at least, a viable technique.

Structural intermediates in many biological reactions at physiological temperatures occur on timescales of much less than one second and therefore cannot be tracked by 'normal' PX, using quasimonochromatic light and the oscillation method described in Section 5.8.2. Instead, Laue diffraction using polychromatic light can be applied. It was thought for a long time that the energy overlapping problem in Laue diffraction would be a fatal flaw, until in 1984, Keith Moffat *et al*. published a landmark paper, in which it was demonstrated that this need not be an insurmountable problem [13].

To understand their argument, it is instructive to again consider the Ewald construction for Laue diffraction, shown in Figure 5.43. The polychromatic x-ray beam has a range

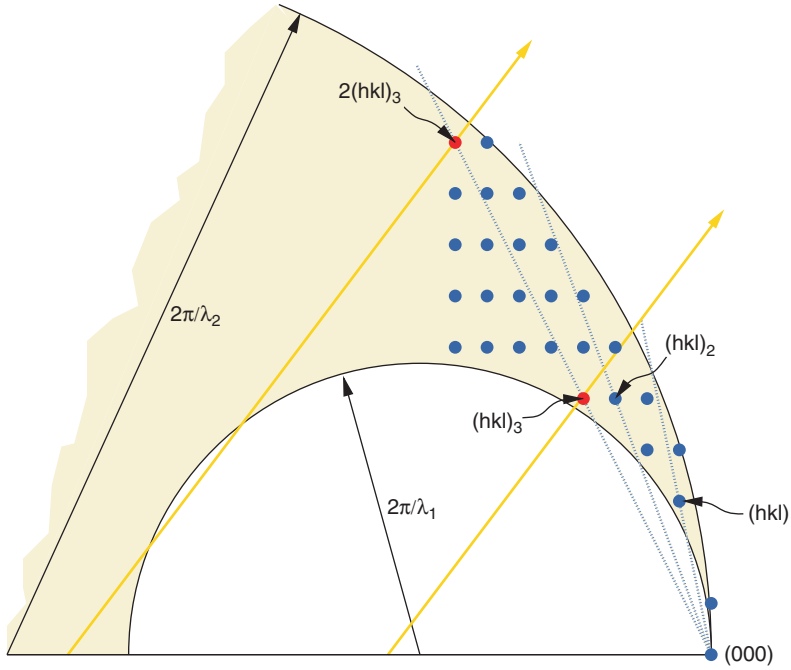

Figure 5.43 *An Ewald construction in Laue diffraction. A polychromatic beam of x-rays spanning an energy range E_1 to E_2, with corresponding wavelengths λ_1 and λ_2 illuminates a protein crystal. All the diffraction maxima lying in the reciprocal-space volume, highlighted in yellow in between the Ewald spheres of radius $2\pi/\lambda_1$ and $2\pi/\lambda_2$ and touching each other at the (000) direct beam maximum, will satisfy the Bragg condition. Three different diffraction maxima $(hkl)_1$, $(hkl)_2$, and $(hkl)_3$ are highlighted. Any other diffraction maxima which lie on the blue radial lines which join these to the (000) point are overtones which have the same diffraction angle and will therefore overlap, as shown for the two points shown in red.*

of energies between E_1 and E_2, with corresponding wavelengths λ_1 and λ_2 illuminating a protein crystal. Those reciprocal lattice points that lie in that part of reciprocal space between the two Ewald spheres with radii $2\pi/\lambda_1$ and $2\pi/\lambda_2$ will simultaneously produce diffraction maxima. It should be clear that higher-order reflections of a given reflection (*hkl*) lie on a straight line in reciprocal space that passes through both the (*hkl*) and (000) reciprocal lattice points. Depending on the sizes of the two limiting Ewald spheres, the higher-order reflections can lie outside the region accessed by the polychromatic beam (shown as the yellow shaded area in Figure 5.43) – the narrower the bandwidth of the x-rays, the less likely it will be that one or more higher-order peaks will be recorded, although this comes with a loss in the total number of diffraction maxima (structure factors) which can be simultaneously recorded. There is therefore a trade-off to consider – the wider the bandwidth, the more Laue reflections are recorded and the shorter the exposure time can be, although quantitative analysis becomes increasingly difficult due to the overlapping order problem. For narrower bandwidths, more complete data sets can be obtained by recording Laue patterns at different orientations of the crystal relative to the x-ray beam.

The range of wavelengths used for time-resolved Laue diffraction of protein crystals has a natural limit at the long-wavelength end, given by the absorption coefficient, which becomes unacceptably high above approximately $\lambda = 2.5$ Å (5 keV), though this of course also depends on the crystal dimensions.

Proteins which undergo reversible photochemical changes lend themselves particularly to time-resolved studies. Indeed, the first major breakthrough in this field was made by Moffat and co-workers in 1996, where they reported on the dynamics of the photolysis of the carbon-monoxide complex of myoglobin on a nanosecond timescale [27] using a pulsed laser and individual x-ray pulses from the electron bunches within the synchrotron storage ring. The temporal resolution was limited by the nanosecond laser pulses used to activate the reaction. The general setup is shown in Figure 5.44.

More recently, photocatalysed reactions using femtosecond laser pulses in a similar (but mutated) system provided 'movies' of the reaction dynamics on a 150 picosecond timescale, whereby now the limit to the temporal resolution was given by the x-ray pulse width in the single-bunch mode used for this experiment [28].

Further improvements in resolution down to the femtosecond timescale are eagerly anticipated with the advent of x-ray FELs, which will offer much higher flux and pulse lengths of the order of a few femtoseconds, similar to the shortest pulses that can be generated by modern lasers.

5.12 Ultrafast Diffraction Using Femtoslicing

The natural time-structure of third-generation synchrotron facilities is given by the duration of the electron bunches within the storage ring, which is of the order of 10 ps (see Chapter 3). XFELs promise the ability to probe timescales some three to four orders of magnitude smaller, but these facilities are still very rare and even more expensive.

To bridge this gap, the idea of 'femtoslicing' a relativistic electron beam with a femtosecond, high-intensity laser pulse was first proposed and demonstrated at the Advanced Light Source (ALS) in the mid- to late-1990s [29], and was implemented also at BESSY in Berlin, Germany, to generate soft x-ray pulses in the range of 1 to 2 keV (12–6 Å).

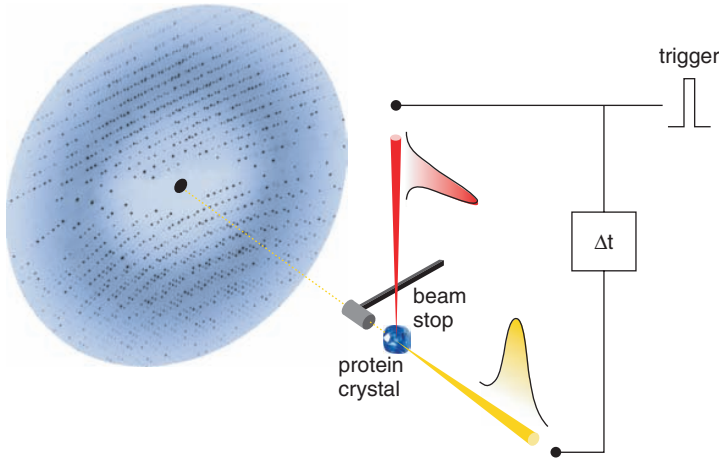

Figure 5.44 *Experimental setup for photoactivated protein reactions. The protein crystal is activated by a short and intense laser pulse (shown in red) and then irradiated after a variable delay Δt with a single x-ray pulse (yellow) from the synchrotron source, operating in either single-bunch or hybrid mode. If the chemical process is reversible, this can be repeated at the same Δt until an acceptable signal-to-noise ratio is achieved.*

The first hard x-ray femtoslicing source came online in 2006 at the Swiss Light Source, and is capable of going down to wavelengths of 1 Å [30].

Femtoslicing functions by the electron beam interacting with the co-propagating laser beam in a modulator (see also Section 3.8.2). The modulator is resonantly tuned so that the periodicity of the magnetic poles λ_m matches the laser periodicity λ_l, that is

$$\lambda_l = \frac{\lambda_m}{2\gamma^2}\left(1 + K^2/2\right). \tag{5.24}$$

In the case of the setup at the Swiss Light Source, the modulator contains 17 periods of length $\lambda_m = 138$ mm, the K-value of the modulator wiggler is approximately 23, and the laser wavelength is $\lambda_l = 805$ nm. The laser beam has a peak electrical field strength of $10^{10}-10^{11}$ V m^{-1} (Figure 5.45). This couples to the electron beam, thereby modulating its energy across a narrow slice of the electron bunch. Through energy dispersion, this femtosliced section is spatially separated from the main part of the bunch using a three-dipole chicane and then travels through a short-period undulator to produce x-ray pulses of less than 100 fs width.

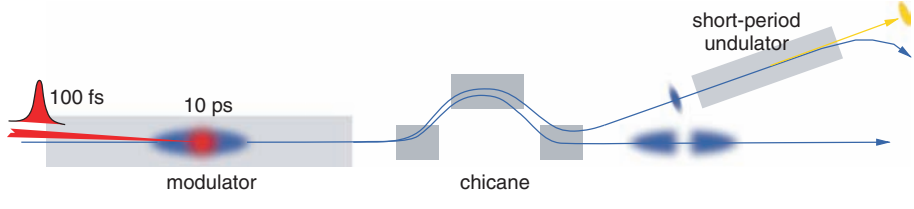

Figure 5.45 *Generation of femtosecond x-ray pulses via femtoslicing.*

These sources provide exceedingly modest fluxes compared to those promised by x-ray free-electron lasers. They are, however, very much cheaper to implement, and have provided valuable insights for future studies into how such ultrafast 'pump-probe' experiments should be carried out.

Femtoslicing has become possible due to advances both in femtosecond laser technology (which is gradually evolving from requiring a PhD in laser physics to operate, to becoming a 'turn-key' tool) and in undulator quality, whereby the spatial and temporal stability of both the x-ray and laser beams is crucial to the success of the technique.

An early illustrative example of femtoslicing using hard x-rays (1.73 Å) was performed on single crystals of bismuth [31]. A schematic of the experimental setup is shown in Figure 5.46. The sliced x-ray femtosecond pulse provides only 180 photons per pulse but at a rate of 1 kHz. The intense infrared laser pulse excites phonon modes in the crystal lattice, which 'swings' in response. Because the excitation is through the crystal surface, the propagation of the energy via the phonons into the bulk of the crystal takes some time. This could be monitored by varying the incident angle of the x-ray beam between 0.4°, just below the critical angle, for which the penetration depth is approximately 5 nm; and 2.0° (penetration depth > 100 nm). By observing how the phonon-vibration amplitude fell away as more and more of the bulk of the crystal was probed, the investigators were able to estimate the diffusion rate for the highly excited carriers of $2.3 \pm 0.3\,\mathrm{cm^2\,s^{-1}}$.

5.13 Surface Diffraction

5.13.1 Introduction

Surface x-ray diffraction (SXRD) is concerned with the study of the structure of crystalline surfaces and interfaces [32, 33]. SXRD studies are important because the surface

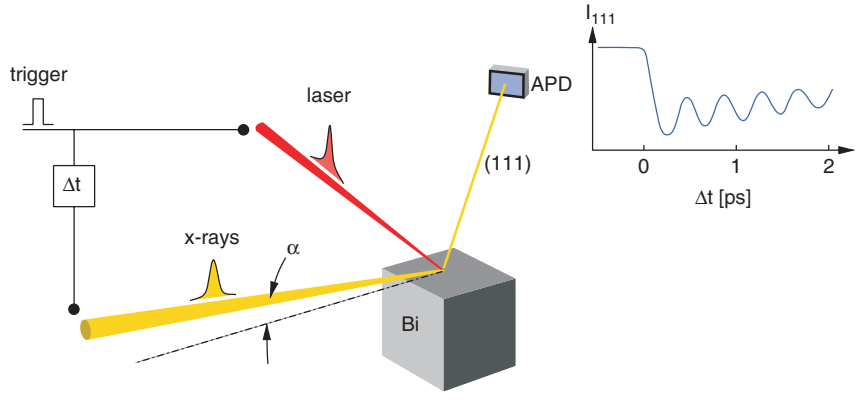

Figure 5.46 *A single crystal of bismuth metal is excited using a femtosecond laser. An avalanche photodiode is placed at the (111) reflection of the crystal under equilibrium conditions. The transient change in the crystal structure is monitored by recording the (111) diffraction intensity as a function of the delay time Δt between the laser pulse and the 140 fs x-ray pulse. This experiment was performed using different grazing-incidence angles of the x-ray pulse between $\alpha = 0.4$ and 2.0°.*

regions of crystals and crystalline interfaces between two heterogeneous materials rarely have the same structure as the bulk, and it is here that many physical and chemical processes take place. Indeed, we define the surface and interface regions as being those which differ from those of the bulk, which typically involve a depth of three or four monolayers or, otherwise expressed, a depth of approximately 2 nm, the extent of (screened) Coulomb interactions in covalent or ionic solids. We would therefore like to characterize the structure of these surface regions to better understand and possibly exploit their physical and chemical properties.

In creating a surface, the near-surface atoms invariably reposition themselves to minimize the electrostatic energy of the electrons of the uppermost atoms that have failed to bond (because they have fewer neighbours than do bulk atoms). Concerted movements of whole atomic layers perpendicular to the plane alone are called relaxations, while larger periodic structures in the plane at or near the surface are called reconstructions (see Figure 5.47).

Over and above this, any atoms chemically or physically absorbed on the surface will also affect the surface properties. It is therefore imperative in SXRD that the surface is well-defined, so that reliable conclusions can be drawn from the experimental data. This means that most surface diffraction experiments are performed in ultra-high vacuum chambers equipped with beryllium windows, which are reasonably transparent to x-rays above energies of about 4 keV. Notable exceptions are studies of interfaces between solids or a solid and a liquid.[12] Analysis of SXRD data is often facilitated by complementary information gleaned by other surface techniques such as scanning tunnelling microscopy, reflection high-energy electron-diffraction, low-energy electron-diffraction, and x-ray photoelectron spectroscopy.

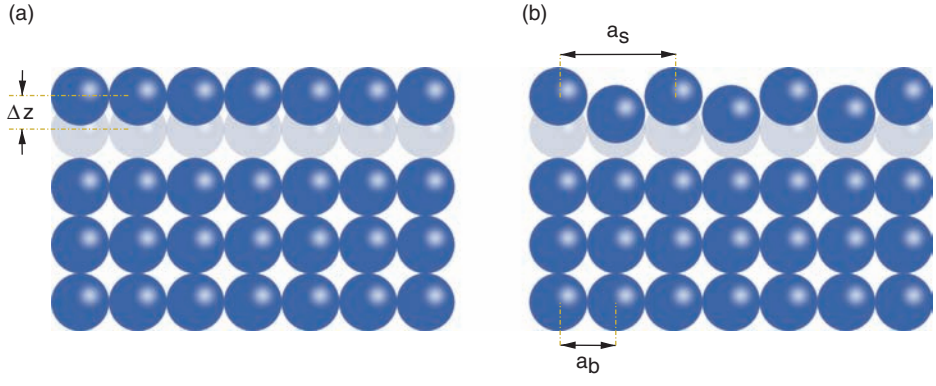

Figure 5.47 *(a) Movements perpendicular to the surface of the uppermost atomic layer(s) Δz produce surface structures referred to as relaxations. (b) Rearrangements of atoms in the plane, unequal movements out-of-plane, or different chemical compositions of the surface atoms can lead to periodic structures in the plane a_s that are not equal in size to the bulk periodicity a_b, in which case one speaks of reconstructions. In this simple schematic, the surface 'net' periodicity is twice as large as that of the bulk.*

[12] On a technical note, however, it is normally desirable to immerse the sample in either a modest vacuum or in a He-gas atmosphere, in order to minimize air scatter immediately above where the x-rays impinge on the sample.

5.13.2 Crystal Truncation Rods

In the simplest derivation of diffraction patterns, two assumptions are made – first that one is operating in the kinematical limit, that is, single-scattering, and second that the crystal is infinitely large. This results in the diffraction peaks being infinitely narrow (known as 'delta functions'). In reality, of course, all diffraction spectra are smeared out to a certain degree because there is partial absorption and extinction; and dynamical multiple scattering cannot be ignored close to the Bragg peaks (see Section 5.4.4).

In addition, crystals are finite in extent and one therefore measures a finite sample volume. The diffraction pattern of a finite crystal can be generated by convolving the Fourier transform of an infinitely large crystal structure (i.e. its 'ideal' diffraction pattern) with the Fourier transform of the function describing the boundary of the real crystal (called the 'shape function'). In most cases (for example, in bulk single-crystal diffraction), the boundary function is irregular on an atomic scale and hence its Fourier transform, the shape function, is exceedingly narrow and broadening effects other than that due to convolution with the shape function will determine the linewidths of the diffraction peaks.

However, a single crystal terminated with an atomically flat surface has a step function as the boundary function (see Figure 5.48). This has an FT showing a $1/k_z$-relationship

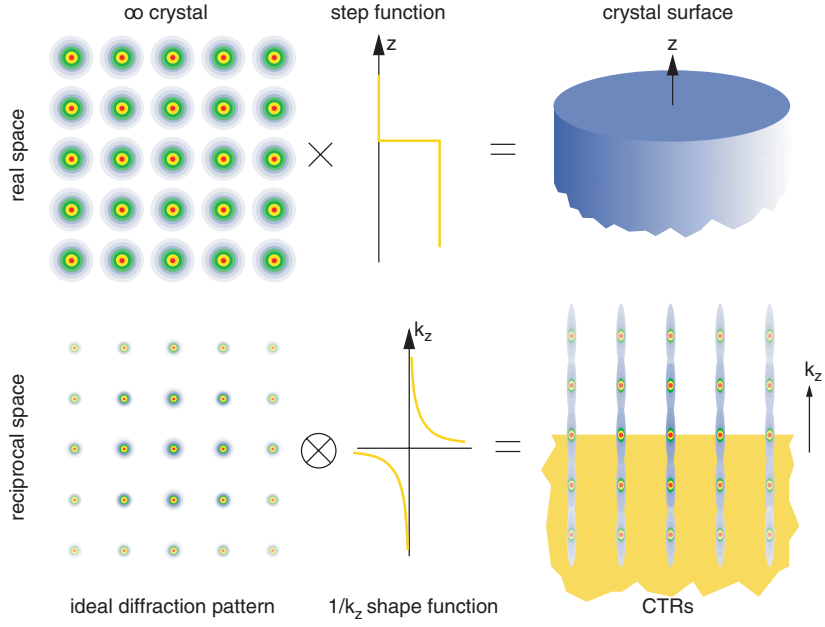

Figure 5.48 *The generation of CTRs, pictorially explained using the convolution theorem. An atomically flat crystal surface can be mathematically described as a* semi-*infinite crystal, whereby an infinitely large crystal is multiplied by a step function, having a value of unity for $z \leq 0$ and zero for $z > 0$. In reciprocal space, the idealized diffraction pattern of the infinite crystal is convoluted with the Fourier transform of the step function, which has a $1/k_z$-dependence. The result of truncating the crystal in this manner is that the diffraction signal perpendicular to the surface is continuous, hence crystal truncation rods (CTRs).*

that extends significantly in reciprocal space. Hence, the convolution of this with the 'ideal' diffraction pattern results in the latter being smeared out to produce a continuous signal in the direction perpendicular to the sample surface. These are crystal truncation rods (CTRs).

The signal intensity between Bragg peaks is so weak because the vector sum of the scattering amplitudes from the different atomic layers is almost completely destructive. One can easily show that the intensity exactly in between Bragg points is of the same order of magnitude as that from a single atomic layer. The Argand diagram of Figure 5.49 shows this schematically. Hence, any shifts in the atomic positions of the upper layers

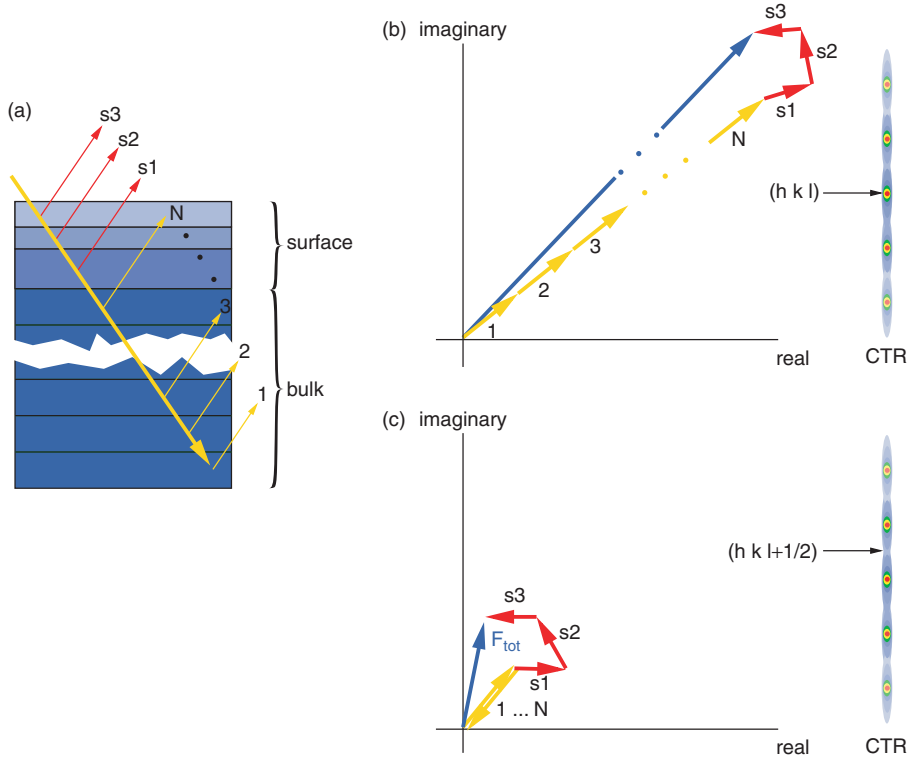

Figure 5.49 *A phase diagram depicting the differing contributions from the surface region at Bragg maxima and between Bragg maxima. (a) Consider scattering from a crystal with a thickness of N monolayers, where N is a large number. (b) In the kinematical approximation, each monolayer of the bulk crystal contributes the same amount (the yellow arrows) to the total scattering amplitude (shown as the blue vector), and at Bragg peaks, they all have the same phase, that is, they add up in a straight line. Because N is large, the relative impact of scattering from the uppermost layers s1, s2, and s3 (shown in red, and which contribute differently from the bulk monolayers in both magnitude and phase, due to their different structure) on F_{tot}, is minimal. (c) At the anti-Bragg point, bulk contributions from successive monolayers are π out of phase with each other and cancel each other out. Therefore, deviations from this destructive interference, as in the vectors s1, s2, and s3, have a large* relative *effect on the magnitude of F_{tot}. Thus, where signal is weakest, the relative effect of surface changes in the structure is most prominent.*

from their bulk positions, due to surface reconstructions and/or relaxations, will have a marked effect on the form and magnitude of the scattered amplitudes in portions of the CTRs away from the Bragg maxima. For large numbers of scattering planes (N in Figure 5.49), interference becomes very quickly destructive away from the Bragg maxima, and a great deal of information can be gained in almost all the region along CTRs in between Bragg peaks.

Recording CTRs therefore provides an exceptionally sensitive method for unravelling the structure of crystalline surfaces. But here's the rub – because most of the useful information on the near-surface structure of a material is extracted from the weakest portions of the CTR in between the Bragg maxima, this information is particularly difficult to record – there might be intensity ratios between the Bragg maxima and the weakest portions of the CTR of as much as eight orders of magnitude.[13] Hence, most surface diffraction experiments require the high flux available at third-generation synchrotron facilities and detectors with a high dynamic range and low background noise. Even then, great care must be taken to minimize diffuse scatter, typically by working in vacuum or a He-atmosphere and by a judicious use of beam-defining slits. On the other hand, although x-rays are only weakly scattered by a monolayer of crystalline material, this has the benefit that SXRD readily satisfies the kinematical approximation, simplifying analysis compared, for example, to surface-analytical techniques using electrons, such as low-energy electron-diffraction (LEED), for which a dynamical approach and an associated much greater computational effort is required in order to obtain quantitative data from the signal intensities.

CTRs become interesting when the top few layers deviate from the structure of the bulk in the direction of the surface normal. This causes the CTR to become skewed. As the deviation of the uppermost layer z increases, so does the amount of skewing. This is shown for $\Delta z/a = 0, \pm 0.05$, and ± 0.1 in Figure 5.50(a).

CTRs of thin-film systems contain other important information about the film properties apart from the detailed atomic structure, including degree of crystallinity, film thickness, density, and interface and surface sharpness. An example of a 4-unit-cell thick $YBa_2Cu_3O_{7-x}$ film grown on $SrTiO_3$ is shown in Figure 5.51.

5.13.3 Superstructure Rods

Superstructure rods (SSRs) are a signature of reconstructions. A surface structure which is, for example, $m \times n$ times larger in the plane of the surface than the bulk structure, will have SSRs in reciprocal space separated by $2\pi/ma$ and $2\pi/nb$, whereby m and n are integers and a and b are the in-plane lattice constants of the bulk material (see Figure 5.52).

The intensity and form of SSRs are in general very different from those of CTRs. Because they are manifestations of surface effects, the number of scatterers compared to those contributing towards CTRs is much smaller, and SSRs are therefore weak, having intensities of the same order of magnitude as that of the weakest regions of the CTRs

[13] This value can be estimated using a back-of-an-envelope calculation and the fact that the intensity of the Bragg peak is approximately proportional to the square of the number of involved scattering planes. This is typically of the order of $N = 10^4$. Remembering that the weakest signal has an intensity of the order of that produced by a single scattering plane, we obtain a ratio of $(10^4)^2 = 10^8$.

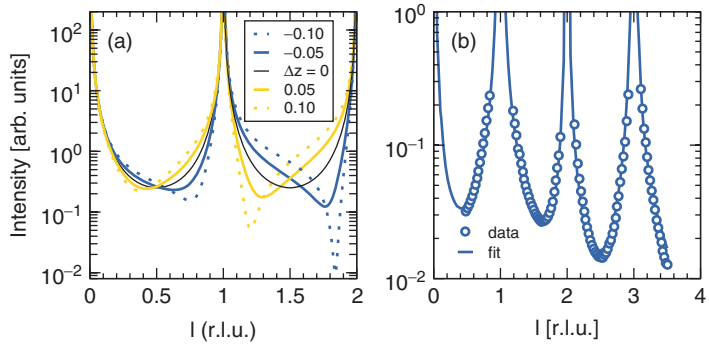

Figure 5.50 *Surface relaxations detected by SXRD (a) The effect of the displacement of the uppermost layer (Δz, given as a fraction of the interlayer spacing) becomes more pronounced as Δz is increased. (b) Relaxation of the topmost atomic layer of Ru(0001) is best fit by a contraction of $2.2 \pm 0.1\%$ relative to the bulk value of 2.141 Å. Open circles: experimental data, solid line: fit. r.l.u. = reciprocal lattice unit. Adapted from [34] with permission of the Institute of Physics.*

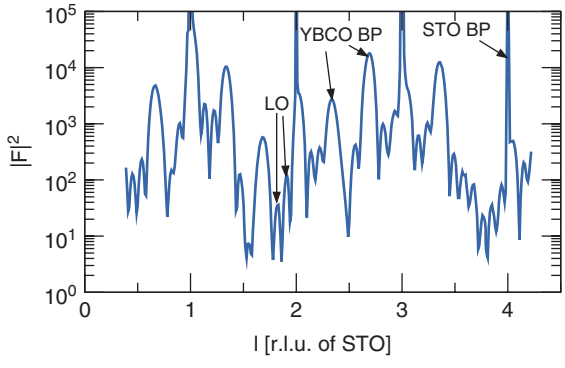

Figure 5.51 *The 21l-CTR of a 4-unit-cell thick film of $YBa_2Cu_3O_{7-x}$ (YBCO) grown on $SrTiO_3$ (STO). The substrate Bragg peaks (STO BP) are very sharp and lie on integer values of the abscissa to within the accuracy of the experiment of 10^{-4} reciprocal lattice units (r.l.u.). Because the YBCO film is so thin (4 unit cells, approximately equivalent to 12 unit cells of STO), its Bragg peaks (YBCO BP) are correspondingly broad. In between the YBCO Bragg peaks are finite-size oscillations, known as Laue oscillations (LO). In contrast to Kiessig fringes, discussed in Section 5.15, these only occur in nonspecular CTRs if the film is crystalline, as they require in-plane momentum transfer and therefore in-plane periodicity. Adapted from [35] with permission of the American Physical Society.*

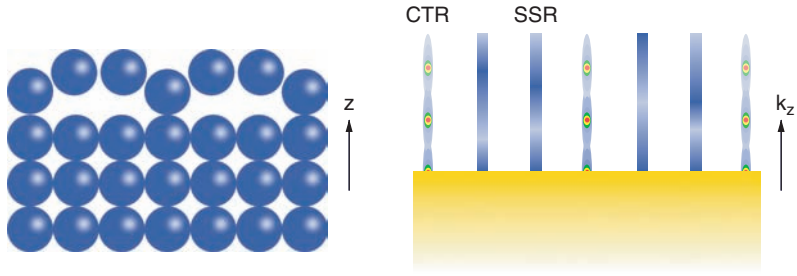

Figure 5.52 *Left: a surface reconstruction with a periodicity three times that of the bulk. Right: in reciprocal space, the CTRs are interspersed with SSRs, whereby the separation between adjacent SSRs is three times less than that between CTRs.*

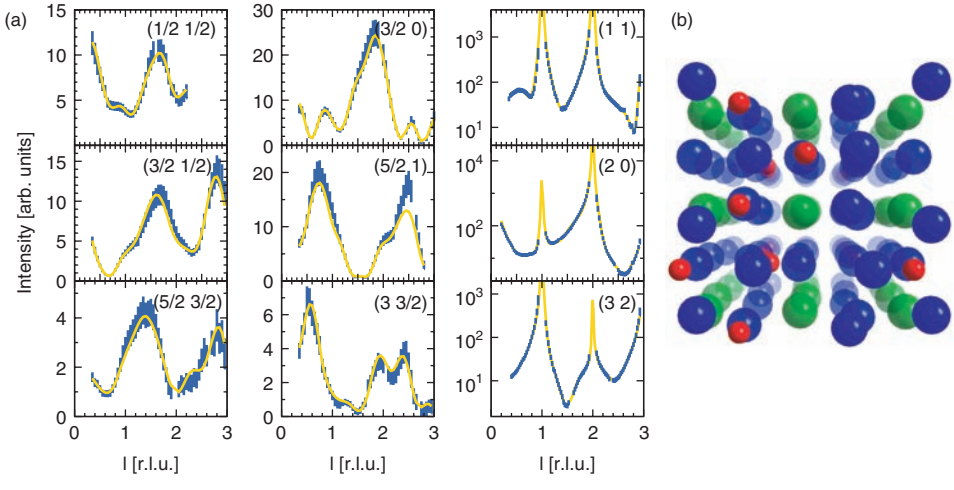

Figure 5.53 *The surface of strontium titanate. (a) Six representative SSRs and three CTRs (blue), plus fits (yellow solid lines), taken from a set of SXRD data recorded of the surface of $SrTiO_3(100)$. Note that the characteristic width of oscillations in the SSRs is approximately 0.5 r.l.u. (b) The structural model of the (2×2) reconstruction that generates the fits for the half-integral–half-integral SSRs. Ti: red, O: blue, Sr: green. Adapted from [36] with permission of the American Physical Society.*

(see, for example, Figure 5.53). In addition, the relative variation in intensity is much more modest than that for CTRs, as only a few layers contribute to the signal.

5.13.4 Data Acquisition

Typically, SXRD measurements are carried out using a grazing-incident angle α, with respect to the sample surface. The choice of α depends on the experiment, sample quality, and scattering strength. For high-quality crystals consisting of heavy elements, such as metal-oxide perovskites, one can normally choose α to be significantly above the critical angle. This has the advantage that the x-rays penetrate deeply into the sample and one can easily subtract the contribution from the bulk during subsequent analysis. For

incident angles close to the critical angle α_c, the surface sensitivity increases rapidly. Exactly at α_c, the reflected wave is perfectly in phase with the incident wave, and the evanescent wave amplitude is approximately twice that of the incident wave. The evanescent *intensity* therefore approaches four times that of the incident beam. The penetration depth is low, and so the bulk contribution is suppressed. Hence, at α_c, the surface sensitivity is highest, which is advantageous for weakly scattering surfaces and also for systems with crystallographically poor bulk material (such as some metal crystals) which produce significant diffuse background.

SXRD can be divided into two activities – in-plane diffraction and out-of-plane measurements. The first is used to study the in-plane positions of the atoms in the uppermost layers, and reveals features such as surface reconstructions. By recording a large set of in-plane peaks, it is possible to obtain a two-dimensional Patterson map (see Section 5.6.2), which often provides important information as to the probable configuration of the atoms in the surface plane. Nonintegral features yield information on the period and orientation of any reconstructed surface net. In this context, qualitative information from LEED measurements or similar techniques can also provide invaluable prior information regarding where one should search for signal in reciprocal space.

The in-plane points can then be used as the starting points for out-of-plane measurements, in the form of CTRs, but also SSRs if the surface is reconstructed. The intensity distribution along a CTR or SSR depends on the structure of the near-surface region perpendicular to that surface. CTR and SSR scans are therefore complementary to in-plane diffraction studies.

Normally, the grazing-incidence beam floods an area of the sample of at least a few hundred square microns and one therefore records a macroscopic average. However, recent work on SiGe-islands grown on silicon using a $3\,\mu m \times 5\,\mu m$ microfocused hard x-ray beam demonstrated that one can map out the positions and characteristics of individual micronsized islands separated by approximately $10\,\mu m$ by probing those structure factors of the islands which do not overlap with those of the substrate. The accuracy of this technique could be confirmed by comparing the maps with scanning electron microscope images of the same samples [36].

5.13.5 Worked Example – The Surface of Strontium Titanate

Strontium titanate ($SrTiO_3$, STO) is the most commonly used substrate crystal for the thin film growth of technologically important perovskite materials such as $LaAlO_3$, $La_{1-x}Sr_xMnO_3$ and $YBa_2Cu_3O_7$. Despite this, there remains a good deal of ambiguity as to the structure of its surface, which seems to depend strongly on the preparation and processing conditions. In the example given here, an STO substrate was etched using an established method, resulting in the surface being over 99% TiO_2-terminated, and being stable in air. This pre-etching process was necessary in order to obtain a well-defined crystalline surface terminated on a single atomic layer. The sample was introduced into an ultra-high vacuum chamber for surface diffraction measurements.

Figure 5.53 shows a representative selection of 6 SSRs and 3 CTRs and their fits from a total data set of 27 rods (containing over 1800 independent data points) of the surface of STO, recorded using 1-Å radiation. The time required to record this data set was about 10 hours. Inspection of the SSRs in Figure 5.53 shows intensity fluctuations

with periodicities of the order of 0.5 r.l.u. This corresponds in real space to deviations of the surface region down to a depth of approximately two unit cells. Therefore, the data was fit to a model of a surface region (i.e. one that differs structurally from that of the bulk) down to a depth of three unit cells and which involves a mixture of surface reconstructions. The fit required approximately $N_{fp} = 200$ fitting parameters [37].

A general feature of the surface of STO which emerged is that it is terminated by an double TiO_2 atomic layer – the Ti-atoms in the upper layer form an unusual zigzag motif (the red balls in Figure 5.53(b)). This Ti-rich surface may provide a reservoir of Ti-atoms that can be incorporated into a film growing on top and thereby affect its physical and electronic properties.

5.14 Resonant X-ray Diffraction

As we have already seen in Section 5.11.3, elastic anomalous scattering has long been employed by crystallographers to help solve the phase problem. It can also be exploited in the technique of resonant x-ray scattering (RXS) to probe unoccupied states close to the vacuum continuum in a second-order process. Consider Figure 5.54. An x-ray photon of energy $h\nu = E_f - E_i$ is absorbed, causing an electron occupying the core level $|i\rangle$ to be excited to the unoccupied state $|f\rangle$ [14] with a certain probability, which then promptly relaxes again to the original core level $|i\rangle$ via emission of an x-ray photon having the same energy as the incident photon. This process is therefore elastic. Crucially, both the absorption strength of the incoming photon and the emission probability of the final photon depend directly on the dipole transition between the lower state and upper state. Quantum selection rules between the two states, and the polarization of the electric field of the x-ray beam determine the dipole strength. RXS is therefore a sensitive spectroscopic probe for the upper state and provides both spatial and site-selective (chemical, spectroscopic) information. This is not normally the case in conventional diffraction – the relative contribution of valence-state electrons in the first-order process

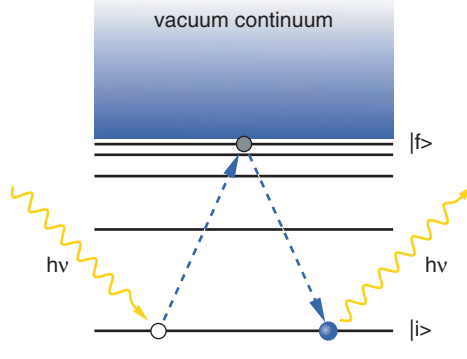

Figure 5.54 *Schematic of the second-order resonant scattering process, which occurs via the intermediate state $|f\rangle$. The photon energy $h\nu = E_f - E_i$.*

[14] Pauli's exclusion principle dictates that the intermediate state $|f\rangle$ must be unoccupied.

of direct scattering is negligibly small as the x-rays are scattered by all the electrons around the atom.

Although the probed state is unoccupied, it has related symmetry properties to occupied valence states – an example might be the so-called $d(3z^2 - r^2)$ and $d(x^2 - y^2)$ e_g states in manganite perovskites [38]. The importance of valence-state electrons can hardly be overstated, as they determine most, if not all, of the physical, structural, and electronic properties of condensed matter.

The material $La_{1-x}Sr_xMnO_3$ (LSMO, see Figure 5.55(a)) is a ferromagnetic compound which exhibits so-called 'colossal magnetoresistivity' (CMR) in the range $0.15 \leq x \leq 0.5$, whereby its resistivity drops by orders of magnitude, driven by the application of an external magnetic field. Importantly, La and Sr are always trivalent and divalent, respectively, but Mn can assume several different valencies. In the compound $LaMnO_3$ ($x = 0$), all the Mn-ions must be trivalent (Mn^{3+}) to maintain charge neutrality. For each substitution of La with Sr, however, one Mn atom becomes tetravalent (Mn^{4+}). Hence, LSMO:$x = 1/2$ contains one half Mn^{4+} and one half Mn^{3+}. The valence states and the orbitals of the Mn^{3+} ions for this compound are arranged in an ordered manner, as shown schematically in Figure 5.55(b).

The unit cell of LSMO is normally considered to be quasicubic, with one chemical unit of $La_{1-x}Sr_xMnO_3$ occupying a simple-cubic unit cell (although in reality, very slight distortions mean that the unit cell is, strictly speaking, orthorhombic with a four-times larger volume). Even if one assumes a pure cubic unit cell, however, charge ordering increases the unit cell size by a factor of two, which takes into account the ordered

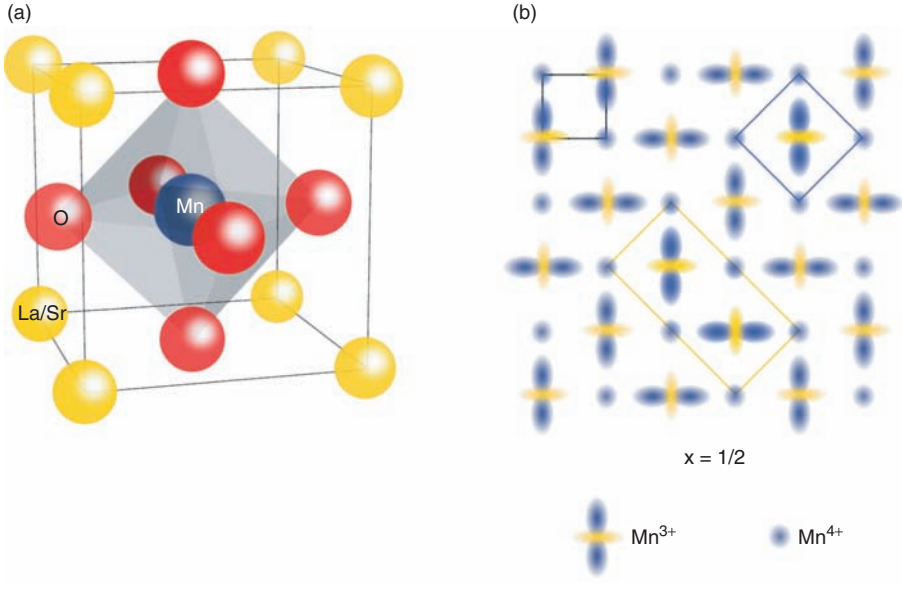

Figure 5.55 *Charge and orbital ordering in the compound $La_{1-x}Sr_xMnO_3$ (a) the perovskite quasicubic structure of LSMO. (b) The charge and orbital order for LSMO ($x = 1/2$). Note that the size of the unit cell is larger than the 'chemical' unit cell (shown in black) when the charge ordering (blue) and orbital ordering (yellow) are taken into account.*

arrangement of the different valence and chemical states of the Mn-ions, (though this is all too often neglected). More subtle is the unit cell which also takes into account the *orbital* ordering – here the Mn^{3+} ions are all chemically identical, and only the direction of the orbitals determines the cell size, which is now *four* times larger than the cubic cell.

Hence, charge- and orbital-ordering cause the unit cell size to increase compared to that of the 'chemical' unit cell, which results in Bragg reflections at noninteger scattering vectors (expressed in units of the quasicubic structure) that one would normally not see in nonresonant diffraction, due to the low scattering strength of the valence electrons. However, the RXS signal is detectable exactly because it specifically probes these upper valence state. Ordering of electron *spin* is yet another degree of freedom that can cause reconstructions.

Note that in this example, it is the $3d$ states of the Mn atom that are responsible for charge-, orbital-, and spin-ordering. The selection rules for electronic dipole transitions induced by the absorption of a single photon dictate that the orbital angular-momentum quantum number l changes by ± 1 (although it is briefly noted that in electric *quadrupole* transitions involving the absorption of two photons, $\Delta l = 0, \pm 2$ and the $3d$ orbital can be accessed from the $1s$ state). Hence, the $3d$ state can only be accessed by a $2p$ state, in other words through an L absorption. Such experiments investigating the orbital properties of the $3d$-transition metals therefore use photon energies of around 500 to 1000 eV [39]. As a result, resonant *soft* x-ray scattering (RESOXS) is at the extreme limit of x-ray diffraction, where absorption by low-Z material is exceedingly high (see Figure 4.5): the entire beamline must be windowless (including the optics and diffractometer), and the vacuum must be less than 10^{-8} mbar. In addition, the Ewald sphere is very small, with a radius of the order of $2\pi/\lambda \sim 0.5\,\text{Å}^{-1}$ and may contain only a single Bragg reflection.

The nature and orientation of the probed state regarding the electron distribution (dipole, quadrupole, etc.) can be distinguished in ordered systems by its dependence on the polarization of the incident x-ray beam.

RXS experiments employ the selected Bragg peak method (Section 5.8.2). Typical procedures involve moving the diffractometer to the Bragg condition for the suspected superlattice reflection (e.g. to the (3/4, 3/4, 0) position for orbital order in $La_{1/2}Sr_{3/2}MnO_4$) and then scanning the photon energy across the absorption edge of interest.

RXS is a relatively new experimental technique, due primarily to the weakness of the signal, which only provided a tolerable signal-to-noise ratio once the brilliance from third-generation synchrotron sources became available. The first results on charge and orbital ordering in the so-called 'half-doped' layered manganite $La_{1/2}Sr_{3/2}MnO_4$ were reported by Murakami *et al*. in 1998 [40] and showed charge ordering with a unit cell $\sqrt{2} \times \sqrt{2}$ times larger in the a–b plane (and rotated by 45°) and orbital ordering with a unit cell that was $2\sqrt{2} \times \sqrt{2}$ times larger in the a–b plane and also rotated by 45°. The energy spectrum of the latter at the (3/4, 3/4, 0) reflection is shown in Figure 5.56.

5.15 X-ray Reflectometry

5.15.1 Introduction

X-ray reflectometry (XRR) is a nondestructive technique used to precisely determine the thickness, roughness, and density of surfaces, thin films, interfaces, and multilayers.

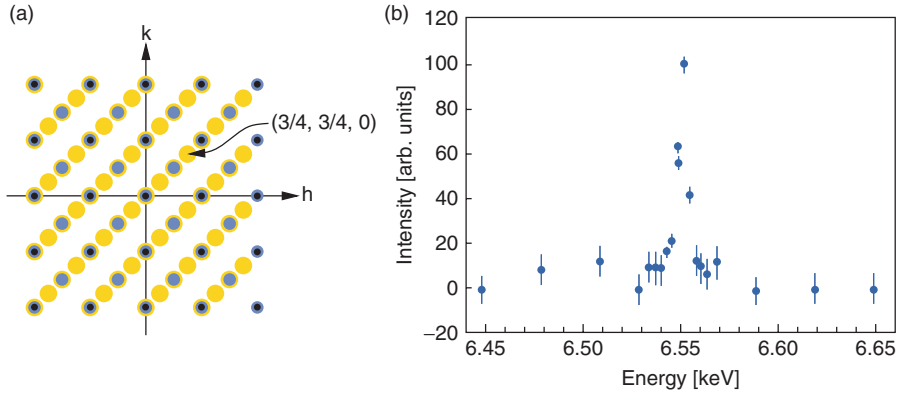

Figure 5.56 *RXS data for the layered perovskite $La_{1/2}Sr_{3/2}MnO_4$. (a) The superlattice reflections due to charge ordering (blue) and orbital ordering (yellow) are shown relative to the reciprocal lattice of the conventional 'chemical' unit cell (black). The real-space configuration of the charge- and orbital-ordering is the same as that shown in Figure 5.55(b). (b) The energy dependence of the orbital-ordering superlattice peak at (3/4, 3/4, 0) near the Mn K-edge in $La_{1/2}Sr_{3/2}MnO_4$. Adapted from [40] with permission of the American Physical Society.*

XRR is performed simply (and with synchrotron radiation also rapidly) by measuring the specularly reflected x-ray intensity as a function of the grazing-incidence angle, typically up to approximately 2°. This allows one to determine film thicknesses, multilayer periodicities, and electron-density profiles over a range spanning approximately 1 to 500 nm, whereby the lower limit is determined by the maximum recorded incident angle, and the upper limit by the divergence of the incident x-ray beam in the plane perpendicular to the sample surface.

XRR works equally well for amorphous, polycrystalline or indeed liquid materials as for crystalline materials, as it essentially probes contrast in the average electron density rather than scattering from individual atoms, and can be thought of as a special form of small-angle x-ray scattering, discussed in detail in Section 5.16.

Most XRR experiments start at angles slightly below α_c, the critical angle for total external reflection (Equation (2.18)), typically somewhere near 0.1°. The footprint on a sample of an x-ray beam with a cross-sectional height of 0.1 mm at this incident angle is therefore of the order of 5 mm. Typical linear dimensions of XRR samples are hence a few mm to a few cm. XRR scans can often span five orders of magnitude or more in the recorded reflectivity intensity, and are as such ideally suited to intense x-ray sources like synchrotrons.

5.15.2 Reflection of X-rays and the Fresnel Equations

Consider Figure 5.57. We describe the incident, reflected, and transmitted waves as

$$\begin{aligned} \psi_i &= a_i e^{i\mathbf{k}_i\cdot\mathbf{r}}, \\ \psi_r &= a_r e^{i\mathbf{k}_r\cdot\mathbf{r}}, \\ \psi_t &= a_t e^{i\mathbf{k}_t\cdot\mathbf{r}}, \end{aligned} \tag{5.25}$$

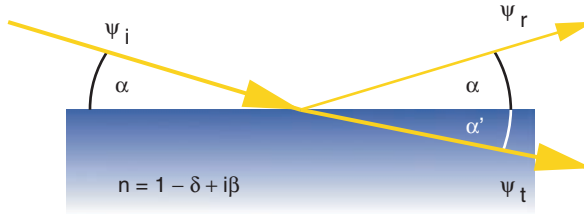

Figure 5.57 *X-ray reflectivity from a surface. The incident beam ψ_i impinges on the surface with an angle α and is reflected specularly (ψ_r). The transmitted beam ψ_t is refracted to a lower angle α', because the refractive index n of the medium is less than unity.*

whereby i, r, and t refer to the incident, reflected, and transmitted beams, respectively. Note that $|\mathbf{k}_i| = |\mathbf{k}_r| = k$, the wavevector magnitude in vacuum, while $|\mathbf{k}_t| = nk$ is the magnitude of the wavevector in the medium of refractive index n.

We have already reasoned in Chapter 2 that below α_c, the critical angle for total external reflection, x-rays are totally reflected. From Equations (2.15) and (2.18), we see that

$$\alpha_c = \lambda\sqrt{\frac{\rho r_0}{\pi}} \tag{5.26}$$

and hence α_c scales with the square-root of the electron density, and as such is an accurate measure of this quantity. Interestingly, essentially all condensed matter has an average electron density somewhere between 0.5 and 5 electrons per cubic Angstrom, so a useful rule-of-thumb is that the critical angle expressed in degrees lies somewhere in between 0.12 and 0.4 times the wavelength given in Angstrom.

At all angles, the reflectivity is determined by the boundary conditions that the normal components of the electric field of the electromagnetic wave and its derivative at the interface must be continuous functions. From Equations (5.25), we therefore obtain for $\mathbf{r} = 0$

$$a_i + a_r = a_t$$

and

$$a_i\mathbf{k}_i + a_r\mathbf{k}_r = a_t\mathbf{k}_t,$$

from which the Fresnel equations for reflectivity and transmittivity can be derived. Here we are primarily interested in the reflectivity amplitude, which above the critical angle is given by

$$r = \frac{a_r}{a_i} = \frac{\alpha - \alpha'}{\alpha + \alpha'}, \tag{5.27}$$

where α and α' are the incident and transmitted angles, respectively (see Figure 5.57). This quantity changes subtly, depending on the polarization of the x-ray electric field relative to the reflecting surface. The difference between parallel or perpendicular polarization is very small, and we limit ourselves here to the case whereby the electric field

is quasiperpendicular to the surface.[15] In this case, the reflection amplitude is

$$r = \frac{\sin\alpha - (n^2 - \cos^2\alpha)^{1/2}}{\sin\alpha + (n^2 - \cos^2\alpha)^{1/2}}. \tag{5.28}$$

We use the simplifications $n = 1 - \delta$, $\sin\alpha \approx \alpha$, and $\cos\alpha \approx 1 - \alpha^2/2$ for small angles and thereby obtain

$$r = \frac{1 - \left(1 - 2\delta/\alpha^2\right)^{1/2}}{1 + \left(1 - 2\delta/\alpha^2\right)^{1/2}}. \tag{5.29}$$

The measured reflectivity *intensity* is $R = |r|^2$. We see immediately that at the critical angle $\alpha_c = \sqrt{2\delta}$, both terms within the brackets of Equation (5.29) are equal to zero and the reflectivity is 100%, as expected. Above α_c, however, the reflectivity drops off very rapidly, as already shown in Figure 2.13.

Once the incident angle becomes significantly larger than α_c, that is, when $2\delta/\alpha^2 \ll 1$, we can expand the terms in brackets in Equation (5.29) by recognizing that

$$\left(1 - \frac{2\delta}{\alpha^2}\right)^{1/2} \approx 1 - \frac{\delta}{\alpha^2}.$$

The reflectivity is then accurately approximated by

$$R = r^2 = \frac{\delta^2}{4\alpha^4} = \left(\frac{\alpha_c}{2\alpha}\right)^4 \tag{5.30}$$

and the reflectivity (a) falls off with the inverse fourth power of incident angle, and (b) is proportional to the square of the electron density ρ (see Equation (5.26) and Figure 5.58).

As soon as the surface has any roughness on a scale similar to or larger than the wavelength of x-radiation being used, the reflectivity drops still more sharply. For a root-mean-square vertical roughness σ, the reflectivity curve for a flat surface R_f is modified for $\alpha \gg \alpha_c$ such that

$$R_r = R_f \cdot \exp\left[-(2k\alpha\sigma)^2\right], \tag{5.31}$$

whereby R_r is the reflectivity for the rough surface, and is shown in Figure 5.58.[16]

In the above, we have ignored two important phenomena, namely absorption effects (by assuming that $|\beta| \ll |\delta|$), and possible multiple scattering of the x-rays, which can be significant for very low incident angles. It lies outside the scope of this book to describe these effects quantitatively, and the reader is referred to the literature for rigorous discussions [41].

[15] This case is referred to as *s*-polarized light, whereby an electric field which is exactly perpendicular to the surface must by definition have its direction of propagation parallel to the surface, that is, $\alpha = 0^\circ$. Hence. the need to qualify with the expression *quasi*perpendicular for XRR, whereby the incident angle is always small, if not exactly zero.

[16] Again, one can apply the convolution theorem (Section 5.5.1) to understand this: the rough surface is the convolution of the flat surface with the roughness function (a Gaussian), hence the reflectivity curve for the rough surface is simply that for the flat surface multiplied by the Fourier transform of the roughness function.

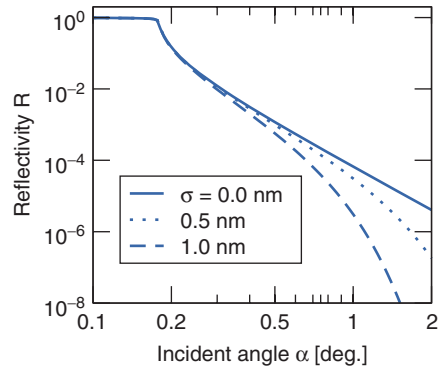

Figure 5.58 *Reflectivity curves for a silicon surface with different roughnesses. Note that on this double-logarithmic plot the reflectivity curve of the flat surface and for $\alpha \gg \alpha_c$ has a gradient of −4, as predicted by Equation (5.30).*

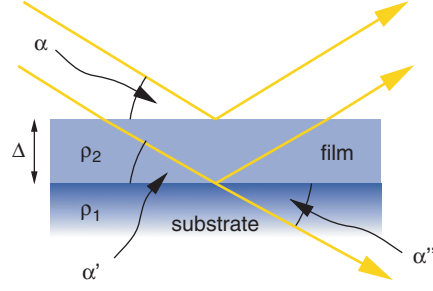

Figure 5.59 *Kiessig fringes occur when that part of the incident beam reflected from the upper surface of a thin layer interferes with that part reflected from the layer–substrate interface.*

5.15.3 Thin Films and Multilayers

Consider a system consisting of a perfectly flat slab of density ρ_1 (the 'substrate') covered with an equally flat thin layer of thickness Δ and density ρ_2 (the 'film', see Figure 5.59). As in the case of a homogeneous slab on its own, x-rays will be partially reflected from the upper film surface. In addition, if $\rho_1 \neq \rho_2$, the interface between the film and substrate will provide contrast for any x-rays impinging on it, and they will be partially reflected here, too. Interference between the two reflected beams cause the reflectivity curves of such systems to be modulated with so-called 'Kiessig fringes' [42].

The reflectivity of a single layer on a substrate is given by

$$R = \left| \frac{r_1 + r_2 \exp(i4\pi\alpha'\Delta/\lambda)}{1 + r_1 r_2 \exp(i4\pi\alpha'\Delta/\lambda)} \right|, \tag{5.32}$$

whereby

$$r_1 = \frac{\alpha - \alpha'}{\alpha + \alpha'} \exp\left[-2 \left(\frac{2\pi}{\lambda} \right)^2 \alpha\alpha'\sigma_1^2 \right], \tag{5.33}$$

$$r_2 = \frac{\alpha' - \alpha''}{\alpha' + \alpha''} \exp\left[-2\left(\frac{2\pi}{\lambda}\right)^2 \alpha'\alpha''\sigma_2^2\right]. \tag{5.34}$$

α'' is the angle of the x-ray beam transmitted into the substrate, while σ_1 and σ_2 are the roughnesses at the surface and substrate, respectively.

The optical path difference OPD is, to a high degree of accuracy, given by

$$\text{OPD} = 2\Delta \sin\alpha'. \tag{5.35}$$

This should equal an integral number of wavelengths for constructive interference. From Snell's law, we also know that

$$\cos\alpha = n\cos\alpha'$$

and, using the first two terms of the Taylor expansion for the cosine function, we obtain

$$1 - \alpha^2/2 = (1 - \delta)\left[1 - \alpha'^2/2\right].$$

If we multiply this out and ignore the term $\alpha^2\delta/2$, we obtain

$$\alpha' \approx (\alpha^2 - 2\delta)^{1/2}, \tag{5.36}$$

from which it follows that constructive interference occurs when

$$\alpha'^2 = \frac{\lambda^2}{4\Delta^2} m^2 + 2\delta, \tag{5.37}$$

whereby m is an integer. Hence a plot of α'^2 against m^2 yields a straight line with a gradient $\lambda^2/4\Delta^2$ which crosses the ordinate at 2δ. This provides both the film thickness and density. One potential problem, particularly for thicker films, is obtaining a reliable value for m, as the Kiessig fringes only occur above α_c, for which m can be much greater than unity. However, if we consider those fringes at relatively high incident angles, that is, where $\alpha \gg \alpha_c$, Equation (5.37) reduces to

$$\alpha' = \frac{m\lambda}{2\Delta} \tag{5.38}$$

and the angular separation between adjacent fringes is simply $\lambda/2\Delta$.

Both the modulation depth and the phase of the Kiessig fringes are affected by the contrast (that is, the difference in electron densities) of the film and substrate. Consider the simulated reflectivity curves shown in Figure 5.60(a). The modulation depth of Si on Au is smaller than that of Au on Si, as scattering from the film's top surface is weaker in the former case. Note also that the Kiessig fringes are π out of phase relative to each other. This is explained by the fact that when reflected off an interface to a medium with lower electron density (e.g. the Au-film/Si-substrate interface), the phase of the reflected beam is preserved, while reflections off surfaces to denser material (Si-film/Au-substrate) undergo a phase flip of π radians. When the difference in contrast between film and substrate is smaller, the Kiessig fringe amplitude drops accordingly, as also shown in Figure 5.60(a) for C on Si (with electron densities of 0.66 and 0.70 electrons per cubic Angstrom, respectively).

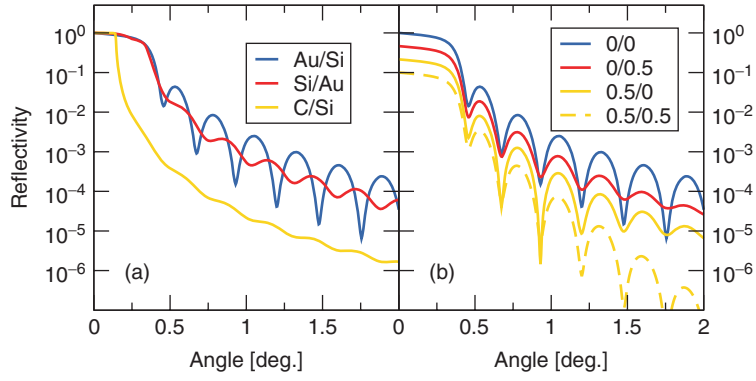

Figure 5.60 *(a) The effect of differing contrast on the reflectivity of thin films. Three film/substrate systems are shown: Au on Si; Si on Au; and C on Si. All films are 10 nm thick and therefore have the same Kiessig-fringe periodicity. There is no substrate or film roughness, and 1 Å radiation was used. (b) The effect of roughness on the reflectivity of a 10-nm thin film of gold on silicon. The roughnesses in the legend are ordered substrate/film and given in nm.*

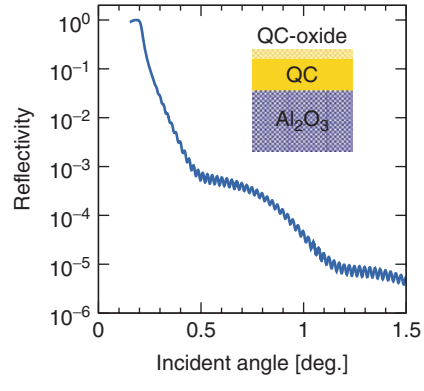

Figure 5.61 *An example of an experimental reflectivity curve of a quasicrystal (QC) thin film grown on sapphire (Al_2O_3). In addition to the high-frequency Kiessig fringes produced by the bulk film, an additional low-frequency modulation can be observed, which is caused by the uppermost 4 nm of the QC-film having been oxidized by exposure to air. Adapted from [43] with permission of Elsevier.*

Figure 5.61 shows an example of the reflectivity curve of a 120 nm-thick quasicrystalline icosahedral Ti-Ni-Zr thin film grown on Al_2O_3, exposed to air for several days [43]. In addition to the high-frequency Kiessig fringes produced by the relatively thick film, there is a much lower-frequency modulation. This is caused by the top 4 nm of the film becoming oxidized. The best fit has a density of the oxide of 4.2 g cm^{-3}.

The important parameters of periodic multilayer structures are shown in Figure 5.62(a). From a scattering perspective, these systems can be thought of as being two-dimensional 'artificial crystals', with lattice dimensions given now by Δ instead of the crystal interplanar spacing d, and the basis (which for crystals is given by the arrangement of

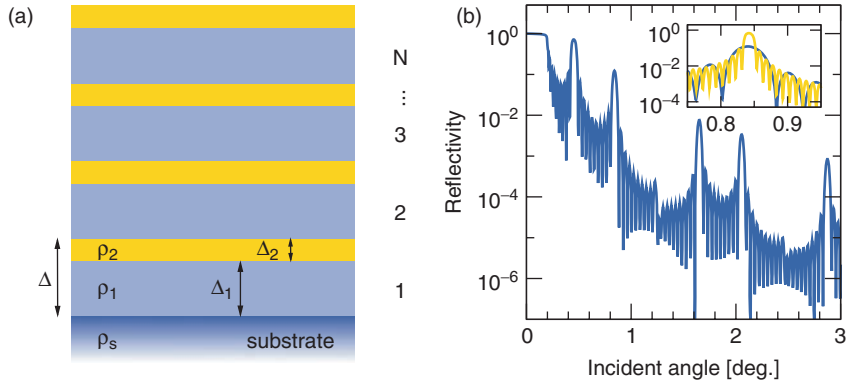

Figure 5.62 *Multilayer x-ray reflectivity. (a) Schematic figure of a multilayer, showing the defining parameters. (b) Simulation of a multilayer reflectivity curve at 12.398 keV for a ten-layer Mo/Si multilayer. $\Delta = 7$ nm, $\Delta_1/\Delta = 1/3$. Note that every third reflectivity peak is missing. Inset: a blowup of the second reflectivity peak near 0.85°. By increasing the number of layers from ten (blue curve) to 40 (yellow curve), the peak becomes both narrower and more intense.*

the atoms within the unit cell) being determined by ρ_1, ρ_2, Δ_1 and Δ_2. Their reflectivity curves therefore have similar characteristics to those of regular $\theta-2\theta$ diffraction spectra of heteroepitaxial films (see Figure 5.62(b)). The more multilayer periods there are, the sharper and more intense are the reflection maxima. An important feature is the ratio of thicknesses of the two sublayers. Remember that certain reflections ('systematic absences') can be missing in crystals due to destructive interference between scatterers within the unit cell, as already mentioned in Section 5.4.3. For the same reasons, multilayers with a sublayer ratio Δ_1/Δ equal to a reciprocal integer $1/M$ will have systematic absences for the orders M, $2M$, $3M$ ....

Multilayers can therefore be used as Bragg reflectors at angles far above the critical angle of either of the constituent materials. As already touched upon in Chapter 4, one particularly interesting application of multilayers is the fabrication of monochromator elements and compact x-ray mirrors in x-ray optical systems. Conventional mirrors at synchrotron beamlines have critical angles of a small fraction of a degree, and need to be of the order of a metre long in order to capture the footprint of the entire x-ray beam. As can be seen in Figure 5.62(b), interference maxima with reflectivities close to unity can be found at incident angles of the order of one degree, meaning that multilayer mirrors can be an order of magnitude smaller. In general such systems consist of layers composed of a thicker, low-Z material (to avoid excessive absorption) and a thinner high-Z material (to provide the necessary contrast for efficient reflection).

Because a relatively small number of scattering 'planes' are involved (given by N in Figure 5.62(a)), the bandwidth of reflectivity maxima from multilayer monochromators are large. Instead of bandwidths $\Delta\lambda/\lambda \sim 10^{-4}$ typical for crystal monochromators, multilayer monochromators have bandwidths of the order of a percent, thereby increasing the transmitted flux from the source accordingly. This is important for computed tomography and some other imaging techniques where flux is at a premium and the spectral purity of the x-radiation is of secondary importance.

5.15.4 Worked Example – Monitoring Monolayer-for-Monolayer Thin Film Growth

XRR can be used to study changes in the thickness or morphology of a thin film, particularly important for in-situ studies during heteroepitaxial film growth.

As discussed in Section 5.13, the *relative* change in the specular reflection intensity due to atomic-scale changes of the surface region is largest where the absolute signal is weakest, exactly in between Bragg peaks (for example, at the (0 0 1/2) position, see Figure 5.49(c)). The reason for this is that most of the signal cancels out, as successive layers provide scattering vectors that are 180° out of phase with one another. Hence, any changes in one of those layers (in this case, the top, growing layer, see Figure 5.63)

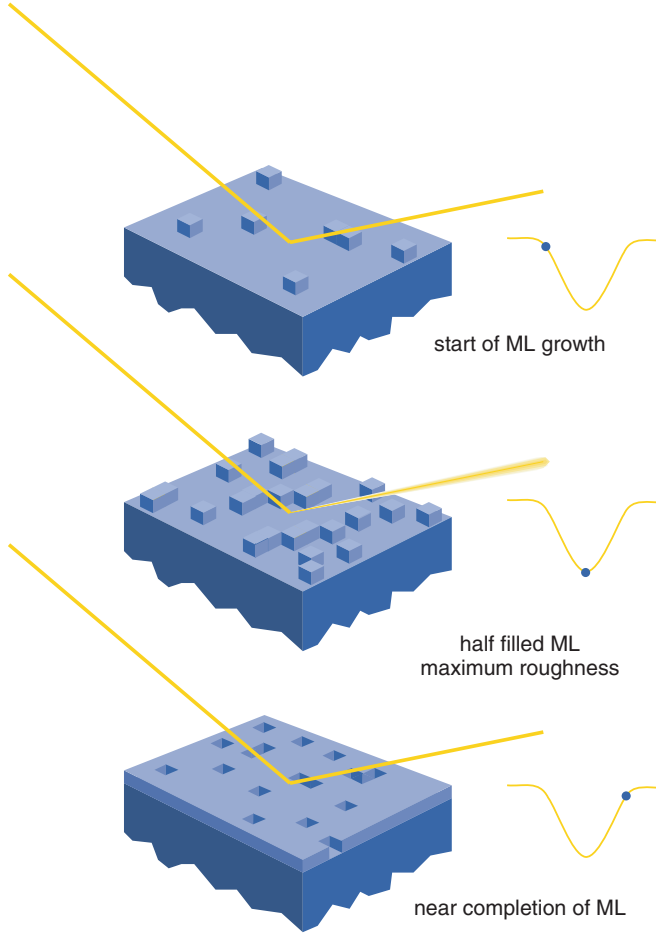

Figure 5.63 *The intensity of the specularly reflected x-ray beam is affected by atomic-scale roughness. Monolayer-for-monolayer growth can therefore be monitored using the specular reflection, whereby each oscillation corresponds to the deposition of a single monolayer.*

will have a large *relative* effect on the signal. So-called roughness oscillations, with a periodicity equal to the time needed to grow a single monolayer, will therefore occur. This phenomenon is also regularly exploited in thin-film growth chambers using laboratory-based reflection high-energy electron-diffraction (RHEED) equipment.

In XRR, a second type of oscillation will also occur, due to the passage of Kiessig fringes across the detector. Let us assume we are at the (0 0 1/2) position of the specular rod. From Bragg's law, this corresponds to an angle

$$\sin\theta_{001/2} = \frac{\lambda}{4d_{001}}. \tag{5.39}$$

But Kiessig fringe maxima occur at this detector angle when

$$m\lambda = 2nd_{001}\sin\theta_{001/2}, \tag{5.40}$$

where m is an integer and nd is the film thickness (n being the number of monolayers). Hence, constructive interference occurs when

$$n = 2m, \tag{5.41}$$

that is, when n is even, while destructive interference occurs for odd n. Note that these thickness oscillations cannot be observed in RHEED, as the high-energy electrons cannot penetrate below the upper surface of the film to the interface. Hence, films which grow in a step-flow mode (i.e. one in which the impinging atoms diffuse to terrace edges rapidly compared to the timescale for the growth of a complete monolayer and therefore exhibit no changes in surface roughness) will show thickness oscillations but no roughness oscillations in XRR, and indeed no oscillations at all in RHEED.

An example of the presence of both types of oscillation is shown in Figure 5.64 for the heteroepitaxial growth of $La_xSr_{1-x}MnO_3$ on $SrTiO_3(001)$ using pulsed laser deposition [44].

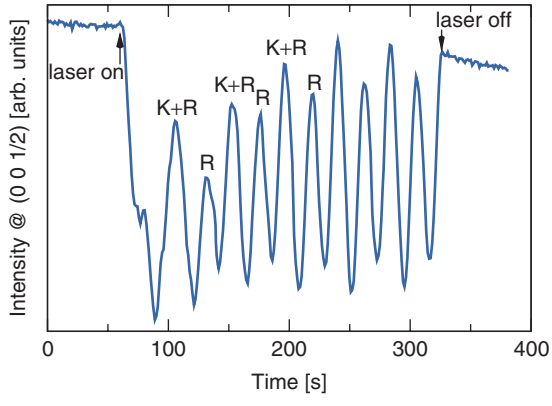

Figure 5.64 *The specular reflectivity intensity at the (0 0 1/2) position during pulsed laser deposition of $La_xSr_{1-x}MnO_3$ on $SrTiO_3(001)$. Note that the roughness oscillations (R) have a frequency twice that of the Kiessig (K) oscillations. Adapted from [44] with permission of the American Physical Society.*

5.16 Small-Angle X-ray Scattering

5.16.1 Introduction

Small-angle x-ray scattering (SAXS) is concerned with determining the general structural features of systems that have typical dimensions between a few nanometres and the micronscale. Such information can be extremely valuable in disciplines such as colloidal science, liquid crystal technology, biochemistry, and cell biology. In particular, it allows one to study systems which lack long-range order, and is therefore complementary to x-ray diffraction [45, 46].

We have seen that in XRD, the scattering angles, given by the Bragg equation, are measured typically in a few tenths of a degree to several tens of degrees. For radiation wavelengths of the order of an Angstrom, and objects of, say, a few to a few hundred nanometres, the scattering angles are typically less than 0.1°. This is the regime of SAXS. Because the scattering angles involved are so small, the distances at the detector of the scattered signal from the direct beam are to a high degree of accuracy directly proportional to the scattering vector, which greatly facilitates analysis.

The scattering characteristics of objects at these larger size-scales are qualitatively different from those produced in x-ray diffraction at atomic scales, which are characterized by high degrees of order. In contrast, on the nano- to micronscale, the structure is generally less well organized and may be composed of complex and nonuniform building blocks. This is especially true when considering biological samples. The resulting scattering patterns in SAXS can therefore more often than not appear to be comparatively nondescript and diffuse. Nonetheless, important information can be extracted from SAXS data, with regards to characteristic size, surface-area-to-volume ratios, and cross-sections. As hard x-rays can penetrate water to significant depths, SAXS is particularly interesting for *in-vivo* and biological samples in ambient liquid environments, or in embedded systems, which are excluded to most other nanoscale imaging techniques, such as scanning electron microscopy.

5.16.2 Theory

Unlike x-ray diffraction, SAXS is purely a contrast technique, whereby the scattering signal is generated by differences in the average electron density.

Let us consider the simplest example of scattering from a homogeneous object (Figure 5.65). Consider two points within the object separated by a distance r. If we consider scattering from all points in the object, it is self-evident that, for large enough angles 2θ, there will be no preferred phase difference between the scattered waves, and their superposition will lead to destructive interference. For small enough angles, however, the phase differences will become smaller and, eventually, below a certain angle, even the largest phase difference originating from scattering from the points furthest away from each other in the object will be less than 2π, and the component waves will begin to reinforce one another until, for forward scattering ($2\theta = 0$), all waves are exactly in phase.

Hence, SAXS signal generally lies within an angular range of the order of $2\theta = \lambda/a$, where a is the characteristic linear dimension of the object. In general, small fluctuations of the electron density within the object over distances much smaller than a will not

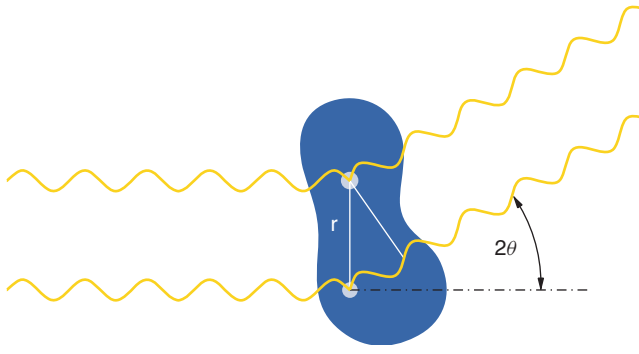

Figure 5.65 *Scattering from a particle. At large scattering angles 2θ, the superposition of scattered waves from all points in the particle will result in destructive interference, as all possible phases will have equal probability. At low enough 2θ-values, however, the phase differences become smaller, resulting in increasingly constructive interference.*

affect the scattering curve and, as a result, one can often assume the object to have a homogeneous density distribution.

5.16.2.1 Far-Field Diffraction and the Q-range

How far must the detector be away from the object in order that the true (square of the) Fourier transform is recorded? Consider Figure 5.66. In the so-called 'Fraunhofer', or 'far-field' regime, the difference δ in the optical path difference (OPD) between (a) two wavelets emanating from opposite sides of the object converging on the detector at a given point P and (b) that between two parallel waves propagating in the same direction, should be much less than the wavelength of light being used.

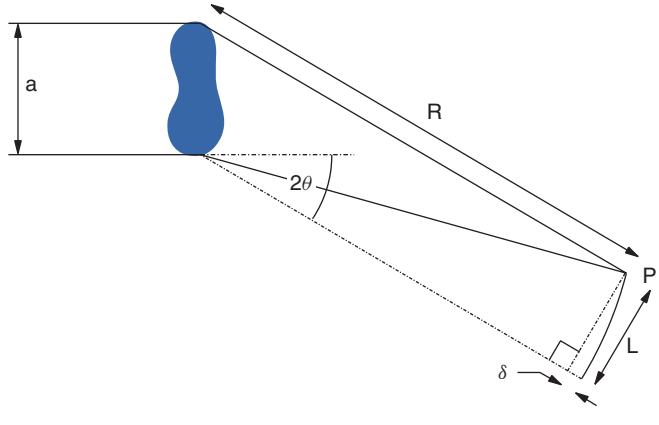

Figure 5.66 *The Fraunhofer condition: the difference δ in the OPD between two converging rays on the detector at point P a distance R from the object of size a and that of two parallel rays should be significantly less than the wavelength of the x-rays being employed.*

We first make the assumption that 2θ is small, and hence that $L \approx a$. We then use Pythagoras' theorem such that

$$(R-\delta)^2 + a^2 = R^2$$
$$\rightsquigarrow \left(1-\frac{\delta}{R}\right)^2 = 1-\frac{a^2}{R^2}.$$

But $\delta/R \ll 1$, and so

$$\left(1-\frac{\delta}{R}\right)^2 \approx 1-\frac{2\delta}{R}$$

and therefore

$$\delta \approx \frac{a^2}{2R}.$$

Because we want $\delta \ll \lambda$, we finally obtain the condition that

$$\frac{a^2}{2R\lambda} \ll 1. \tag{5.42}$$

If we insert typical values of $a = 1\,\mu\text{m}$, $\lambda = 1\,\text{Å}$ into Equation (5.42), we find that the Fraunhofer condition is met at distances even as short as a centimetre. However, typically, object–detector distances of the order of several metres are more common. Why should this be so? The answer lies with detector technology – the further away the detector, the better resolved are elements of the diffraction pattern separated by δQ.

A measurement at a given Q-value allows one to investigate electron-density fluctuations in the sample on a distance scale of $D = 2\pi/Q$. Consider Figure 5.67. At low Q, the observation window is very large, and information about interactions between the particles in the system can be investigated. At intermediate Q the window is of the order of the particle size, while at high Q – in the so-called Porod's region (see below) – contrast occurs primarily at the interface between the particles and their surroundings.

5.16.2.2 Scattering Curves

It is possible to compute the scattering curve for any specific object shape, it being the square of the Fourier transform of the object's electron-density distribution. For

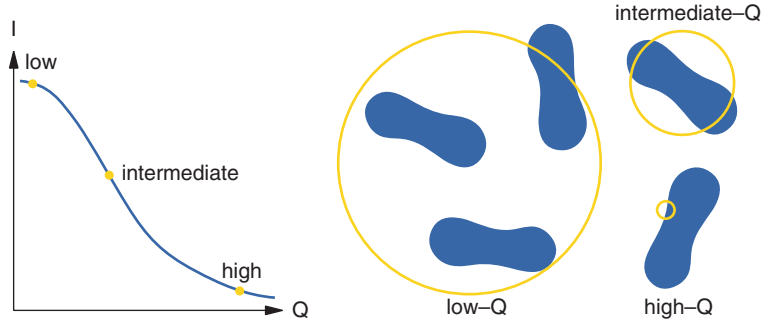

Figure 5.67 *The three Q-domains of SAXS.*

an ensemble of identical, anisotropically shaped, but randomly oriented objects, one must calculate the average scattering signal from all possible orientations. This assumes that the objects are sufficiently separated from one another (i.e. they are 'dilute') that multiple scattering by two or more objects is negligible, and that one can simply add up incoherently the intensities of scattering from individual objects.

Often, the scattering curves for differently shaped objects appear at first sight to be somewhat nondescript and very similar. It is possible, however, to obtain important structural parameters directly from them. We discuss these now in more detail. The scattering functions for the three commonly used shapes of a sphere of radius r, a rod of radius r and length $2h$, and a flat disk of radius r and thickness $2h$ are, respectively

$$I(Q, \text{sphere}) = I_0 \left\{ \frac{3[\sin(Qr) - Qr\ \cos(Qr)]}{(Qr)^3} \right\}^2, \tag{5.43}$$

$$I(Q, \text{rod}) = I_0 \pi\ \exp\left[\frac{-(Qr)^2}{4}\right] (2Qh)^{-1}, \tag{5.44}$$

$$I(Q, \text{disk}) = 2I_0\ \exp\left[\frac{-(Qh)^2}{3}\right] (Qr)^{-2}. \tag{5.45}$$

Note that the scattering function for a collection of spheres oscillates and has its first minimum when $\tan(Qr - 1) = Qr$. This is satisfied for $Qr = 4.493$ radians, which allows one to accurately determine the sphere radius. An example of such a system is given in Figure 5.68 for SiO_2 spheres, which were found to have a diameter of 82.3 nm, from the position of the first minimum in the scattering curve. Only systems which contain

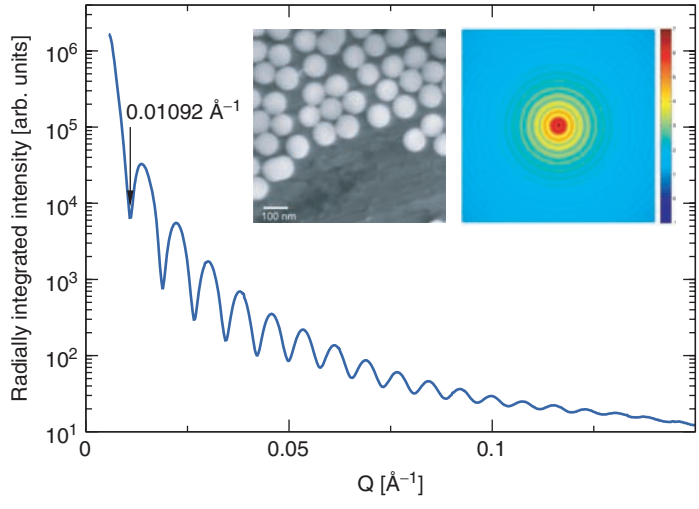

Figure 5.68 *The SAXS signal from an ensemble of SiO_2 spheres, all with the same nominal diameter of 80 nm and shown in the inset scanning electron micrograph, shows clear oscillations. Analysis of the curve produced by radial integration of the image shows that the first minimum lies at 0.01092 Å, corresponding to an average diameter of 82.3 nm. Courtesy Dillip K. Satapathy, Confined Fluids Group, Paul Scherrer Institut.*

nearly spherical particles with a very narrow range of sizes (so-called 'monodisperse') show such oscillations in SAXS.

5.16.2.3 The Radius of Gyration and Guinier's Law

The radius of gyration is identical to the radius of inertia in mechanics, and is defined as the root-mean-square of the distances of all the electrons in the object from their centre of gravity. On a more intuitive level, one can think of it being the radius of a spherical shell of infinitesimal shell thickness and having the same mass (or, more accurately, number of electrons, all packed within the shell thickness) as the object under investigation, and is a measure for its average spatial extent. Indeed, one of the most important radii of gyration used in SAXS is that of a hollow spherical shell, as this can often satisfactorily approximate the shape of biological cells in fluid media. It is given by

$$R^2 = \frac{3}{5}\frac{r_2^5 - r_1^5}{r_2^3 - r_1^3}, \tag{5.46}$$

where R is the radius of gyration, and r_1 and r_2 are the inner and outer shell radii, respectively. We can expand the numerator and denominator such that

$$R^2 = \frac{3}{5}\frac{(r_2 - r_1)(r_2^4 + r_2^3 r_1 + r_2^2 r_1^2 + r_2 r_1^3 + r_1^4)}{(r_2 - r_1)(r_2^2 + r_2 r_1 + r_1^2)}$$

$$= \frac{3}{5}\frac{(r_2^4 + r_2^3 r_1 + r_2^2 r_1^2 + r_2 r_1^3 + r_1^4)}{(r_2^2 + r_2 r_1 + r_1^2)}.$$

If we then set $r_2 = r_1 = r$, as in the case of a hollow shell with infinitesimal shell width, we immediately obtain $R = r$, as asserted above.

Let us now consider the low scattering-vector regime, that is, where $Q \ll 2\pi/R$. In this region, we can only resolve density fluctuations which vary on a scale $D = 2\pi/Q$, which is larger than the characteristic length scale of the particle. Hence, the details of the particles' shape are lost, and we can only obtain information on their overall size, in other words, the radius of gyration. It can be shown that in this region the scattering intensity $I(Q)$, regardless of the particle shape, is accurately given by

$$I(Q) = I(0)\ \exp(-Q^2R^2/3). \tag{5.47}$$

In the low-Q region, a plot of $\ln[I(Q)]$ against Q^2 therefore exhibits a linear descent, directly yielding the radius of gyration. This is Guinier's law. The radii of gyration of some commonly used geometrical shapes to describe systems in SAXS are given in Table 5.1. Note that the equations for the hollow elliptical cylinder and the hollow ellipsoid are the most general, from which the remaining six expressions can be easily derived.

Note that Equations (5.43) to (5.45) all reduce to the same Guinier approximation (Equation (5.47)) for $Qr \ll 1$ and $Qh \ll 1$. Importantly, for the scattering profile of an ensemble of rodlike structures of nearly equal cross-section (but not necessarily length, assuming the lengths of the rods always remain much larger than their diameters), plots of $\ln[Q\,I(Q)]$ against Q^2 yield a gradient of $-r^2$, which is proportional to the cross-section and independent of the rod length (see Equation (5.44)). Similarly, for the case of a flat disk, plotting $\ln[Q^2\,I(Q)]$ against Q^2 yields the disk thickness from the gradient of $-h^2/3$ (see Equation (5.45)).

Table 5.1 *Formulae describing the radius of gyration R for some commonly used homogeneous triaxial bodies as models in SAXS.*

Object	R^2
Solid sphere radius r	$\frac{3}{5}r^2$
Hollow sphere radii r_1 and r_2	$\frac{3}{5}\frac{r_2^5 - r_1^5}{r_2^3 - r_1^3}$
Hollow tube radii r_1, r_2, height $2h$	$\frac{r_1^2 + r_2^2}{2} + \frac{h^2}{3}$
Solid rod radii r, height $2h$	$\frac{r^2}{2} + \frac{h^2}{3}$
Solid ellipsoid semi-axes a, b, c	$\frac{a^2 + b^2 + c^2}{5}$
Hollow ellipsoid outer semi-axes a, b, c, inner semi-axes αa, βb, γc	$\frac{(1-\alpha^3\beta\gamma)a^2 + (1-\alpha\beta^3\gamma)b^2 + (1-\alpha\beta\gamma^3)c^2}{5(1-\alpha\beta\gamma)}$
Solid elliptical cylinder semi-axes a, b, height $2h$	$\frac{a^2 + b^2}{4} + \frac{h^2}{3}$
Hollow elliptical cylinder outer semi-axes a, b, outer height $2h$, inner semi-axes αa, βb, inner height $2\gamma h$	$\frac{3(1-\alpha^3\beta\gamma)a^2 + 3(1-\alpha\beta^3\gamma)b^2 + 4(1-\alpha\beta\gamma^3)h^2}{12(1-\alpha\beta\gamma)}$

5.16.2.4 Porod's Law

The scattering amplitude is proportional to the contrast, or difference in the electron density Δn_e between the scatterer and its surroundings, and therefore the scattering intensity is proportional to $(\Delta n_e)^2$. Additionally, the intensity is proportional to the total number of scatterers in the irradiated volume, N_p, assuming they are separated from one another by more than the coherence length of the x-rays being used. In other words

$$I(Q) \propto N_p\, n_e^2 \tag{5.48}$$

At higher Q-values, toward the tail of the scattering curve, the signal comes from differences in contrast (electron density) over correspondingly small distances, given by a characteristic length $2\pi/Q$ (see the high-Q region shown in Figure 5.67).

For particles within which the average electron density varies significantly less than the difference in electron density as one crosses the particle's surface, scattering will occur primarily at the interface between the particle and the medium in which it is suspended.[17]

[17] For example, the electron densities of water and a typical protein are 0.33 and 0.44 $e^- Å^{-3}$, respectively.

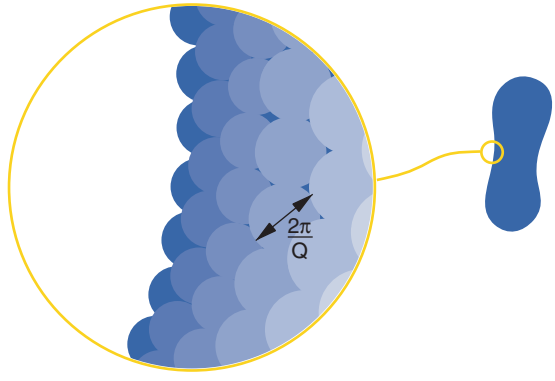

Figure 5.69 *At high Q, the observation window, of size $2\pi/Q$, is much smaller than the scattering particles, and contrast is only obtained across the interface of the particles and the medium in which they reside. Hence, the interface can be thought of as an array of spherical scatterers of diameter $2\pi/Q$. This leads to Porod's law.*

Consider Figure 5.69. We divide up the surface of a particle into spherical elements having this characteristic size. The number N_p of such spheres is equal to the surface area of the particle divided by the cross-sectional area of the scattering spheres, and is therefore proportional to Q^2. On the other hand, the number of electrons per scattering sphere (equivalent to Δn_e above) is proportional to the cube of its characteristic size, or $1/Q^3$. From Equation (5.48), we thus obtain for large Q

$$I(Q) \propto Q^{-4}. \tag{5.49}$$

A more rigorous derivation results in

$$I(Q) = K 2\pi (\Delta\rho)^2 S Q^{-4}, \tag{5.50}$$

where K is a constant and S is the surface area of the particle. This is Porod's law, and can be used to measure the surface area of nanoscale structures.

One can also arrive at the same dependence if one assumes that at sufficiently high Q the probed interface accurately approximates a flat surface. We have already derived the inverse fourth-power dependence of the scattering intensity on the incident angle α for reflectivity from surfaces, given by Equation (5.30). We reasoned that in SAXS α is proportional to Q, hence we obtain again in this manner Porod's law.

5.16.3 Practical Considerations

One of the most stringent practical requirements for the success of SAXS experiments is the minimization of vibrations and long-term drifts. Vibrations on the Hz-scale or higher can originate from passing traffic, air-conditioning fans, mechanical vacuum pumps or humming from high-power electrical components, for example. Long-term drifts occur if the temperature of the experimental hutch is allowed to fluctuate by more than approximately $\pm 1\,^{\circ}$. In general, lateral shifts in the incident-beam–sample should be significantly

less than the sample dimensions, while the detector should remain stable to within less than the linear pixel size, typically measured in microns to a few tens of microns.

The pixel size itself determines the largest scale which can be probed in the SAXS experiment, while the smallest sample dimensions are determined in principle by the size of the detector. In most instances, however, because the scattering intensity drops as $1/Q^4$ (Porod's law), at large Q-values close to the detector edge the signal is so weak that the dynamic range of the detector forbids reliable measurements. The very high brilliance of third-generation facilities has, however, significantly pushed back this limitation.

Air scatter should also be avoided, as this can swamp low-intensity SAXS signal. Modern SAXS stations are equipped with monstrous vacuum flight tubes pumped down to the mbar regime or below (Figure 5.70), in between the sample at one end and the detector at the other.

Lastly, the beam divergence should be compatible with the pixel resolution. This is normally achieved by focusing the beam at the detector and, if necessary, slitting down the beam near the horizontal focusing element.

5.16.4 Worked Example – The Shape of Memory

The flow of calcium (Ca^{2+}) ions across cell membranes control and affect many crucial phenomena such as fertilization, and is also thought to be responsible for the most primitive process in the formation of long-term memory in all animal species [47, 48]. How the cell responds to the change and rate of change of concentration of Ca^{2+} and translates them into a cellular biochemical response is therefore of crucial importance in understanding some of these most basic physiological mechanisms.

It has long been known that a key component in the process of Ca^{2+}-flow is the enzyme known as Ca^{2+}/calmodulin protein kinase (CaMKII) that controls the transfer of phosphate groups. The entire, or *holo*enzyme is a large structure of approximately

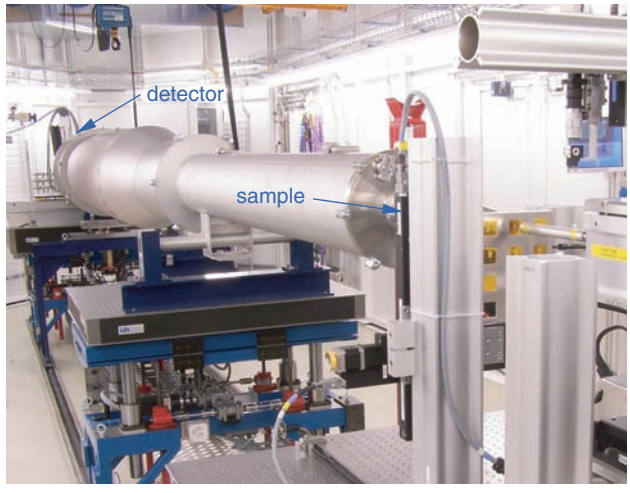

Figure 5.70 *The 7-m flight tube installed at the cSAXS beamline, Swiss Light Source. Courtesy Oliver Bunk and Andreas Menzel, Paul Scherrer Institut.*

600 kDa mass, consisting of 12 kinase domains and a central 'association domain' that links the kinase domains together.[18] The 'glue' that binds the kinase domains to the central association domain are regulatory segments that also inhibit the phosphorylating activity of the kinase.

When Ca^{2+} enters the cell, it binds to a small protein called calmodulin (CaM, short for calcium-modulated protein), which in turn then binds to the regulatory segment and initiates the kinase-domain activity. Importantly, CaM not only transduces the initial presence of Ca^{2+} into phosphorylation by the kinase, but also causes 'autophosphorylation' of the regulatory segment. This 'jams open' the kinase activity, thereby allowing phosphorylation to continue even after the calcium signal finishes.

The detailed structure of the kinase domain and the overall shape of the holoenzyme were investigated using protein crystallography and SAXS, respectively [49]. The results from crystallography show that the kinase domains form dimeric structures which block the kinase activity.

Until this study, two models had been proposed for the structure of the holoenzyme, shown as cartoons in Figure 5.71(a). The mechanism for controlling the response to Ca^{2+} is unknown, but surely intimately depends on the arrangement of the kinase domains in the holoenzyme.

Attempts to determine the full structure of the holoenzyme using crystallography were unsuccessful. SAXS data, however, at $Q \leq 0.04\,\text{Å}^{-1}$ (length scales larger than 150 Å) revealed from Guinier's law a radius of gyration of $R_g = 72\,\text{Å}$. The mass and radii of the kinase rings and central association domain were known, hence it was simple to determine the kinase-ring separation which best suits the experimentally determined

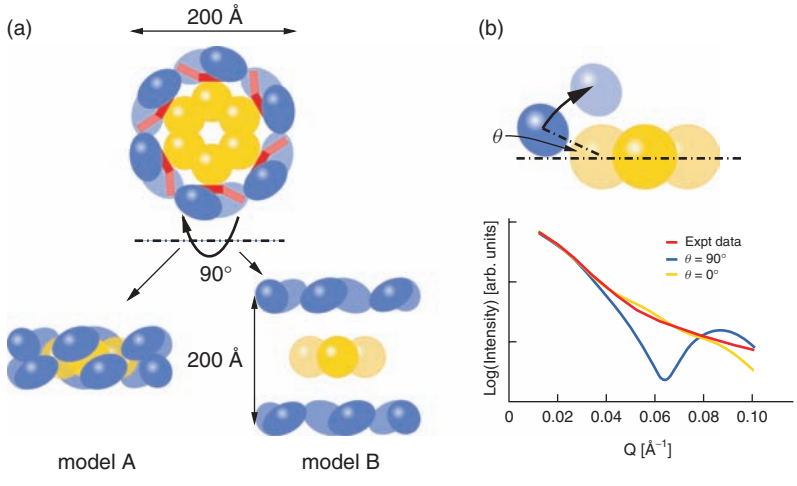

Figure 5.71 *SAXS data of the CaMKII holoenzyme. (a) Cartoon of the two proposed models. The central association domain is shown in yellow, while the kinase domains (blue) are bound to this via regulatory segments (red). The radius of gyration R_g is only consistent with model A. (b) SAXS data at larger Q fit best to a model in which the kinase domains are in the plane of the rings ($\theta = 0$). Adapted from [49] with permission of Elsevier.*

[18] A kinase is an enzyme responsible for the transfer of phosphate groups, a key process in energy transfer.

radius of gyration. This turned out to be 0 Å, that is, the kinase domains are packed closely together, as in model A of Figure 5.71(a). The same radius of gyration can be obtained, however, by rotating the kinase domains above and below the association domain. The SAXS signal was modelled for different rotation angles θ, shown in Figure 5.71(b). The best fit was for $\theta = 0^\circ$, particularly in the mid-Q-range of approximately 0.06 to 0.1 $Å^{-1}$, which probes length scales similar to the holoenzyme's dimensions. Note that for $\theta = 90^\circ$, the holoenzyme would fairly accurately approximate a sphere, resulting in a modulation of the SAXS signal. No such modulation could be observed, however.

Further SAXS data recorded for the holoenzyme in the presence of Ca^{2+}/CaM showed R_g to increase from 72 to 90 Å. This was interpreted as being due to the release of the kinase domains from their centrally docked positions into a much looser structure.

The transition from tight packing of the kinase domains in the absence of Ca^{2+}/CaM to the more open structure when Ca^{2+}/CaM is introduced was consistent with crystallographic and other biophysical measurements.

5.16.5 Grazing-Incidence SAXS

Grazing-incidence SAXS, or GISAXS, is a powerful nondestructive tool to study nanoscale structures on surfaces, interfaces, and thin films [50].

One can think of GISAXS as marrying reflectivity with 'normal' SAXS. By allowing the x-ray beam to impinge on a sample at an angle that is grazing with respect to its average surface (i.e. close to the critical angle for total external reflection), only the surface region is probed, along with any structures that might have formed on it. Because the grazing beam samples a large fraction of the surface, measured in several hundred or even thousands of square microns, averaged statistical information over the sample surface is obtained. Typical applications include investigations of quantum-dot arrays, kinetic or time-resolved studies (such as changes in morphology of nanostructures during thin film growth [51]), as well as the observation of spontaneous formation of self-organized nanostructures in films of polymers. Indeed, as an in-situ technique for a wide range of environmental parameters (including, most importantly, pressure and temperature) it is invaluable, because direct imaging methods such as atomic force microscopy or scanning tunnelling microscopy are essentially impossible to use under these conditions. In addition, GISAXS can be used to investigate buried and internal structures. Increasingly, GISAXS is being applied to catalytic and biological systems, such as proteins attached to the surfaces of lipid layers.

A typical GISAXS setup is shown in Figure 5.72. The component of the scattering vector **Q** in plane (the x-direction in Figure 5.72), is given by

$$Q_x = |k| \sin 2\theta \cos \alpha_f \tag{5.51}$$

and yields information on the widths of the objects in the surface plane, while the out-of-plane component

$$Q_y = |k| \sin \alpha_i + \sin \alpha_f \tag{5.52}$$

does likewise for the object heights. Any correlations in the neighbour-to-neighbour distance D distribution of the nanoparticles will also be highlighted by interference maxima in plane, separated by $2\pi/D$.

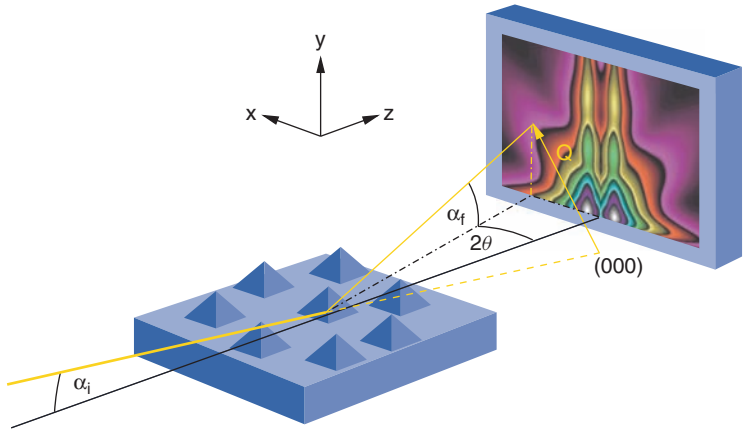

Figure 5.72 *A grazing-incidence x-ray beam is scattered by nanostructures on a surface. The small-angle component is recorded on an area detector. The simulated GISAXS image has been adapted from [51] with permission of AAAS.*

5.17 Concluding Remarks

The periodic nature of crystalline systems lends them physical and electronic properties that can only be understood in terms of that periodicity and the configuration of the atoms within each period. At the same time, this periodicity provides the necessary scattering conditions for diffraction phenomena. X-ray diffraction is a nondestructive structural tool with unsurpassed resolution, capable of determining dimensions with femtometre accuracy.

The primary obstacle to reconstructing structural information from diffraction data is the ubiquitous phase problem. The approach to how this is solved essentially encapsulates the whole art of crystallography. An impressive arsenal of tools has been developed over the last century to crack this problem, and in no more a breathtaking manner than that used in protein crystallography, which nowadays can determine the structure and provide insights into the functionality of biological systems consisting of tens to hundreds of thousands of atoms – the progress from solving the structures of diamond and diopside by the Braggs in the first decades of the twentieth century to determining the form of enzymes, ribosomes, and other biological systems to a resolution of a few Angstroms eighty years later must count as one of the most remarkable in the history of modern science.

Diffraction and elastic-scattering techniques continue to play leading roles in modern synchrotron science. The reason is clear – which other methods are capable of obtaining atomic-scale images of objects consisting of as many as several hundred thousand atoms? Indeed, with the advent of the free-electron laser, scattering methods will undoubtedly enjoy a renewed prominence, particularly in the field of ultrafast time-resolved studies.

References

[1] B. D. Cullity, *Elements of x-ray diffraction*. Addison-Wesley, 1978.

[2] B. E. Warren, *X-ray diffraction*. Dover, 1990.

[3] C. Giacovazzo, ed., *Fundamentals of crystallography*. Oxford Science, 2000.

[4] J. R. Helliwell and P. M. Rentzepis, eds., *Time-resolved diffraction*. Oxford Scientific, Second ed., 2007.

[5] B. Warren and W. L. Bragg, "The structure of diopside, $CaMg(SiO_3)_2$," *Zeit. Kristallogr*., 69, pp. 168–193, 1928.

[6] M. F. Perutz, "How W. L. Bragg invented x-ray analysis," *Acta Crystallogr. A*, 46, pp. 633–643, 1990.

[7] P. Strickland, ed., *International Tables for X-ray Crystallography*. International Union of Crystallography, 2006.

[8] H. M. Rietveld, "A profile refinement method for nuclear and magnetic structures," *J. Appl. Crystallogr*., 2, pp. 65–71, 1969.

[9] S. Marchesini "A unified evaluation of iterative projection algorithms for phase retrieval," *Rev. Sci. Instrum*., 78, p. 049901, 2007.

[10] R. W. Gerchberg and W. O. Saxton, "Practical algorithm for determination of phase from image and diffraction plane pictures," *Optik*, 35, pp. 237–246, 1972.

[11] J. R. Fienup, "Phase retrieval algorithms - a comparison," *Appl. Optics*, 21, pp. 2758–2769, 1982.

[12] V. Elser, "Phase retrieval by iterated projections," *J. Opt. Soc. Am. A*, 20, pp. 40–55, 2003.

[13] K. Moffat, D. Szebenyi, and D. Bilderback, "X-ray Laue diffraction from protein crystals," *Science*, 223, pp. 1423–1425, 1984.

[14] H. Shimizu, M. Iwamoto, T. Konno, A. Nihei, Y. C. Sasaki, and S. Oiki, "Global twisting motion of single molecular KcsA potassium channel upon gating," *Cell*, 132, pp. 67–78, 2008.

[15] M. Müller, B. Murphy, M. Burghammer, C. Riekel, M. Roberts, M. Papiz, D. Clarke, J. Gunneweg, and E. Pantos, "Identification of ancient textile fibres from Khirbet Qumran caves using synchrotron radiation microbeam diffraction," *Spectrochem. Acta B*, 59, pp. 1669–1674, 2004.

[16] M. Müller, B. Murphy, M. Burghammer, I. Snigireva, C. Riekel, J. Gunneweg, and E. Pantos, "Identification of single archaeological textile fibres from the cave of letters using synchrotron radiation microbeam diffraction and microfluorescence," *Appl. Phys. A*, 83, pp. 183–188, 2006.

[17] K. Fadenberger, I. E. Gunduz, C. Tsotsos, M. Kokonou, S. Gravani, S. Brandstetter, A. Bergamaschi, B. Schmitt, P. H. Mayrhofer, C. C. Doumanidis, and C. Rebholz, "In situ observation of rapid reactions in nanoscale Ni-Al multilayer foils using synchrotron radiation," *Appl. Phys. Lett*., 97, p. 144101, 2010.

[18] International Centre for Diffraction Data website. http://www.icdd.com/.

[19] G. Oszlanyi and A. Suto, "Ab initio structure solution by charge flipping," *Acta Crystallogr. A*, 60, pp. 134–141, 2004.

[20] S. J. L. Billinge and M. G. Kanatzidis, "Beyond crystallography: the study of disorder, nanocrystallinity and crystallographically challenged materials with pair distribution functions," *Chem. Comm*., 7, pp. 749–760, 2004.

[21] A. Bhattacharya, "Protein structures: Structures of desire," *Nature*, 459, pp. 24–27, 2009.

[22] J. M. Cork, "The crystal structure of some of the alums," *Philos. Mag*., 4, pp. 688–698, 1927.

[23] M. F. Perutz, M. G. Rossmann, A. F. Cullis, H. Muirhead, G. Will, and A. C. T. North, "Structure of haemoglobin: A three-dimensional Fourier synthesis at 5.5 Å resolution, obtained by x-ray analysis," *Nature*, 185, pp. 416–422, 1960.

[24] A. A. V. Albertini, A. K. Wernimont, T. Muziol, R. B. G. Ravelli, C. R. Clapier, G. Schoehn, W. Weissenhorn, and R. W. H. Ruigrok, "Crystal structure of the rabies virus nucleoprotein-RNA complex," *Science*, 313, pp. 360–363, 2006.

[25] R. B. Von Dreele, "Combined Rietveld and stereochemical restraint refinement of a protein crystal structure," *J. Appl. Crystallogr*., 32, pp. 1084–1089, 1999.

[26] I. Margiolaki and J. P. Wright, "Powder crystallography on macromolecules," *Acta. Crystallogr. A*, 64, pp. 169–180, 2008.

[27] V. Srajer, T.-Y. Teng, T. Ursby, C. Pradervand, Z. Ren, S. Adachi, W. Schildkamp, D. Bourgeois, M. Wulff, and K. Moffat, "Photolysis of the carbon monoxide complex of myoglobin: nanosecond time-resolved crystallography," *Science*, 274, pp. 1726–1729, 1996.

[28] F. Schotte, M. H. Lim, T. A. Jackson, A. V. Smirnov, J. Soman, J. S. Olson, G. N. Phillips Jr., M. Wulff, and P. A. Anfinrud, "Watching a protein as it functions with 150-ps time-resolved x-ray crystallography," *Science*, 300, pp. 1944–1947, 2003.

[29] A. A. Zholents and M. S. Zolotorev, "Femtosecond x-ray pulses of synchrotron radiation," *Phys. Rev. Lett*., 76, pp. 912–915, 1996.

[30] P. Beaud, S. L. Johnson, A. Streun, R. Abela, D. Abramsohn, D. Grolimund, F. Krasniqi, T. Schmidt, V. Schlott, and G. Ingold, "Spatiotemporal stability of a femtosecond hard x-ray undulator source studied by control of coherent optical phonons," *Phys. Rev. Lett*., 99, p. 174801, 2007.

[31] S. L. Johnson, P. Beaud, C. J. Milne, F. S. Krasniqi, E. S. Zijlstra, M. E. Garcia, M. Kaiser, D. Grolimund, R. Abela, and G. Ingold, "Nanoscale depth-resolved coherent femtosecond motion in laser-excited bismuth," *Phys. Rev. Lett*., 100, p. 155501, 2008.

[32] R. Feidenhans'l, "Surface-structure determination by x-ray diffraction," *Surf. Sci. Rep*., 10, pp. 105–188, 1989.

[33] I. K. Robinson and D. J. Tweet, "Surface x-ray-diffraction," *Rep. Prog. Phys*., 55, pp. 599–651, 1992.

[34] D. Martoccia, M. Björck, C. M. Schlepütz, T. Brugger, S. A. Pauli, B. D. Patterson, T. Greber, and P. R. Willmott, "Graphene on Ru(0001) – a corrugated and chiral structure," *New J. Phys*., 12, p. 043028, 2010.

[35] C. M. Schlepütz, M. Björck, E. Koller, S. A. Pauli, D. Martoccia, O. Fischer, and P. R. Willmott, "Structure of ultrathin heteroepitaxial superconducting $YBa_2Cu_3O_{7-x}$ films," *Phys. Rev. B*, 81, p. 174520, 2010.

[36] R. Herger, P. R. Willmott, O. Bunk, C. M. Schlepütz, B. D. Patterson, and B. Delley, "Surface of strontium titanate," *Phys. Rev. Lett*., 98, p. 076102, 2007.

[37] C. Mocuta, J. Stangl, K. Mundboth, T. H. Metzger, G. Bauer, I. A. Vartanyants, M. Schmidbauer, and T. Boeck, "Beyond the ensemble average: X-ray microdiffraction analysis of single SiGe islands," *Phys. Rev. B*, 77, p. 245425, 2008.

[38] Y. Murakami, J. P. Hill, D. Gibbs, M. Blume, I. Koyama, M. Tanaka, H. Kawata, T. Arima, Y. Tokura, K. Hirota, and Y. Endoh, "Resonant x-ray scattering from orbital ordering in $LaMnO_3$," *Phys. Rev. Lett*., 81, pp. 582–585, 1998.

[39] P. D. Hatton, S. B. Wilkins, T. A. W. Beale, T. K. Johal, D. Prabhakaran, and A. T. Boothroyd, "Resonant soft x-ray diffraction – in extremis," *J. Synchrotron Rad*., 12, pp. 434–441, 2005.

[40] Y. Murakami, H. Kawada, H. Kawata, M. Tanaka, T. Arima, Y. Moritomo, and Y. Tokura, "Direct observation of charge and orbital ordering in $La_{0.5}Sr_{1.5}MnO_4$," *Phys. Rev. Lett*., 80, pp. 1932–1935, 1998.

[41] L. G. Parratt, "Surface studies of solids by total reflection of x-rays," *Phys. Rev*., 95, pp. 359–369, 1954.

[42] H. Kiessig, "Untersuchung zur Totalreflexion von Röntgenstrahlen," *Ann. der Phys*., 10, pp. 715–768, 1931.

[43] P. R. Willmott, C. M. Schlepütz, B. D. Patterson, R. Herger, M. Lange, D. Meister, D. Maden, C. Broennimann, E. F. Eikenberry, G. Hülsen, and A. Al-Adwan, "In situ studies of complex PLD-grown films using hard X-ray surface diffraction," *Appl. Surf. Sci*., 247, pp. 188–196, 2005.

[44] P. R. Willmott, R. Herger, C. M. Schlepütz, D. Martoccia, and B. D. Patterson, "Energetic surface smoothing of complex metal-oxide thin films," *Phys. Rev. Lett*., 96, p. 176102, 2006.

[45] O. Glatter and O. Kratky, eds., *Small-angle x-ray scattering*. Academic Press, 1982.

[46] P. Lindner and T. Zemb, eds., *Neutrons, x-rays, and light: scattering methods applied to soft condensed matter*. North-Holland Elsevier, 2002.

[47] M. A. Lynch, "Long-term potentiation and memory," *Physiolog. Rev*., 84, pp. 87–136, 2004.

[48] K. Limbäck-Stokin, E. Korzus, R. Nagaoka-Yasuda, and M. Mayford, "Nuclear calcium/calmodulin regulates memory consolidation," *J. Neurosci*., 24, pp. 10858–10867, 2004.

[49] O. S. Rosenberg, S. Deindl, R.-J. Sung, A. C. Nairn, and J. Kuriyan, "Structure of the autoinhibited kinase domain of CaMKII and SAXS analysis of the holoenzyme," *Cell*, 123, pp. 849–860, 2005.

[50] G. Renaud, R. Lazzari, and F. Leroy, "Probing surface and interface morphology with grazing incidence small angle x-ray scattering," *Surf. Sci. Rep*., 64, pp. 255–380, 2009.

[51] G. Renaud, R. Lazzari, C. Revenant, A. Barbier, M. Noblet, O. Ulrich, F. Leroy, J. Jupille, Y. Borensztein, C. R. Henry, J.-P. Deville, F. Scheurer, J. Mane-Mane, and O. Fruchart, "Real-time monitoring of growing nanoparticles," *Science*, 300, pp. 1416–1419, 2003.

6
Spectroscopic Techniques

6.1 Introduction

In this chapter, we will discuss the experimental methods which probe on the one hand the absorption of ultraviolet and x-ray light by matter, and on the other the subsequent mechanisms, such as fluorescence and photoelectron emission.

Spectroscopic methods in general measure the response of a system as a function of energy. The energy that is scanned may be that of the incident beam (whatever that beam might be – photons, electrons, neutrons, etc.), or the energy of the outgoing particles (for example, photons in x-ray fluorescence, or electrons in x-ray photoelectron spectroscopy). Some techniques also consider the *direction* of the outgoing beam, such as in angle-resolved photoelectron spectroscopy, in which case one is interested in the momentum as well as the energy. Still higher levels of sophistication are obtained by mapping the spectroscopic response of heterogeneous samples as a function of position across them. Examples include scanning transmission x-ray microscopy and x-ray fluorescence mapping.

All the spectroscopic techniques described in this chapter can be divided into three categories. In every case, the incident radiation is synchrotron light, which is absorbed, resulting in excitation of an electron. X-ray absorption spectroscopies (XAS) are concerned with the change in response of a system as a function of the incident photon energy and as such can only be performed at synchrotrons. In general, a core electron is excited to an unfilled valence state. This state can then relax via emission of a photon, or may be radiationless, leading to the ejection of photoelectrons, Auger electrons and a cascade of low-energy secondary electrons. As such, x-ray absorption spectroscopies probe the unoccupied density of states of the system (see Figure 6.1(a)).

In contrast, x-ray photoelectron spectroscopies (XPS) involve the collection and measurement of the energies of directly ejected electrons, yielding information about the core- or valence state from which the electrons originated. Hence, each point in an XAS spectrum can be thought of as being the energy-integrated intensity of an XPS spectrum recorded using the photon energy associated with that point in the XAS spectrum

An Introduction to Synchrotron Radiation: Techniques and Applications, First Edition. Philip Willmott.

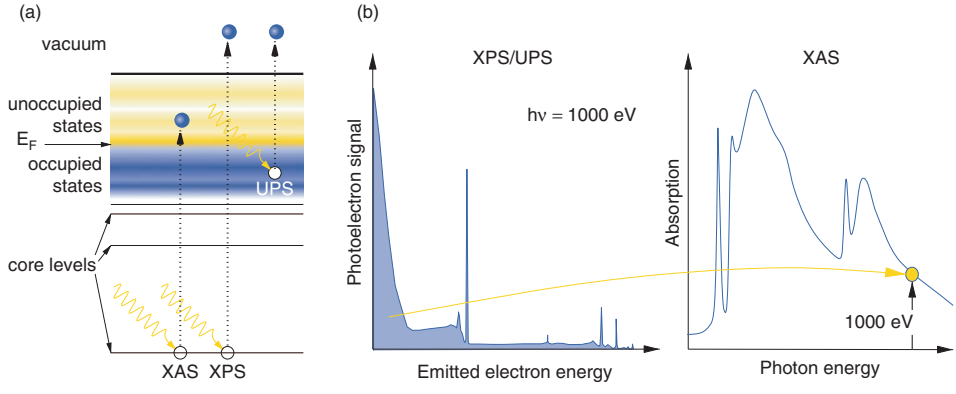

Figure 6.1 *The relationship between absorption and photoelectron spectroscopies. (a) XAS promotes core-level electrons to unoccupied valence-state levels and thereby probes these upper states by varying the incident photon energy. In contrast, in XPS and UPS, the electrons are directly promoted into the vacuum and their signal intensity as a function of emitted electron energy (or binding energy) is recorded for a fixed incident photon energy. (b) Each point in an XAS spectrum can be thought of as the integrated XPS signal for the photon energy of interest.*

(Figure 6.1(b)), that is

$$I_{\mathrm{XAS}} = \int I_{\mathrm{XPS}}(E, h\nu)dE. \tag{6.1}$$

X-ray fluorescence (XRF) on the other hand records the spectral response of the emitted photons produced by relaxation of excited states formed by photoabsorption at a given incident photon energy.[1]

The information obtained by the broad palette of synchrotron-based spectroscopies are to a greater or lesser extent interlinked – because both x-ray fluorescence and photoelectron spectroscopy rely on the absorption of x-rays there is a direct correlation between them and absorption spectroscopy. We therefore begin this chapter with x-ray absorption.

6.2 X-ray Absorption Processes

As we have discussed in Section 2.6.3, the absorbing power of a material is given by the absorption coefficient μ, which describes the exponential drop in intensity of an incident beam passing through a medium. μ depends on the types of atoms constituting the medium, how they are distributed, the nature of their bonding, magnetism, the light polarization and the wavelength of the x-rays. In the simplest case, μ is determined by comparing the intensities of the beam entering and after it emerges from a sample of

[1] Strictly speaking, XRF is also possible and is often carried out using an incident beam of quasimonochromatic electrons in an electron microscope, in which case the technique is called energy-dispersive x-ray spectroscopy, EDAXS.

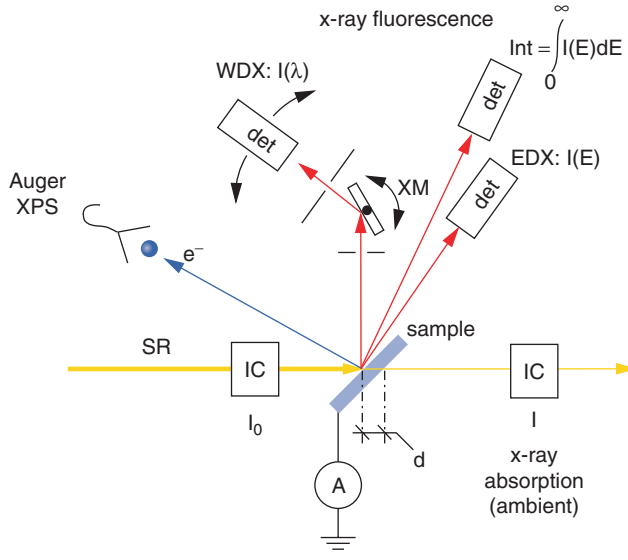

Figure 6.2 *X-ray absorption and x-ray fluorescence experiments. Monochromatic synchrotron radiation (SR) is allowed to impinge on a sample. X-ray absorption spectra can be performed by measuring the amount of light that passes through a thin sample. The x-ray intensities before entering the sample (I_0) and after (I) are measured using ionization chambers (IC) or other beam-intensity monitors, and is particularly suited for samples that cannot be placed in vacuum, such as in biological or catalytic experiments. The total electron current (A) can also be used to indirectly determine the absorption spectrum. In this case, the sample and detectors must be in vacuum. X-ray fluorescence spectra can be recorded, either using a crystal monochromator (XM) in wavelength-dispersive spectra (WDX), or by using a dispersive solid-state semiconductor device (EDX). The integrated fluorescence yield can also be used as a measure of absorption strength. Unwanted elastically scattered x-rays are best suppressed by placing the detector in the plane perpendicular to the polarization of the synchrotron radiation.*

known thickness d (see Figure 6.2), such that

$$\frac{I}{I_0} = e^{-\mu d}. \tag{6.2}$$

This is the Beer–Lambert law for linear absorption.

The relative absorption spectrum can also be indirectly acquired by detecting the yield of secondary electrons. In the case of liquid or solid samples, the measurements must be corrected for reflections from the boundaries of the solid. It is briefly mentioned that reflectivities are normally very low, except for ultraviolet photons up to approximately 100 eV and x-ray photons incident at shallow angles (i.e. near or below the critical angle, see Equation (2.18)).

Conservation of energy demands that the energy of a system after absorbing a photon is increased by the photon energy $h\nu$, that is

$$E_f = E_i + h\nu. \tag{6.3}$$

Secondly, Pauli's exclusion principle requires that the final quantum state E_f was unoccupied before absorption took place. Quantum mechanics shows us that the probability of an optical transition by the absorption of a photon from an initial state (labelled as the *state vector* or *ket* $|\psi_i\rangle$ in the commonly used Dirac formalism) to a final state ($|\psi_f\rangle$) is proportional to the square of the transition dipole matrix element

$$P_{if} \propto \left|\langle\psi_f|er|\psi_i\rangle\right|^2, \tag{6.4}$$

which is shorthand for

$$P_{if} \propto \left|\int \psi_f^* \, er \, \psi_i dV\right|^2, \tag{6.5}$$

where er is the transition dipole operator, ψ^* is the complex conjugate of the electron wavefunction ψ and integration is over all space. When a photon has an energy equal to the difference in energy between two states involved in an allowed electronic transition, it is said to be *resonant* with that transition.

There may be other isoenergetic transitions involving other quantum states, and so the total probability for absorption of a photon of energy $h\nu$ is the sum of the probabilities for all such transitions. Hence the study of the variation of μ with photon energy provides extremely valuable information on the electronic structure of the material and can only be performed using tunable synchrotron radiation.

6.2.1 Energy Level Schemes of Atoms, Molecules and Solids

Now we have discussed the basic ideas behind electronic transitions induced by absorption of a photon (optical transitions), it is instructive to consider the general features of electronic states in different states of matter. We begin with the simplest system, the isolated atom, of which the simplest of these is the hydrogen atom.

The energy required to remove the single electron from the lowest energy level of hydrogen (i.e. the $1s$ state, principal quantum number $n = 1$) is $E_1 = 13.6\,\text{eV}$. Normally, the energy scale is so defined that the electron is said to have zero energy exactly at the ionization threshold, and therefore the discrete energy states below this have negative energies. The energy of a level n in a hydrogen atom is given by

$$E_n = -\frac{R_H}{n^2}, \tag{6.6}$$

where $R_H = 13.6\,\text{eV}$ is the Rydberg constant and is equal to the binding energy of an electron in the $1s$ state. The energy of a photon resonant with a transition from the $1s$-state of hydrogen to an excited state n is therefore

$$R_H \left(1 - \frac{1}{n^2}\right). \tag{6.7}$$

As n becomes large, the energy of the state approaches the ionization threshold asymptotically. States with large n are called Rydberg states (see Figure 6.3(a)) and are closely spaced in energy. Electrons excited to above the ionization threshold (IT) are completely free from the constraints of the system and can therefore assume any kinetic

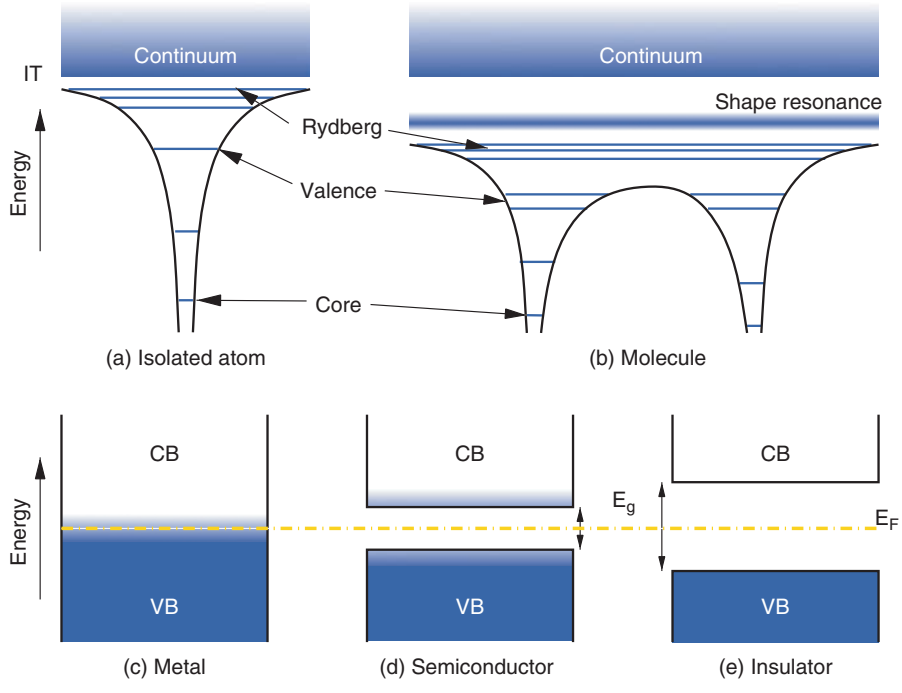

Figure 6.3 *Schematic electron-energy level schemes for different systems. (a) Isolated atoms contain core-level state (most tightly bound) electrons, valence-state electrons, and excited-state electrons, of which those with the highest energy bound levels occupy closely lying Rydberg states. (b) In molecules, the energy of the system is lowered by sharing or exchange of valence electrons. The valence states are therefore generally lower in energy than for the individual atoms. Condensed matter (c)–(e) generally exhibits continuous bands of electron-state* density *due to the interactions between many mutually interacting valence electrons (see text). The boundary between the bound-state valence band (VB) and the unbound-state conduction band (CB) at absolute zero temperature is the Fermi energy, E_F. A metal has no gap between the valence and conduction bands and will electrically conduct at zero temperatures. Semiconductors and insulators exhibit energy gaps (E_g) around the Fermi level in which there are no electronic states. The magnitude of E_g is determined by the strength of the bonding of the valence electrons. The gap in semiconductors is sufficiently small that a small fraction of the valence electrons can be thermally excited to the conduction band. In insulators, E_g is so large that the conduction band remains entirely unoccupied.*

energy (i.e. they are no longer quantized). Above the IT, there is a true continuum of electronic states.

For atoms heavier than hydrogen, the binding energy of the 1s state increases approximately as Z^2, where Z is the nuclear charge, as was shown in Figure 2.2. This is because the Coulombic potential at a given radius is proportional to the product of the two charges involved (i.e. $Ze \times e$), while the radius of the 1s state is proportional to $1/Z$, (the Coulombic potential being itself inversely proportional to the radius). The reason this dependence is only approximate is due to the presence of the other $Z - 1$ electrons

of the atom which will interact with the electron as it is ejected from the atom. Z ranges in magnitude from 1 to 100, hence optical transitions in heavy atoms can result in the absorption of photons with energies as high as approximately 100 keV. For example, the binding energy of the $1s$ (K-state) of uranium ($Z = 92$) is 115.6 keV. Synchrotron sources cover this range of photon energies.

The electronic state of an atom partaking in chemical bonding is different from that of the same atom in an isolated state, due to the redistribution of (primarily) the outermost electrons in the valence or conduction band in chemical bonding (Figure 6.3(b)). Absorption spectra associated with the excitation of a core-level electron (which is normally shielded by outer electrons from the effect of chemical bonding) are only marginally affected by the nature of the chemical bond, as this latter involves the more weakly bound valence electrons. Nonetheless, we will see in Section 6.9 that the small changes in the binding energy (so-called 'chemical shifts') of core-electrons due to differences in the distribution of the valence electrons caused by chemical bonding can, however, provide valuable information via the energy spectra of ejected photoelectrons on the chemical nature of the sample under investigation.

Molecules also show absorption features called 'shape resonances'. Although these are quasi-unbound (and are hence fairly diffuse), they still 'feel' the influence of the electrostatic field of the molecule and are associated with electron states that are physically close to the molecule.

The absorption spectra of condensed matter in the x-ray region associated with the ejection of core electrons exhibit well-defined and (to a greater or lesser extent) sharp absorption edges. However, as we will see in Sections 6.4 and 6.5, subtle (and sometimes not so subtle) differences compared to absorption spectra from the same elements found in isolated atoms and molecules can be evident. These are caused by interactions of the excited electron with other scattering atoms within the local surroundings of the absorbing atom.

The energy-level schemes of condensed matter samples differ most notably from those of isolated systems (atoms and small molecules) with regards to the valence states. In condensed matter, the valence electrons can physically extend over large distances, thereby interacting with other electrons in a complex manner. Consider the interaction of two electrons which, when they are in isolated atoms, have similar binding energies. When these two electrons approach one another, they interact to form two new 'eigenstates',[2] which are a combined configuration of these electrons with a lower energy than the original isolated state, and another with a higher energy. In general, the lower state is associated with a bonding configuration, while the upper state is unbound, or 'antibonding'.

Let us extend this concept to a large number of such electrons (of the order of Avogadro's number, 6×10^{23}), as one might find in a macroscopic solid or liquid. Now, instead of there being one upper and one lower energy state, there are of the order of 10^{23} lower and upper states. Instead of being well-separated as in isolated atoms, these states are extremely densely packed (with separations of the order of 10^{-23} eV) and therefore form quasicontinuous bands of electrons.

[2] 'Eigen' is German for 'actual', or 'real'.

In the case of a metal, the valence band and the conduction band butt up against each other. At absolute zero temperature, the valence band is fully occupied and the conduction band is empty – the boundary between the valence and conduction band is known as the Fermi energy, E_F (Figure 6.3(c)). At finite temperatures, those electrons in the valence band that lie within about kT of E_F can be thermally promoted into the conduction band. kT is approximately 25 meV at room temperature.

Metals are therefore characterized by containing a 'gas' of delocalized and nearly free valence electrons. However, for systems in which the attraction of the valence electron to the ionic core increases, the valence and conduction bands separate in energy, leaving a 'gap' in between, in which there are no allowed states. For small energy gaps (E_g in Figure 6.3) up to a few eV the system is a semiconductor and a small fraction of the valence electrons can be thermally promoted to the conduction band. There is no specific lower limit for the bandgap which thereby makes a material an insulator. Rather an insulator is better defined as a material which melts or decomposes before a detectable fraction of its valence electrons can be promoted into the conduction band.

Remembering Pauli's exclusion principle, we immediately recognize that metals, semiconductors and insulators must have different absorption signatures. Electrons residing in metals can be excited by electromagnetic radiation of any energy and this is the reason why metals reflect light so well in the visible and infrared region. Semiconductors above absolute zero have a certain (though often very small) fraction of electrons thermally promoted to the conduction band, which is why semiconductors do conduct, albeit poorly. Interband transitions are allowed, but require that $h\nu \geq E_g$. In insulators, in contrast, no intraband transitions are possible, while the minimum photon energy required to induce an interband transition is larger than that in semiconductors.

6.2.2 Absorption Features

Absorption spectra have several characteristic features. As mentioned in Chapter 2, μ varies approximately as the inverse third power of the photon energy. This trend is interrupted by steplike increases in absorption as the photon energy matches the ionization potential of an occupied electron state in the atom.

6.2.2.1 Core-Level Absorption Edges

The K-edge results from the increase in absorption due to resonance of the photon energy with the ionization threshold of the K- or $1s$-state (see Figure 2.2), while the L- and M-edges are those associated with ionization of the $n = 2$ and $n = 3$ states, respectively. All absorption edges except the K-edges exhibit three or more energetically closely lying 'sub-edges' associated with nondegenerate quantum states. These arise due to different (quantized) possibilities of coupling the nonzero orbital- and spin angular momentum components of the electronic states (see Section 2.6.3). The reason these are not seen in K-edges is that in this case, the orbital angular momentum is zero, as $l = 0$ (see Figure 6.4) and there can therefore be no coupling.

There can be particularly strong fluctuations in intensity around the absorption edge, known as the x-ray absorption near-edge structure (XANES), discussed in Section 6.4. Starting approximately 50 eV above the edges, the absorption profile exhibits structure

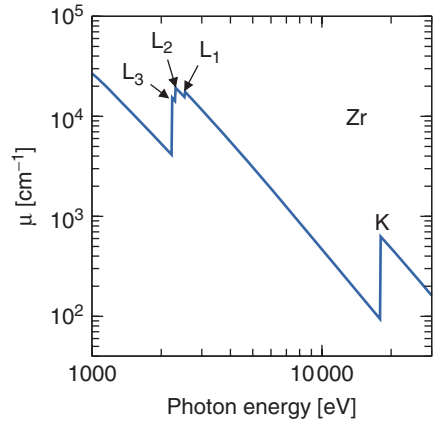

Figure 6.4 *The absorption coefficient of Zr across its K and L-edges. Note that the L-edge has three components, corresponding to ionization from the 2s, $2p_{1/2}$ and $2p_{3/2}$ levels, denoted as L_1, L_2 and L_3, respectively.*

that can extend several hundred eV or more. This 'extended x-ray absorption fine structure' is discussed in detail in Section 6.5.

6.2.2.2 Delayed Onset

The shape of the absorption edges depends on the initial state being excited. As a rule of thumb, those states with a low angular-momentum quantum number l (i.e. s-states with $l = 0$ and p-states with $l = 1$) have sharp rising edges across only a few electron volts. States with larger l values have broader absorption features, with the maximum several tens of eV higher in energy than the onset of the absorption. This so-called 'delayed onset' can be explained qualitatively as follows. Quantum states with high principal quantum number tend to remain away from the nucleus – classically, one would expect them to have a large circular orbit, while s and p orbitals spend some of their time very close to the nucleus.

Referring back to Equation (6.5), it can be seen that the maximum probability for absorption occurs if the integrand of the transition matrix element is as symmetric as possible, as we integrate over all space. The dipole transition operator r is, however, antisymmetric, so for the integrand in Equation (6.5) to be symmetric, we require the overlap integral $\langle \psi_i | \psi_f \rangle$ to be as antisymmetric as possible, that is, the two involved states should be dissimilar in symmetry.[3] Vacuum states just above the ionization threshold also have circularlike orbitals that keep the electron away from the nucleus, and hence the transition probability between initial states with high n and final vacuum states just above threshold is low. At still higher excess energies, however, the electron behaves more like a plane wave and the transition probability therefore increases. An example of this delayed onset is shown for gold in Figure 6.5.

[3] Note, for example, that because they have the same angular symmetry, the probability of transition from the 1s to the 2s level in the hydrogen atom is zero.

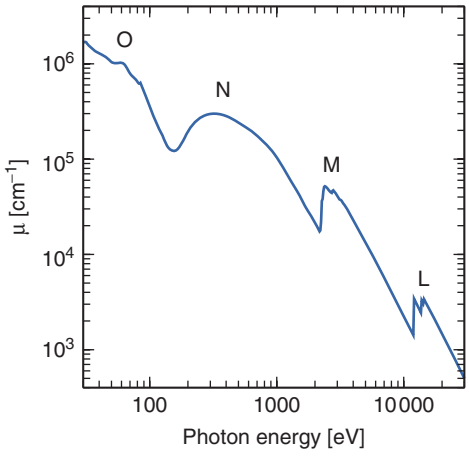

Figure 6.5 *The 'delayed onset' can be clearly seen in the absorption spectrum of gold. The L-edge and, to a lesser extent, the M-edge are sharp, while those absorption edges at lower excitation energies become broader and less well-defined.*

6.3 Photoelectron Energies, Wavelengths and Absorption Regions

The kinetic energy $\mathcal{E}_e$ of a photoelectron ejected from an atom by absorption of a photon is equal to the photon energy $h\nu$ minus the binding energy of the electron E_B

$$\mathcal{E}_e = h\nu - E_B. \tag{6.8}$$

But the de Broglie relationship describing particle-wave duality states that a particle with momentum p has an associated wavelength $h/p = \hbar k$, where h is Planck's constant and k is the wavevector $2\pi/\lambda_e$.

On the other hand, the kinetic energy and momentum of a nonrelativistic particle of mass m are related by

$$E = \frac{p^2}{2m}. \tag{6.9}$$

Hence the photoelectron's linear momentum $\hbar k_e$ is related to its energy $\mathcal{E}_e$ by

$$\mathcal{E}_e = h\nu - E_B = \frac{\hbar^2 k_e^2}{2m_e}, \tag{6.10}$$

and so

$$\lambda_e = h\,[2m_e(h\nu - E_B)]^{-1/2}, \tag{6.11}$$

or, in practical units

$$\lambda_e[\text{Å}] = \frac{12.28}{\sqrt{\mathcal{E}_e\,[\text{eV}]}}, \tag{6.12}$$

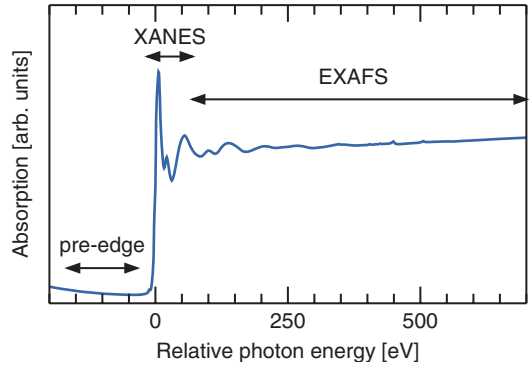

Figure 6.6 *The important regions of a typical x-ray absorption spectrum around a core-electron absorption edge. The energy scale is relative to the centre of the absorption edge. Some features may be found in the pre-edge region due to transitions of core-electrons to empty bound states. In the region around the absorption edge, one obtains x-ray absorption near-edge structure (XANES), dominated by transitions to unoccupied bound states. The absorption spectrum at energies from about 100 eV above the absorption edge often shows a series of oscillations up to as much as 1000 eV. This extended x-ray absorption fine structure (EXAFS) contains information about the local structure around the absorbing atom.*

or

$$k[\text{Å}^{-1}] = 0.512\sqrt{\mathcal{E}_e\,[\text{eV}]}. \tag{6.13}$$

Note that the wavelength of photoelectrons with energies below ~ 100 eV are comparable to or larger than interatomic distances. We will see that this energy provides a natural divide between the two related but distinct methods XANES (also called NEXAFS) and EXAFS, shown in Figure 6.6. A word of caution: the dividing energy between XANES and EXAFS cannot be universally defined, as the transition is gradual and in any case shifts according to the typical nearest-neighbour distances in the system under investigation.

The mean free path of electrons in condensed matter depends strongly on their kinetic energy and only weakly on the nature of the type of matter it is travelling through. The electrons can be both elastically and inelastically scattered by several processes. Particularly when they have energies between 20 and 1000 eV, electrons can be strongly inelastically scattered by excitation of plasmons (concerted movements of bound electrons relative to the nuclear cores). Their mean free path, and therefore their escape depth in this energy region is consequently very small, of the order of a few Å. This trend is more or less the same for all condensed matter, and is given by a 'universal curve' (see Figure 6.7).

We now turn to the lower-energy technique, XANES, then continue with a discussion of EXAFS.

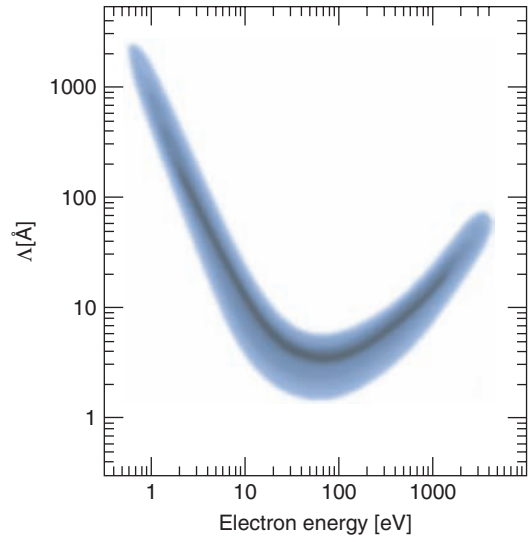

Figure 6.7 *The escape depth Λ of electrons in condensed matter depends strongly on the electron's kinetic energy and relatively weakly on the sort of material through which it is travelling. The blue curve covers the escape depths for the large majority of condensed matter.*

6.4 X-ray Absorption Near-Edge Structure, XANES

6.4.1 Introduction

The absorption cross-section for photoexcitation of a core-electron to vacuum is not a smooth function, but exhibits oscillations in the neighbourhood of the absorption edge (see Figure 6.6).

The features in the immediate vicinity of an absorption edge are caused by excited bound states and also the quasicontinuum just above the ionization threshold, which exhibit large fluctuations in their density of states. This signal is referred to as near-edge x-ray absorption fine structure (NEXAFS) or x-ray absorption near-edge structure (XANES) and images the transition to unoccupied orbitals. As such, it reflects the electronic structure of the material. In principle, NEXAFS and XANES are synonymous and here we will make no distinction (and for consistency's sake, plump for XANES).

For the low photoelectron energies typical of XANES, the mean free path becomes significantly larger than at higher energies (see Figure 6.7). XANES therefore 'sees' a large range of absorber–scatterer distances. This results in multiple scattering of the photoelectron, which means theoretical modelling of XANES still remains a significant challenge. XANES spectra are therefore mostly *interpreted* and not quantitatively analysed, although progress is being made in theoretical modelling and quantitative interpretation [1, 2].

There is an important aspect underlying the detection of a XANES signal. After the absorption of an x-ray photon by a core level, a photoelectron is generated. Importantly, this might not be ejected into the vacuum continuum above the ionization threshold, but might only have enough energy to be promoted to an unoccupied but *bound* level. The core hole resulting from the absorption process can be filled either via an Auger process or by capture of an electron from another shell followed by emission of a fluorescent photon. The difference between XANES and traditional photoemission experiments is that in the latter, the initial photoelectron itself is measured (see Section 6.9), while in XANES the fluorescent photon, Auger electron or inelastically scattered photoelectron is recorded.

This distinction may sound trivial but it is actually significant: in photoemission spectroscopy the final state of the directly emitted electron captured in the detector is an extended, free-electron state. As such, only excitations beyond the ionization threshold are possible. In XANES, one can measure the final state of bound photoelectrons, since the photoelectron itself need not be detected. Hence, the effect of measuring fluorescent photons, Auger electrons and directly emitted electrons is to sum over all possible final states of the photoelectrons, meaning that what XANES measures is the total density of states of the initial core level with all final states, consistent with conservation rules. The distinction is critical, because in spectroscopy final states are more susceptible to many-body effects than initial states, meaning that XANES spectra are more easily calculable than photoemission spectra.

6.4.2 The XANES Signal

What is the source for these large fluctuations in absorption intensity in the XANES regime? There are two major factors, namely the intrinsic probability of a transition from one state to another (Equation (6.5)); and the density of states in which the excited state is embedded. These can be encapsulated in the famous Fermi's golden rule[4]

$$P_{if} = \frac{2\pi}{\hbar} \left| \langle f | H' | i \rangle \right|^2 \rho, \tag{6.14}$$

where $\langle f | H' | i \rangle$ is the transition matrix element that we have already met in Equation (6.4), H' is the perturbation Hamiltonian responsible for the transition (which in the simplest case of a dipole transition to a continuum state in the vacuum is the dipole operator $e\mathbf{r}$) and ρ is the density of states in which the final state is embedded. Let us now discuss how these factors affect XANES features.

The lowest-energy XANES signals actually occur *below* the absorption edge, and are referred to as pre-edge transitions. Pre-edge features are strongly influenced by the symmetry of the local environment around the absorbing atom. To understand this, one must first appreciate that orbitals in molecules and solid-state systems can assume the character of a mixture of atomic orbitals and are related to the symmetry of the local environment, such as the simple example of the famous tetragonally symmetric sp^3-orbitals of diamond formed by hybridization of one $2s$- and three $2p$-atomic orbitals.

[4] Although popularized by Fermi, who coined the phrase 'golden rule', most of the theory behind the rule was done by P. A. M. Dirac.

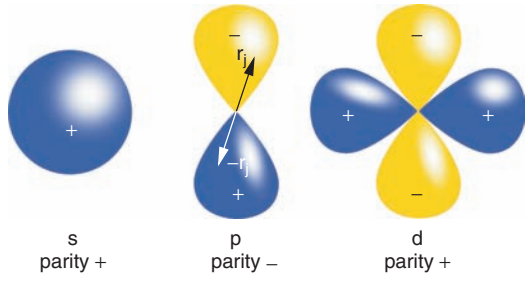

Figure 6.8 *Parity and symmetry. The parity of an electron orbital is either positive or negative, depending on how it is transformed when moving all the elements j of the orbital's amplitude wavefunction from r_j to $-r_j$. So, for example, p-orbitals are antisymmetric with negative parity, while d-orbitals are symmetric and have positive parity.*

Those orbitals around an atom which are allowed depend on the symmetry of the local environment. In general orbitals exhibit a so-called 'parity' of $(-1)^l$. An orbital has positive parity and is symmetric if it is invariant under the operation of moving each volume element of the orbital amplitude wavefunction from $\mathbf{r}$ to $-\mathbf{r}$ – it is said to be inversion symmetric. Hence, the spherically symmetric s-states have positive parity (as do the d-orbitals), while the p- and f-states are antisymmetric and have therefore negative parity (Figure 6.8).

One can also have local environments which are inversion symmetric, such as any of the atom sites in a rocksalt crystal. In such environments, atomic orbitals with different parities cannot hybridize, as the symmetry of such hybridized orbitals does not match that of the environment. So, for example, an sp^3-orbital is incompatible with an inversion-symmetric environment.

Consider as an example a material for which a transition between a K-state ($1s$, $l = 0$, symmetric) and a bound excited state of mostly (symmetric) d-character is observed with high intensity. We have already argued that in the dipole approximation, only p-states (antisymmetric) are accessible from s-states. Hence the upper bound state must have both p-type (antisymmetric) and d-type character. From this we learn that the environment around the absorbing atom cannot be inversion symmetric because of the partial p-character. The presence of bound-state resonant absorption peaks therefore provides important information on the local symmetry of the system. Qualitative information regarding bond lengths can also be easily obtained, as the intensity of pre-edge peaks drops sharply with increasing bond length, due to the exponential envelope decay character of the orbital wavefunctions.

Depending on the strength of interaction with neighbouring atoms and molecules, excited electronic states immediately above the highest occupied state of a system can maintain to a large degree the same structure they have in isolated molecules (see Figure 6.9). This is particularly true for molecules absorbed with low coverage on surfaces or other systems, where 'crosstalk' or 'scrambling' of their individual character caused by interaction with neighbouring atoms or molecules is less pronounced and hence the density of states is sparser. The large majority of XANES experiments are performed on organic compounds, where more often than not, any interaction between the molecule of interest and the 'substrate' on which it is absorbed is small.

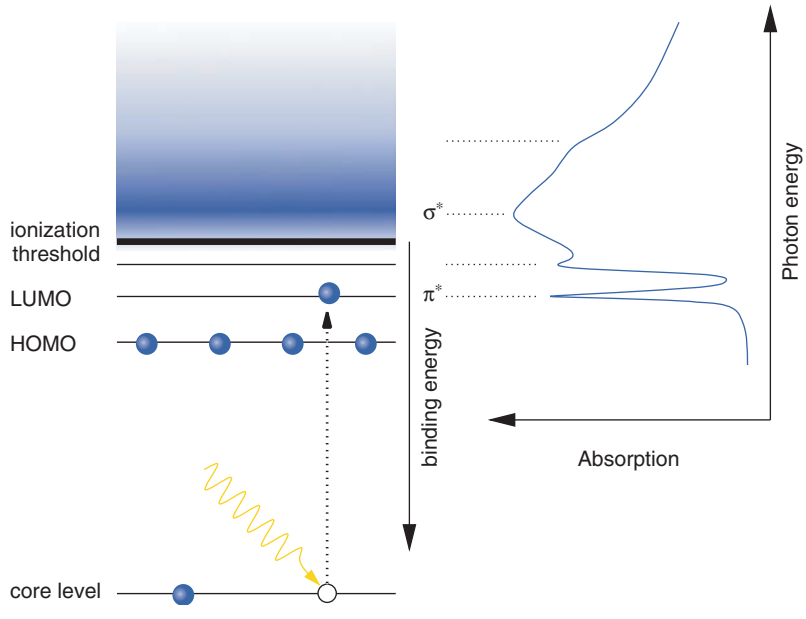

Figure 6.9 *The XANES signal of a core-level electron may consist of sharp bound features corresponding to unoccupied bound states marginally below the ionization threshold (I.T.), of which excitons and white lines are examples, and broader features up to 30 or 40 eV above the I.T. In the case of unsaturated organic compounds (i.e. those with double or triple carbon bonds), the lowest unoccupied molecular orbital (LUMO) immediately above the highest occupied molecular orbital, (HOMO) is invariably a π^* orbital, while the lowest quasi-unbound state above the ionization threshold is σ^*. Their energies in carbon XANES spectra are approximately 286 and 295 eV, respectively, though crucially, these can shift by as much as two or three eV, according to the local chemical environment, a feature exploited to identify organic compounds using XANES.*

The nomenclature used by chemists for the highest occupied state and lowest unoccupied state in isolated molecules is 'highest occupied molecular orbital' (HOMO) and 'lowest unoccupied molecular orbital' (LUMO), respectively. In contrast to 'normal' solid-state orbitals within quasicontinuous energy bands, these orbitals may therefore be fairly discrete for both bound and unbound states. Hence, illumination with photons of the appropriate energy can promote a core-electron to an unoccupied but isolated state. Importantly, because these unoccupied states maintain much of their 'isolated' character, the selection rules governing such discrete-to-discrete transitions are far stricter than normally encountered by solid-state physicists.

The intensity of transitions to bound-state unoccupied orbitals can be high, depending on their density and selection rules. In early experiments in which the spectra were recorded on photographic strips, these bound, excited-state features were easily identifiable as sharp and strong 'white lines', and this is how they were christened. A very illustrative example of how the intensity of white lines can change from material to material is shown in Figure 6.10 for the XANES spectra of the five 5*d*-metals Re, Os,

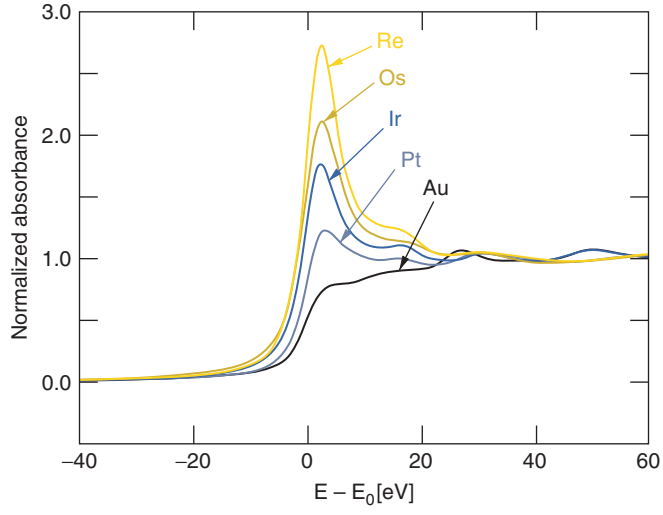

Figure 6.10 *XANES spectra comparing the L_3 absorption edges of five of the 5d-metals, rhenium to gold. Adapted from [3] with permission of the American Chemical Society.*

Ir, Pt and Au [3]. The ground-state electronic configuration of Re is $[\mathrm{Xe}]\,6s^2\,4f^{14}\,5d^5$. As one moves to the right in the periodic table from Re, the remaining five available $5d$-states are filled. In gold, these states are all occupied. In the elements Re to Pt, the bound but unoccupied $5d$-states, which are below the ionization threshold, are seen as intense white lines. As gold has no such unoccupied states, excitation is directly into the unbound continuum.

The probability of a photon being absorbed depends on the orientation of the electric field vector (the polarization, **P**) of the electromagnetic field of the x-rays relative to the polarization direction of the orbital being excited.[5] So, for example, the π-orbital of the formate radical shown in Figure 6.11 lies perpendicular to the molecular plane, and is most efficiently excited if **P** has the same orientation. Because the electric field is always perpendicular to the propagation direction of the x-ray beam, this means that the latter must lie in the molecular plane. The formate radical binds strongly to the surface of Cu(110). The orientation of the molecule was unambiguously determined by the change in the XANES signal with x-ray polarization orientation and incident angle of the x-rays on the surface [4]. Importantly, when the x-rays strike the surface at a glancing angle (20°) and P is parallel to the $\langle 1\bar{1}0\rangle$ crystal axis of the copper, excitation of the π-orbital is suppressed, but is strongly enhanced if the sample is rotated so that P is parallel to the $\langle 001\rangle$ direction. This leads to the conclusion that the formate radical is oriented with the molecular plane perpendicular to the surface and parallel to the $\langle 1\bar{1}0\rangle$ direction (Figure 6.11).

Hence, for systems with well-defined structural orientations (i.e. not polycrystalline material or powders), XANES can provide direct and simply interpreted information on

[5] This dependence of the absorption spectral response on the orientation of the linear polarization is called x-ray linear dichroism (XLD). The observation of XLD generally indicates an asymmetry in the electron charge density.

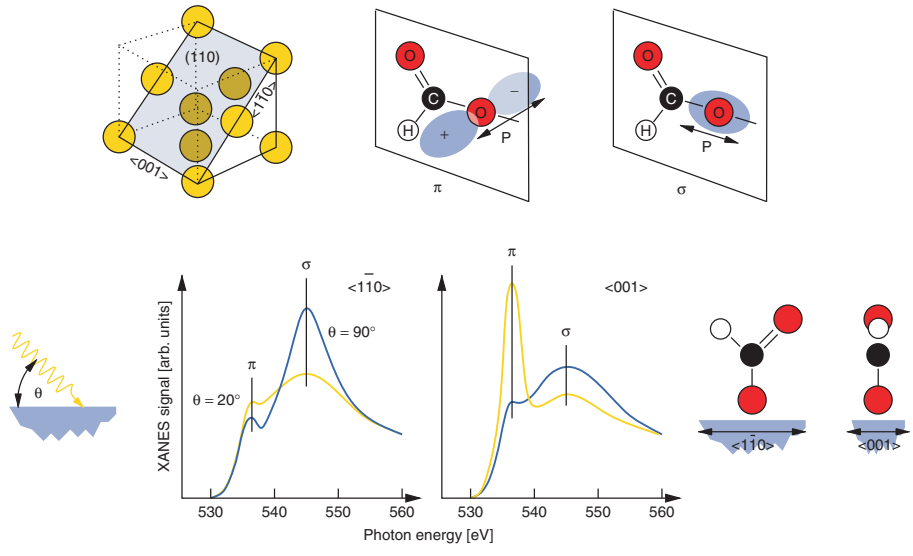

Figure 6.11 *XANES spectra of the formate radical absorbed on the Cu(110) surface change according to the orientation of the x-ray polarization vector (P) relative to the molecule. The yellow curves are for grazing-incidence radiation, the blue curves for normal incidence. From these spectra, the orientation of the radical on the surface could be unambiguously identified. Adapted from [4] with permission of the American Physical Society.*

the configuration via the dependence of the transition probability on the polarization of the exciting photons.

6.4.3 Worked Example – Preservation of the Seventeenth-Century Warship *Vasa*

The seventeenth-century Swedish warship *Vasa* sunk to a depth of about 30 m within the first mile of her maiden voyage in the brackish waters of Stockholm harbour in 1628. The salty and sewage-laden anaerobic conditions of the harbour waters, however, prevented wood-consuming plants, fauna or bacteria, from attacking the ship, thereby preserving the *Vasa* in excellent condition. The *Vasa* was raised in 1961 and, after nearly three decades of preservation treatment, she was put on display in Stockholm in 1990 (see Figure 6.12).

The first signs of problems were detected in the especially wet summer in Stockholm of 2000. Stockholm museum had become unusually humid, due to the high throughput of damp and bedraggled visitors. Small crusty patches of salt were detected on the wooden surfaces of the *Vasa*, accompanied by a softening of the woods within the hold of the vessel. pH measurements below 2 indicated high levels of acidity. Samples of these surface deposits were investigated by x-ray powder diffraction, from which several sorts of hydrated sulfates (i.e. compounds including the ionic group SO_4^{2-}), plus the allotrope of elemental sulfur, brimstone (S_8), were detected. It appeared that these materials were leaching out of solution from the bulk of the wood.

Samples of wood were taken from several positions of the ship and investigated using sulfur *K*-edge XANES (Figure 6.13) [5]. This technique had to be used instead of x-ray

Figure 6.12 *The Swedish warship* Vasa *on display in the Vasa Museum, Stockholm, Sweden.*

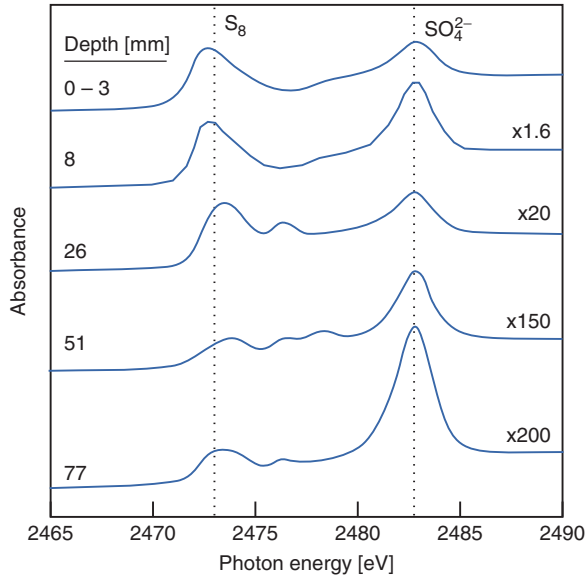

Figure 6.13 *Sulfur K-edge XANES spectra from samples taken at several depths from the woodwork of the* Vasa. *The two major peaks at 2473 and 2483 eV photon energy correspond, respectively, to elemental sulfur and* SO_4^{2-} *(i.e. sulfur in oxidation state +VI). These peaks bracket weaker signals from sulfur compounds in intermediate oxidation states. Adapted from [5] with permission of Macmillan Publishers Ltd.*

diffraction because, on the one hand, the sulfur or its compounds were very poorly crystalline or indeed amorphous or aqueous, and on the other, any XRD patterns would be swamped by spurious signal originating from crystalline debris and the celluloid of the wood. The high chemical specificity of XANES circumvents these problems.

The XANES spectra show several intermediate oxidation states of sulfur, corresponding to the overall oxidation reaction to sulfuric acid

$$S(s) + \frac{3}{2}O_2 + H_2O \rightarrow 2H^+(aq) + SO_4{}^{2-} \tag{6.15}$$

between the signal for the starting elemental sulfur at 2473 eV and the fully oxidized and more tightly bound sulfate at 2483 eV. Depth-profiling of the XANES signal revealed concentrations of elemental sulfur in the near-surface region of the hull woodwork of several mass per cent. It was estimated that the total potential yield of sulfuric acid from this reservoir of brimstone was at least 5000 kg!

What could be the source of this high sulfur content? In native wood, sulfur concentration is measured in parts per million. It is known that water with low concentrations of dissolved oxygen provides favourable conditions for anaerobic bacteria, which produce hydrogen sulfide (H_2S) as a waste product. The H_2S was deposited within the cells of the wood of the *Vasa*, which could then further react (either with water or via sulfide-oxidizing bacteria, though the precise reaction pathway remains unknown) to form elemental sulfur.

Once the *Vasa* was raised, it was immediately exposed to oxygen-rich air, and the above chemical reaction could proceed, which was further catalysed by iron from the rusted bolts of the ship. Damage was not just chemical, but also mechanical – the transformation from elemental sulfur to hydrated sulfate salts is accompanied by a substantial volume increase, which can rupture the cells of the wood.

These findings have had an immediate impact on marine archaeology. First and foremost, a solution to neutralize the oxidation of sulfur that accumulates in the wood of sunken ships must be found, while catalytic agents such as iron should be replaced with inert material. Secondly, there is now a strong current of feeling that no more submerged vessels should be recovered until such a preservation solution is found. For ships that have already been recovered, new conservation guidelines, especially regarding humidity, spray sealing and isolation from iron components, are being established.

6.5 Extended X-ray Absorption Fine Structure, EXAFS

6.5.1 Introduction

Extended x-ray absorption fine structure (EXAFS) signal refers to oscillations in the x-ray absorption coefficient starting at approximately 50 eV above an absorption edge and extending up to several hundred eV. The interpretation and analysis of EXAFS signal was first presented by Sayers, Stern and Lytle in their seminal paper in 1971 [6].

EXAFS partially relies on extensive libraries of reference-compound spectra of well-characterized materials, which provide a comparison for new investigations [7]. Nonetheless, because (in contrast to XANES) multiple scattering is insignificant in

EXAFS, the theory of EXAFS is well described and quantitative comparisons with experimental results can be drawn [6, 8, 9].

The energy of the photoelectrons associated with EXAFS have mean-free paths of the order of a few Angstroms (see Figure 6.7) and hence EXAFS only probes the immediate neighbourhood of the absorbing atom. EXAFS is therefore an important technique in noncrystalline solids and liquids, as it yields information of the short-range structure of materials. It is much used for investigating liquids, clusters, biological materials, low-concentration impurities and catalysts, although it can also be used for well-ordered structures, and is thus complementary to diffraction techniques. The sensitivity of EXAFS is typically below 100 parts per million (ppm) and can be as low as 10 ppm, depending on the element under investigation.

Surface regions, involving perhaps the first five or so monolayers, are exceedingly interesting for materials scientists, as it is normally here that reactions and physical changes take place. Chemi- and physisorption, corrosion, catalytic processes, reconstructions and surface alloying and segregation are typical examples of mechanisms that take place only in the surface region of a condensed-matter sample. However, EXAFS is normally a bulk technique. To obtain surface sensitivity, surface-EXAFS (SEXAFS) uses surface-sensitive detection techniques, the most common of which is Auger electron spectroscopy (AES).

6.5.2 The EXAFS Signal

EXAFS is caused by single backscattering events by nearest-neighbour atoms of the outgoing photoelectron produced through x-ray absorption. The scattered waves interfere with the original photoelectron wave, which results in a change in the electron density at the absorbing atom and consequently leads to a modulation of the absorption strength $\mu(E)$ as the photon energy is scanned (see Figure 6.14).

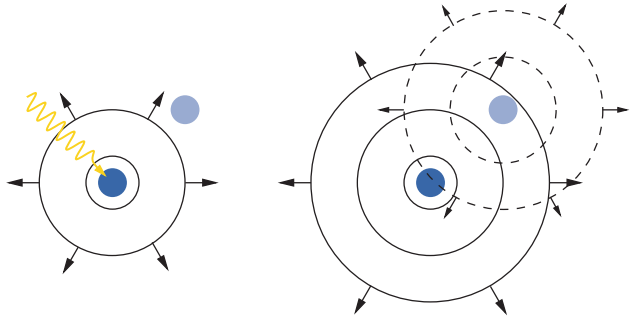

Figure 6.14 *The origin of EXAFS. An x-ray photon is absorbed by an atom, resulting in promotion of a core-level electron to an unoccupied continuum state. As the electron wave propagates out from the excited atom (solid circles), it can be scattered by neighbouring atoms. The scattered waves (dashed circles) interfere with the outgoing waves, thereby modulating the absorption cross-section as a function of photon energy.*

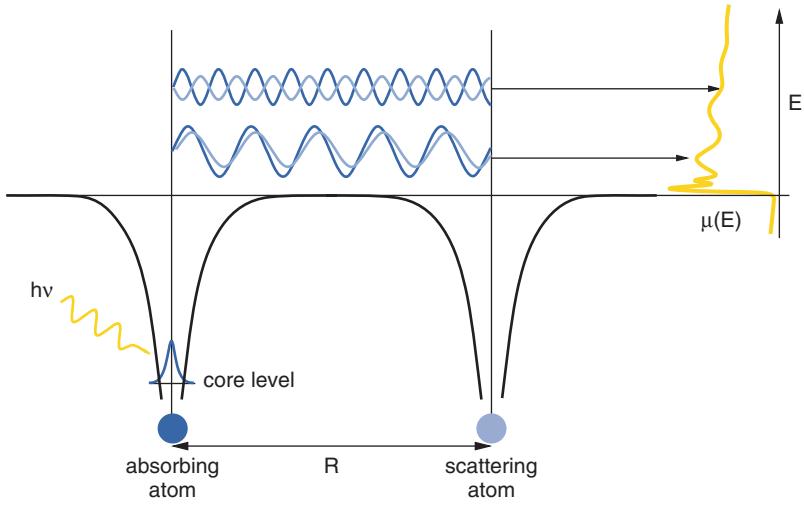

Figure 6.15 *Creation of the EXAFS signal. Photoelectrons promoted to the vacuum continuum by absorption of x-rays can be partially scattered by neighbouring atoms. Depending on the electron wavevector k, interference between the outgoing wave and scattered wave can be constructive or destructive, resulting in a modulation of the absorption spectrum.*

The theory of EXAFS is in principle, straightforward [9]. Here, we cover the basics, explaining the quintessential features. We begin by looking more closely at the interference between the outgoing spherical photoelectron wave and a wave produced by its backscattering by a neighbouring atom at a distance R, as shown schematically in Figure 6.15.

The most important parameters are therefore the material properties (in particular the atomic absorption coefficient μ), the electron wavelength λ_e (Equations (6.11) and (6.12)), the number and type of nearest neighbours and the distance R between the absorbing atom and neighbouring atoms.

Consider first the absorption coefficient of an isolated atom, $\mu_0(k)$, which varies monotonically and smoothly with wavevector (or energy) above the absorption edge. In condensed matter, however, the absorption coefficient $\mu(k)$ is modulated by the so-called EXAFS function $\chi(k)$ because of interference with backscattered waves, and so

$$\mu(k) = \mu_0(k)[1 + \chi(k)], \tag{6.16}$$

which we can rearrange to obtain

$$\chi(k) = \frac{\mu(k) - \mu_0(k)}{\mu_0(k)}. \tag{6.17}$$

$\chi(k)$ is therefore an expression of the effect on an isolated atom's absorption characteristics caused by it no longer being isolated. An illustrative example of the effect of neighbouring atoms is shown in Figure 6.16 for the K-edge of krypton in its gas, liquid and solid phase [10].

Clearly, because the fine structure oscillations are produced by interference between the outgoing and backscattered waves, which is in turn determined by the local atomic structure, we should be able to extract this structural information from the EXAFS signal.

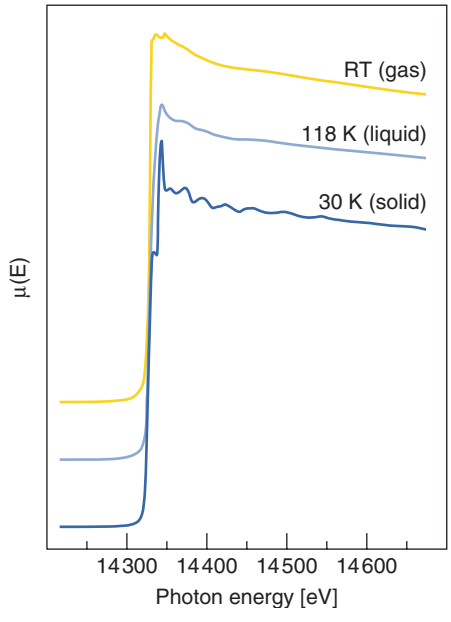

Figure 6.16 *The change in $\mu(E)$ for the K-edge absorption spectra of krypton in its gas, liquid and solid phases. Adapted from [10] with permission of the American Physical Society.*

We now derive an expression for $\chi(k)$, based on the concept of 'shells'. We define a set of concentric shells centred around the absorbing atom, whereby the jth shell passes through N_j identical atoms (see Figure 6.17). First, we should recognize that the EXAFS signal is angular dependent, as formulated by Stern and Heald [11]

$$\chi_j(k,\theta) = \sum_{N_j} 3\langle\cos^2\theta_j\rangle\chi_j^{\text{iso}}(k), \tag{6.18}$$

where θ_j is the angle between the electric field vector and the vector connecting the absorbing atom to the backscattering atom, and $\chi_j^{\text{iso}}(k)$ is the isotropic contribution of the jth shell. In the majority of cases where there is no long-range order and the sample is amorphous or polycrystalline, there is no macroscopic preferred orientation in the sample and the term $\langle\cos^2\theta_j\rangle$ averages to 1/3, in which case, $\chi_j(k)$ reduces to

$$\chi_j(k) = \sum_{N_j} \chi_j^{\text{iso}}(k). \tag{6.19}$$

We now present the standard expression used for $\chi^{\text{iso}}(k)$ and then justify its content. The EXAFS equation is

$$\chi^{\text{iso}}(k) = \sum_j N_j \frac{t_j(k)\sin\left[2kR_j + \delta_j(k)\right]}{R_j^2} \times \exp(-2k^2\sigma_j^2)\,\exp(-2R_j/\Lambda). \tag{6.20}$$

The summation is over all the j considered shells, each shell containing N_j atoms. Shells with radii much larger than a few Angstroms do not normally need to be considered,

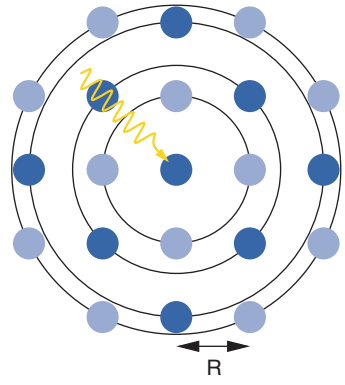

Figure 6.17 *A two-dimensional representation of the first four shells in a regular square array of two different atom types. Here, $N_j = 4, 4, 4,$ and 8; $R_j = R, \sqrt{2}R, 2R,$ and $\sqrt{5}R$, for $j = 1, 2, 3$ and 4, respectively.*

due to the local nature of EXAFS (which is essentially given by the small photoelectron mean-free path and the R_j^{-2} term discussed below).

The factor $t_j(k)$ is the scattering amplitude for the scatterers in the jth shell and has dimensions of area. It depends on the type of backscattering atom and is so small for the first four or five elements that these atoms cannot be detected by EXAFS. t_j increases significantly and exhibits progressively complex oscillatory behaviour as a function of k with increasing atomic number Z [8, 12].

The denominator of R_j^2 takes into account the inverse-square decrease in the electron density of the spherically expanding electron wave from the absorbing atom.

The heart of the equation is the part that describes the oscillations. Let us assume to begin with that there is no phase shift associated with the backscattering of the outwards propagating photoelectron wave. In this case, the backscattered wave interferes with the outgoing wave at the origin constructively if $R = n\lambda_e/4$, where n is an even integer. Conversely, if n is odd, interference is destructive. But $\lambda_e = 2\pi/k(e)$, hence the conditions for constructive and destructive interference can be re-expressed as

$$2kR_j = n\pi \tag{6.21}$$

for n even and odd, respectively. Until now, we have ignored any phase shifts in the backscattering process. Coulomb interactions between the electron and the core of the scattering atom will, however, induce a phase shift $\delta_j(k)$, which must be considered [13, 8]. Hence the oscillatory term is given by

$$\sin\left[2kR_j + \delta_j(k)\right], \tag{6.22}$$

as in Equation (6.20).

The first exponential term in Equation (6.20) accounts for the fact that the neighbouring atoms are not stationary, but vibrating with an amplitude σ_j parallel to $\mathbf{k}$. This is the well-known Debye–Waller factor, and is normally only marginally below unity for samples at room temperature. The final exponential term accounts for inelastic scattering of

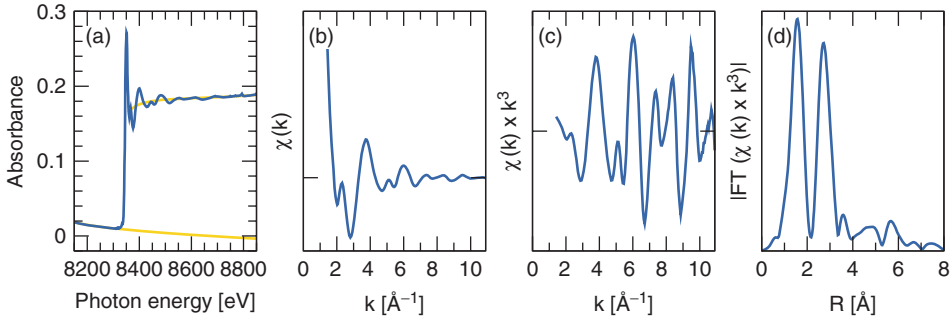

Figure 6.18 *Data analysis of EXAFS spectra. A background using a low-order polynomial is fit to the pre-edge region of the absorbance spectrum $\mu(E)$ in (a), followed by a spline fit to the EXAFS region to obtain $\mu_0(E)$ (both shown here in yellow). The resulting $\chi(k)$ function in (b) is then weighted by multiplication by k^2 or k^3, shown in (c), which is then Fourier-transformed (d) to obtain the interatomic spacings. The example here is for the nickel K-edge for Ni-absorbates in the mineral montmorillonite.*

the electron wave (mainly by plasmons and phonons) on its round trip of $2R_j$, where Λ is the electron's mean-free path length, shown earlier as a function of energy in Figure 6.7.

Data analysis of EXAFS spectra proceeds as follows (Figure 6.18). The oscillatory function $\chi(k)$ is extracted from the raw data $\mu(E)$ by removal of the background signal fitted to the pre-edge signal and the monotonically varying signal $\mu_0(E)$ above the absorption edge. The EXAFS function is then re-expressed as a function of k instead of $\mathcal{E}_e$ using Equation (6.13) and weighted to enhance the weak oscillations at higher k by multiplication of $\chi(k)$ by k^2 or k^3. This function is then Fourier transformed, which converts the data from being expressed in terms of frequencies (proportional to k) to a plot which reveals characteristic lengths, associated with the interatomic distances R_j.

Generally, the phase shifts $\delta_j(k)$ are not incorporated into the Fourier transform, the result of which being that the peaks that emerge tend to be shifted towards lower R-values than the actual bond lengths.

The quality of the EXAFS data and the degree of possible interpretation depend very much on the number of oscillations, which in turn depend on the degree of local ordering and on the absorption strength of the partaking atoms (i.e. $t_j(k)$ and Λ in Equation (6.20)), hence comparison with well-characterized reference samples, preferably of similar chemical nature, can be invaluable in obtaining meaningful information.

6.5.3 Worked Example – Resistance of Lichens to Metallic Pollution

Some lichens have developed the ability to tolerate high concentrations of metals and thereby prosper in heavily contaminated areas unsuitable for other plant species. The biochemical mechanisms responsible for this resilience are largely unknown. In this example, the high sensitivity of EXAFS and its suitability to study pollutant elements in biological systems are highlighted [14]. Metal toxins tend to be bound to functional groups of biological molecules, which are generally uncrystallized or only very poorly crystallized. This makes EXAFS ideally suited to such systems.

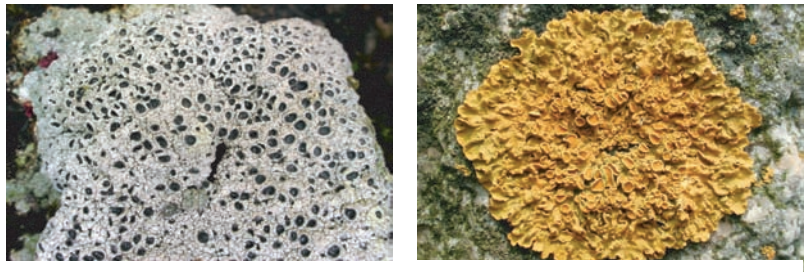

Figure 6.19 *The two lichens* Diploschistes muscorum *and* Xanthoria parietina. *Courtesy Leif Stridvall and Anita Stridvall.*

Two species of lichen were investigated – *Diploschistes muscorum*, harvested in the vicinity of a zinc and lead smelter in Auby, northern France, and *Xanthoria parietina*, sampled on cement poles near a tetraethyl and tetramethyl lead factory near Nantes, France (see Figure 6.19). Chemical tests on *D. muscorum* showed it to contain over 30 g Zn per kilogram and almost 3 g Pb per kilogram. Both these concentrations are approximately three times higher than those recognized as being harmful to 'normal' plants, and as such *D. muscorum* is classified as a 'hyperaccumulator'. *X. parietina* accumulates lead less voraciously, at approximately 0.6 g/kg and is said to be 'tolerant'.

It has been long known that metals collect as inactive complexes of carboxylic groups on the inside fungal walls of *Penicillin chrysogenum*, which has a very similar cell-wall structure to *D. muscorum* and *X. parietina*. It was therefore thought that the two lichens under investigation might use this same mechanism, and so a sample of *P. chrysogenum* was contaminated with Zn and Pb, and used as a reference check.

EXAFS experiments were carried out around the Zn K-edge (9659 eV) for *D. muscorum* and at the Pb L_{III}-edge (13035 eV) for both lichen species. The spectra were recorded in fluorescence or transmission mode, depending on the metal concentration. The results were compared to a relevant selection of lead and zinc compounds which were also recorded, as summarized in Figure 6.20.

From the comparisons, it was possible to conclude that in *D. muscorum*, Pb and Zn are accumulated through an as yet unidentified mechanism in which the synthesis of oxalate ($(C_2O_4)^{2-}$) is enhanced, allowing the precipitation of the toxic metals as insoluble oxalate salts. *X. parietina*, on the other hand, complexed Pb to carboxylic groups of the fungal wall cells in the same manner as does *P. chrysogenum*.

Hence, using EXAFS, two distinct strategies to adapt to these highly adverse environments have been identified. This type of mechanistic information may be particularly important in selecting suitably adapted lichens as biomonitors of contaminated areas, or indeed in their implementation in cleaning the environment in so-called phytoremediation.

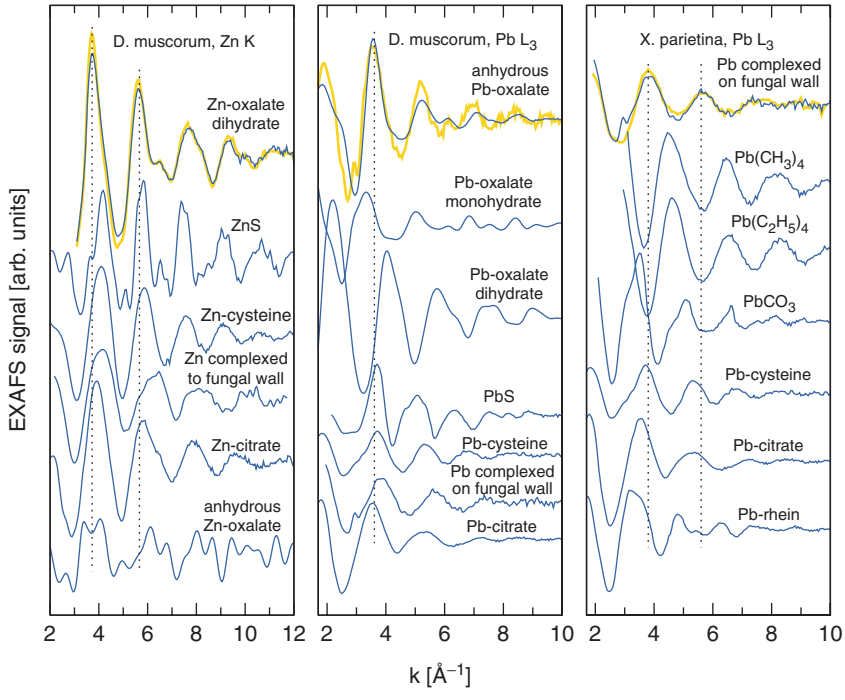

Figure 6.20 *Left panel: EXAFS spectra taken at the Zn K-edge of* D. muscorum *and various other Zn-containing compounds. Middle panel: Comparison of the Pb L_{III}-edge EXAFS spectrum of* D. muscorum *and other Pb-containing compounds. Right panel: Comparison of the Pb L_{III}-edge EXAFS spectrum of* X. parietina *and other Pb-containing compounds. The spectra for the two lichens are shown in yellow, the reference spectra in blue. Adapted from [14] with permission of the American Chemical Society.*

6.5.4 Time-Resolved Absorption Spectroscopy

Conventional EXAFS scans, in which energy ranges of the order of 1000 eV with 1 eV resolution are recorded, require acquisition times of several minutes to a few tens of minutes, depending on the sample type and recording mode. A large fraction of the time is taken up not with data acquisition, but with movements of the monochromator motors between data points in the energy scan, resulting in dead times needed to allow transient vibrations and settling oscillations to decay to an acceptable level.

Chemical kinetics studied using x-ray spectroscopy, especially of catalytic systems, require an improvement in the temporal resolution of at least two orders of magnitude. With this in mind, two distinct approaches to improve the time resolution of XAS have been developed in the last two to three decades, namely energy-dispersive x-ray absorption spectroscopy (EDXAS) and quick-EXAFS (QEXAFS), both briefly described below.

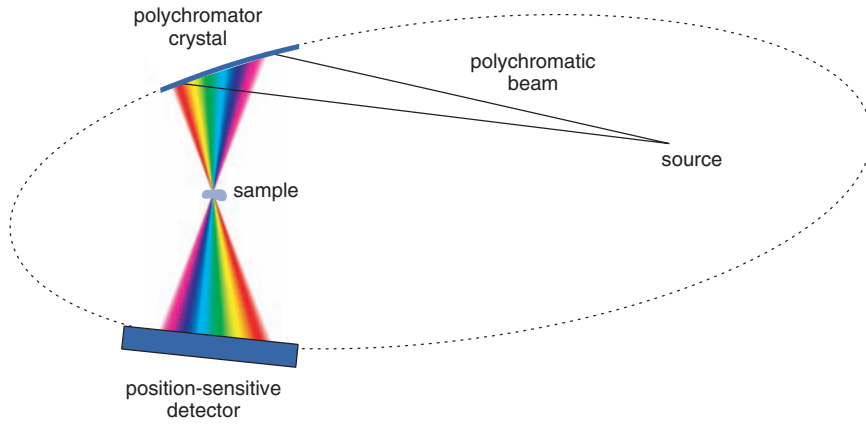

Figure 6.21 *The principle of EDXAS. A polychromatic beam is dispersed into its monochromatic components via a polychromator crystal. The elliptical configuration, in which the source and sample lie at the ellipse's foci and the polychromator crystal on its perimeter, is chosen so that the dispersed x-rays are tightly focused onto the sample and the transmitted signal is spread across a position-sensitive detector.*

6.5.4.1 EDXAS

Energy-dispersive x-ray absorption spectroscopy (EDXAS) is based on the simple idea of spatially dispersing the component parts of a polychromatic beam, then bringing these to a common focus where a sample is placed. The transmitted signal is then recorded using a position-sensitive detector (Figure 6.21). The time-resolution is limited by the readout time of the detector, which for modern devices can be less than a millisecond.

The optics used to disperse the polychromatic incident beam and focus the components onto a common micron-sized spot have developed considerably since the first device was reported in 1981 [15, 16]. In general, they consist of bending a single crystal (commonly made from silicon) into a cylindrical or elliptical shape. This 'polychromator crystal' disperses the incident beam by picking out the different wavelengths according to Bragg's law. Overtones can be suppressed by bouncing the dispersed beam off an x-ray mirror, as discussed in Section 4.3.3.

The ideal focus is achieved using an elliptical curvature of the polychromator crystal and placing the sample and source at the foci of that ellipse (see Figure 6.21). In this manner, submicron spot sizes can be routinely achieved. Because there are no moving parts, the beam stability can be excellent, allowing for spatial mapping. However, even very small distortions of the polychromator crystal due to thermal drifts or other effects will cause the dispersed signal to drift spatially. This sets exceedingly stringent specifications in EDXAS for mechanical and thermal stability. EDXAS is also often used for studies at extreme conditions of pressure using diamond-anvil cells, which have small volumes and limited acceptance angles. A limitation of EDXAS is that it assumes that the sample is homogeneous across the focused spot. If this condition is not met, artefacts will arise in the recorded spectra.

6.5.4.2 QEXAFS

Quick-EXAFS (QEXAFS) was first presented by Ronald Frahm in 1989 [17]. Essentially, the idea is to minimize the number of motor movements, and make these as smooth as possible. In modern QEXAFS stations, a single motor rotates a channel-cut monochromator crystal continuously (often using piezo drives, which are notably free of vibrations), while modern encoder systems are able to monitor the crystal's Bragg angle with sufficient resolution and 'on-the-fly', that is, without stopping the motor. Consider, for example, a Si(111) crystal scanning the energy across the iron K-edge from 7000 to 8000 eV. This requires the Bragg angle to change by marginally more than 2°. The encoder must be therefore able to rapidly read out the angular position with an accuracy measured in thousandths of a degree and gate the detector within the time needed to rotate the crystal through an angle corresponding to the required energy resolution. Within this fairly narrow angular range, the relationship between angular step size and energy step is almost linear.

Modern QEXAFS setups [18, 19] are capable of measuring complete EXAFS spectra within a few tens of milliseconds, meaning that individual data points on any given spectrum can be recorded with up to 100 kHz frequency, depending on the signal-to-noise ratio.

Because in QEXAFS one limits the monochromator movements to a single motor rotation of a channel-cut crystal, the beam will drift vertically during the scan by an amount proportional to the channel height (see Equation (4.11)). Hence the channel-cut crystals used in QEXAFS are normally made to be as compact as possible. The beam drift can be further reduced by using a vertically focusing mirror downstream of the monochromator. If these small movements are unacceptable for certain experiments, a truly fixed-exit monochromator using only one channel-cut crystal is possible, if the two reflecting surfaces are milled to a special shape [20].

An example is shown in Figure 6.22 of the dynamics of redox reactions of ruthenium in alternating oxygen and hydrogen ambient atmospheres, investigated at the SuperXAS beamline of the Swiss Light Source [21].

6.6 Fluorescence Spectroscopies

6.6.1 Introduction

Fluorescence spectroscopies in general are photon-in–photon-out techniques and can therefore be used for both conducting and insulating samples and can often be applied under ambient or indeed *in-vivo* environments, depending on the energy range of interest.

In the following, we will consider three types of experiment, namely 'conventional' x-ray fluorescence (XRF), resonant inelastic x-ray scattering (RIXS) and x-ray standing wave spectroscopy (XSW). RIXS can be considered as a special case of XRF in which the photon has an energy resonant with a transition from a core level to an unoccupied conduction or valence band state instead of having a higher energy which promotes the core-electron to the vacuum.

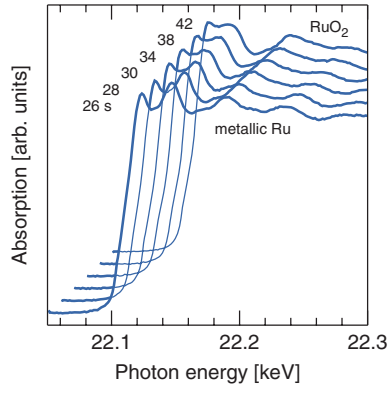

Figure 6.22 *The transformation of metallic ruthenium nanoparticles to partially oxidized ruthenium at approximately 240 °C, as the ambient atmosphere is changed from hydrogen to oxygen. Selected QEXAFS spectra, including their time stamps, are shown over the period of the oxidation transition. The spectral scanning rate is over 3000 eV s^{-1}. The spectra are cascaded for reasons of clarity. Courtesy Christian König and Maarten Nachtegaal, Paul Scherrer Institut.*

XSW, on the other hand, measures the depth-dependent fluorescence[6] from atoms bathed in standing waves produced by interference between incoming and elastically scattered x-ray wavefronts, and as such is a structural technique capable of yielding sub-Angstrom information for two-dimensional systems like crystal surfaces, absorbates or films.

6.6.2 X-ray Fluorescence

X-ray fluorescence (XRF) is a powerful method for determining the chemical composition of objects that are either too thick for transmission-absorption measurements, or are insulating and therefore preclude photoemission techniques. XRF is used widely in geochemistry, archaeology, forensic science and investigations of rare or ancient artefacts. It exploits the 'fingerprint' characteristic radiation of the elements for chemical and elemental analysis.

The basic experimental setup for x-ray fluorescence measurements has been shown in Figure 6.2. Fluorescence can be used as a measure of absorption, as we have already mentioned, if the integrated fluorescence yield is recorded, in other words if the fluorescence signal is not dispersively probed. X-ray fluorescence *spectroscopy* applies to methods which disperse the fluorescence spectrum. We have already discussed in Section 4.6.7 the relative merits of different dispersive photon detector types. In x-ray fluorescence spectromicroscopy, the dispersed XRF signal is mapped out spatially. This is only practical if using an EDX detector (Section 4.6.7), which limits the spectral resolution and the detection sensitivity limit.

XRF signal is strongly absorbed by air for emission below approximately 3000 eV, which corresponds to Ca or thereabouts ($Z = 20$, K emission line at 3692 eV, see also

[6] XSW can be performed with detection via electron emission, although this is less common.

Figure 4.5). For emission signal below this photon energy, the path to the detector should be evacuated to avoid absorption by the residual gas, such that the product of sample–detector distance and pressure is less than approximately 1 Pa m.

6.6.2.1 Worked Example – The Lost Peasant of van Gogh

The paintings of Vincent van Gogh (1853–1890) are among the most influential in the history of modern art. Incredibly, virtually all his works were created in the last ten years of his life; indeed most of the best known were painted in the last *two* years. The last decade of van Gogh's life was overshadowed by poverty and ill-health, accompanied by a rapid and vibrant development in his artistic style. Because he could ill afford to buy new canvases, van Gogh regularly painted over old pictures which he deemed to be less artistically worthy. Consequently, this practice means that there are significant gaps in the record of his progress as an artist.

A typical such example is the 1887 painting 'Patch of Grass', which has been known for some time to have been painted over an earlier composition of a three-quarter profile of a head.[7] The presence of the earlier work was revealed using x-ray transmission radiography and infrared reflectography although, due to the contrast and spatial limitations of these techniques, only very vague facial features could be distinguished, precluding any detailed analysis or artistic interpretation.

In 2008, 'Patch of Grass' was investigated at the microfluorescence beamline L at the HASYLAB in Hamburg, in order to obtain a more detailed image of the hidden portrait (see Figure 6.23) [22]. Microfluorescence was used because of its elemental and chemical specificity. A square region ($17.5 \times 17.5\,\mathrm{cm}^2$) centred around the head was probed using a $0.5 \times 0.5\,\mathrm{mm}^2$ x-ray beam tuned to 38.5 keV. For each position, a fluorescence spectrum was recorded for two seconds, using a high-resolution energy-dispersive Ge-detector, covering an energy range of 3 to 35 keV. Characteristic peaks in the spectra could be associated with easily identifiable elements. As it turned out, the most important of these would be the lead *L*-lines around 15 keV, the antimony $K\alpha_1$ line at 30 keV, and the mercury $L\alpha$ and $L\gamma$ lines at 10 and 14 keV, respectively, as these could be used as tracers.

Notably, the best contrast in the distribution maps were obtained for Hg (especially in the region of the lips and cheeks) and for Sb, which provided the most striking distribution map in which the lighter portions of the skin in the painting are clearly highlighted. The Hg-map could be fairly straightforwardly attributed to the red paint vermillion, which contains mercury sulfide (not used in the predominantly green 'Patch of Grass'). The source for the Sb-map was less clear-cut, as several paints contain antimony. One of the most likely candidates was Naples yellow, which contains the compound $Pb_2Sb_2O_7$, used frequently by van Gogh in his Dutch years. Rather puzzlingly, however, no corresponding map could be established for Pb, which was instead homogeneous and bright.

The reason for this could be explained by investigating the cross-section of a small flake of the painting using a microfocused beam ($1.1 \times 0.3\,\mu\mathrm{m}^2$ spot size) at the ID21 beamline at the ESRF (Figure 6.23(f)). Micro-XRF analysis of the flake showed that its cross-section contained not only the surface paint from 'Patch of Grass' and the original

[7] It is thought, from correspondence between van Gogh and his brother, Theo, that this original portrait was probably painted between 1884 and 1885.

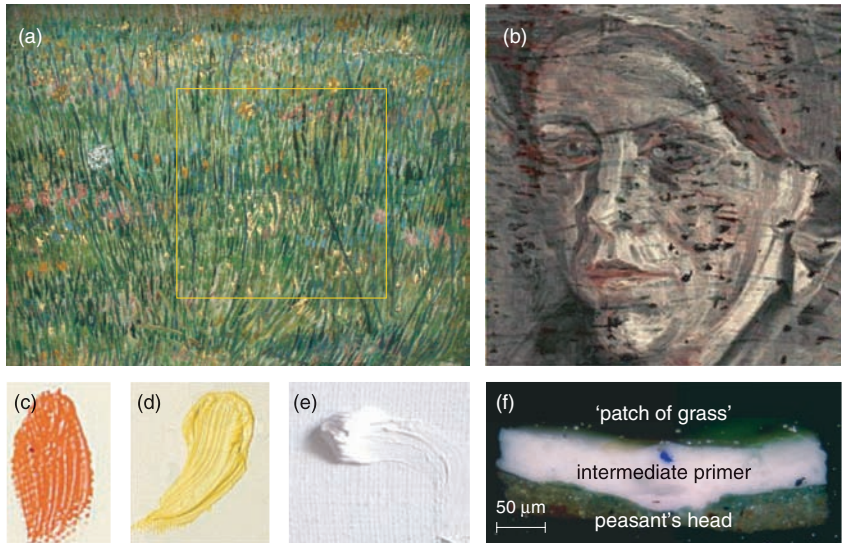

Figure 6.23 *(a) The 1887 floral painting by van Gogh, 'Patch of Grass'. XRF analysis within the area bounded by the yellow box reveals details of a hidden face of a peasant woman from an earlier painting (ca. 1884–1885) when rotated anticlockwise by 90°. (b) A tritonal reconstruction from Hg (from the paint vermillion(c)) and Sb (from the paint Naples yellow (d)) elemental distribution maps of the hidden portrait. (f) The vertical cross-section of a microscopic sample of the painting was investigated using μ-XRF and μ-XANES, revealing an intermediate thick primer layer of lead-white (e). Reprinted from [22] with permission of the American Chemical Society.*

Sb-containing pigment of the hidden portrait, but also a separating layer of about 50 µm thickness, showing strong Pb-fluorescence. It became clear that this is an intermediate primer on top of the hidden face, using 'lead white'. The strength of this signal was so high in the Pb-elemental distribution map that the contribution from the underlying Naples yellow was washed out.

Finally, μ-XANES measurements at the Sb L_3- and L_1-edges were also performed and compared with reference samples of pure Naples yellow, and other Sb-containing pigments. The shape of the Sb-XANES spectrum of the painting and the position of the Sb absorption K-edge, clearly most closely mimic those of Naples yellow.

It is thought that the hidden portrait was one of a series van Gogh painted in 1884 to 1885, using peasant models from the village of Nuenen in the Netherlands. It has been suggested that van Gogh practiced techniques of form, colour and light effects in these works, and as such, they provide invaluable information about his artistic development.

6.6.3 Resonant Inelastic X-ray Scattering

Resonant inelastic x-ray scattering (RIXS) is a rapidly developing experimental method in which photons resonant with electronic transitions are inelastically scattered from matter [23–25]. RIXS has many analogies with Raman spectroscopy in the visible and infrared regime.

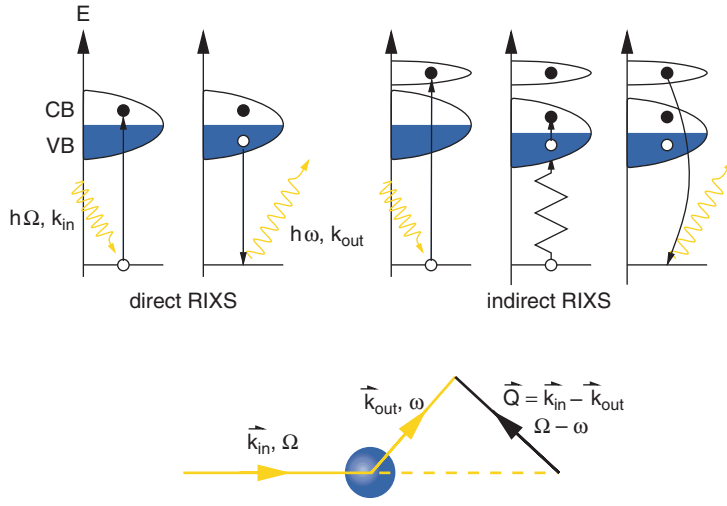

Figure 6.24 *RIXS processes. In direct RIXS, a previously unoccupied conduction-band state is accessed by resonant absorption of a photon, and the core-level hole is subsequently filled by relaxation of an electron from a previously occupied valence-band level, resulting in the production of an electron-hole pair. As this is the only difference between the initial state before excitation and the final state, their momentum and energy differences describe the electron-hole excitation. In indirect RIXS, an electron is promoted to an empty state then relaxes back to its original core state. Importantly, the intermediate state before relaxation can impose a strong electrostatic screening potential, which shakes up electrons in the valence band, thereby exciting an electron-hole pair. In both cases, the energy loss and momentum transfer of the system is transferred to the generated electron-hole pair and is the reason why the valence band is shown to be at lower energies after the initial excitation of the core level.*

One can consider RIXS as being the marriage of conventional x-ray absorption and x-ray fluorescence in a single experiment. RIXS signal can be directly generated following resonant absorption of a photon by an atom promoting an electron to an unoccupied conduction-band state followed by relaxation of a lower-lying state to the core, resulting in the formation of an electron-hole pair. Indirect RIXS signal occurs when absorption and relaxation are between the same states, but when the atom is in the excited intermediate state, a strong electrostatic screening potential around the core hole 'shakes' the atom, resulting in the promotion of an electron from the valence to the conduction band (see Figure 6.24).

Importantly, while XAS and XRF alone yield no information regarding the momenta of the electrons, excluding insights into electronic bandstructure,[8] RIXS does indeed provide this information, at least for wide bandgap materials, and as such is an important adjunct to ARPES, described in Section 6.9.2, as a probe for the electronic bandstructure

[8] The reason why both XAS and XRF cannot probe the electron momenta is that a core hole is involved–in the case of XAS it is in the final state, and in XRF it is in the initial state. This results in a breakdown of the crystal's translational symmetry at exactly the location where the measurement is being made, which has the consequence that k, the momentum, is no longer a good quantum number. In RIXS, neither the initial or final states contain a core vacancy and momentum remains a good quantum number.

of a wide variety of materials. Another important advantage of RIXS over conventional XRF is that it 'bypasses' the intermediate state containing the core hole. As a result, RIXS spectra are not limited in resolution by the natural-linewidth broadening associated with the very short core-hole lifetime. Finally, because excitations are second-order and involve the *virtual* absorption of a photon, certain excitations (such as the $d-d$ excitation between $3d$-states in transition-metal oxides) are allowed in RIXS which are forbidden in direct optical processes.

An important distinction between ARPES and RIXS is that the former can only probe occupied states of a material (by measuring the energy and momentum of an ejected electron formally in an occupied state), while the latter involves normally unoccupied states (as part of the electron-hole pair) and can therefore be used to probe the transitions between occupied and unoccupied bands.[9] This aspect of RIXS was used to advantage as long ago as 2000 in the investigation of the anisotropic nature of the so-called 'Mott gap' in the compound $Ca_2CuO_2Cl_2$ [26].

RIXS is both bulk sensitive and site (i.e. chemically) selective. Because the technique is photon-in–photon-out, one can apply external electric or magnetic fields, experimental 'knobs' excluded to photoemission techniques such as ARPES.

Conservation of momentum and energy (see Figure 6.24) requires that the energy transfer and momentum transferred to the electron-hole pair are given by

$$E = \hbar\Omega - \hbar\omega \tag{6.23}$$

and

$$\mathbf{Q} = \mathbf{k}_{in} - \mathbf{k}_{out}, \tag{6.24}$$

respectively. Note that in the special case of $\omega = \Omega$, there is no net energy loss, and one speaks of 'Rayleigh scattering' or 'resonant elastic x-ray scattering'.

What sort of excitations can RIXS probe? The most modern apparatuses can even access phonon excitations (collective vibrations of the crystal lattice) having energies smaller than 100 meV. One of the most important classes of excitations within the scientific framework of condensed-matter physics, particularly with regards to high-temperature superconductivity, is so-called 'magnons', which describe a collective response of magnetic spins in systems which may show some magnetic order (be it antiferromagnetic, ferromagnetic, or ferrimagnetic). Magnon excitations typically cost between a few meV and a few tenths of an eV, depending on the system. Another very important process in condensed matter, in particular in many metal-oxides, is charge-transfer excitation. These require energies of the order of a few eV.

The RIXS intensity is a second-order process and is in general weak, due to the normally low quantum efficiency of x-ray emission, especially for low-Z elements and for *L*-edge and lower-energy x-ray absorption transitions (see Figure 2.21). One of the main reasons for the increased interest in RIXS in the last decade has been the availability of high-brilliance undulator sources at third-generation synchrotron facilities, allowing one to record the dispersed, weak RIXS signal in reasonable times.

[9] The exception to this is the resonant 'normal' fluorescence part of RIXS, which probes the occupied density of states – the initial state in the photon-emitting relaxation process must be occupied.

A unique feature of RIXS is the very large range of energies and momenta that can be transferred to the system under investigation in the inelastic scattering event. This is because photons instead of electrons (used in photoelectron studies) are used as probes. For a given wavelength λ, or more conveniently, wavenumber k, the ratio of the photon- to electron energy is

$$\begin{aligned}\frac{h\nu}{\mathcal{E}_e} &= \frac{\hbar ck}{\hbar^2 k^2/2m} \\ &= \frac{2mc}{\hbar k} = 517.3/k,\end{aligned} \tag{6.25}$$

whereby k is given in reciprocal Angstroms. So for typical wavevectors of the order of $1\,\text{Å}^{-1}$, the range of energies and scattering vectors $\mathbf{Q}$ available to RIXS surpass other x-ray methods by approximately two orders of magnitude.

The second reason why RIXS has only relatively recently gained so much popularity is due to marked technical improvements in the obtainable resolution. If energy losses well under an electronvolt are to be resolved, one requires a resolving power $E/\Delta E \geq 10\,000$. This has only become a feasible prospect in the first decade of the twenty-first century. The most advanced (soft x-ray) RIXS spectrometer at the time of writing is the SAXES instrument at the ADRESS beamline of the Swiss Light Source, which, depending on the photon energy used, can provide an energy resolution of up to 20 000 [27], an improvement of more than an order of magnitude in less than a decade.

One distinguishes two types of energy-loss features in RIXS, according to whether they vary or not as a function of the incident photon energy $\hbar\Omega$. Features which are independent of $\hbar\Omega$ appear as vertical lines in a two-dimensional plot of energy transfer versus incident energy (Figure 6.25). They represent pair excitations associated with two well-defined states and are therefore products of RIXS processes. Their analogy in the optical regime are Stokes phonons in Raman spectroscopy, from which they obtain their name of Raman peaks. In contrast, those features which exhibit a constant emission photon energy are called x-ray fluorescence- or emission lines.

According to whether one is probing absorption edges in the soft or hard x-ray regime, the equipment used for RIXS differs considerably. Soft x-rays are used to investigate the K-edges of low-Z atoms (up to silicon or thereabouts at 1840 eV,[10]) and the L- and M-edges of the transition metals and heavier elements. Hard x-rays are needed for K-edge RIXS for the elements above sulfur. Soft RIXS instruments use diffraction gratings for both monochromation and signal dispersion, with line spacings around a few thousand per millimetre, operating at grazing incidence [27, 28]. The dispersed signal propagates in ultra-high vacuum to an area detector such as a CCD. The larger the sample–detector distance, the higher the line-density of the grating; and the smaller the pixel size, the higher is the resolution.

Hard RIXS equipment uses crystal monochromators and analysers in much the same manner as that described for WDX experiments in conventional x-ray fluorescence, described in Section 4.6.7 [29].

[10] The intermediate region between soft- and hard x-rays from approximately 2 to 4 keV has recently been coined as the 'tender' x-ray regime ...

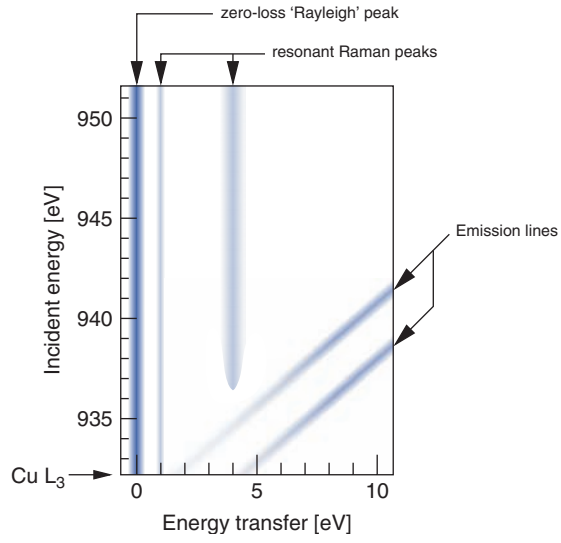

Figure 6.25 *Two-dimensional schematic of a plot showing typical RIXS features. The strong, elastically scattered 'Rayleigh' peak is shown as the vertical line at zero energy transfer. Signals which exhibit constant, nonzero energy transfer are referred to as resonant Raman peaks and are associated with RIXS processes. Features which appear as diagonal stripes occur because the emitted photon has constant energy, and are therefore referred to as emission lines. The strength of these features becomes normally stronger with energy transfer as one moves away from the second-order RIXS process towards the stronger, nonresonant and first-order x-ray fluorescence signal.*

The procedure for RIXS experiments is first to record a conventional x-ray absorption spectrum at the absorption edge of interest. Next, x-ray emission spectra are recorded for different incident photon energies. In order to obtain the dispersion of the RIXS signal as a function of momentum transfer, the sample and detector orientations must be scanned in a concerted manner.

A high-resolution and momentum-dispersed soft RIXS experiment was carried out at the ADRESS beamline of the Swiss Light Source on the compound $Sr_{14}Cu_{24}O_{41}$ [30]. This system is of particular interest, as it exhibits similar behaviour to so-called 'stripe phases' thought to exist in high-temperature superconductors (HTSCs). A deeper understanding of this prototypical system would therefore shed light on the physics of HTSCs.

The compound's structure contains two different types of copper-oxide planes within the plane defined by the a and c crystallographic axes, one of which is shown in Figure 6.26(a). This substructure is a quasi-one-dimensional 'spin ladder'. In the ground state, the 'rungs' of the ladder (highlighted by the ellipse in Figure 6.26(a)) are terminated by copper atoms with opposite spin and the system is antiferromagnetic. However, the system can be excited to a ferromagnetic configuration with parallel spins at the cost of approximately 150 meV. This energy is described by magnetic quanta called 'triplons'.

A single crystal of $Sr_{14}Cu_{24}O_{41}$ was prepared. The material cleaves naturally parallel to the $a-c$ plane shown in Figure 6.26(a). The azimuthal orientation was such that the scattering plane contained the b and c axes (see Figure 6.26(b)). Different components

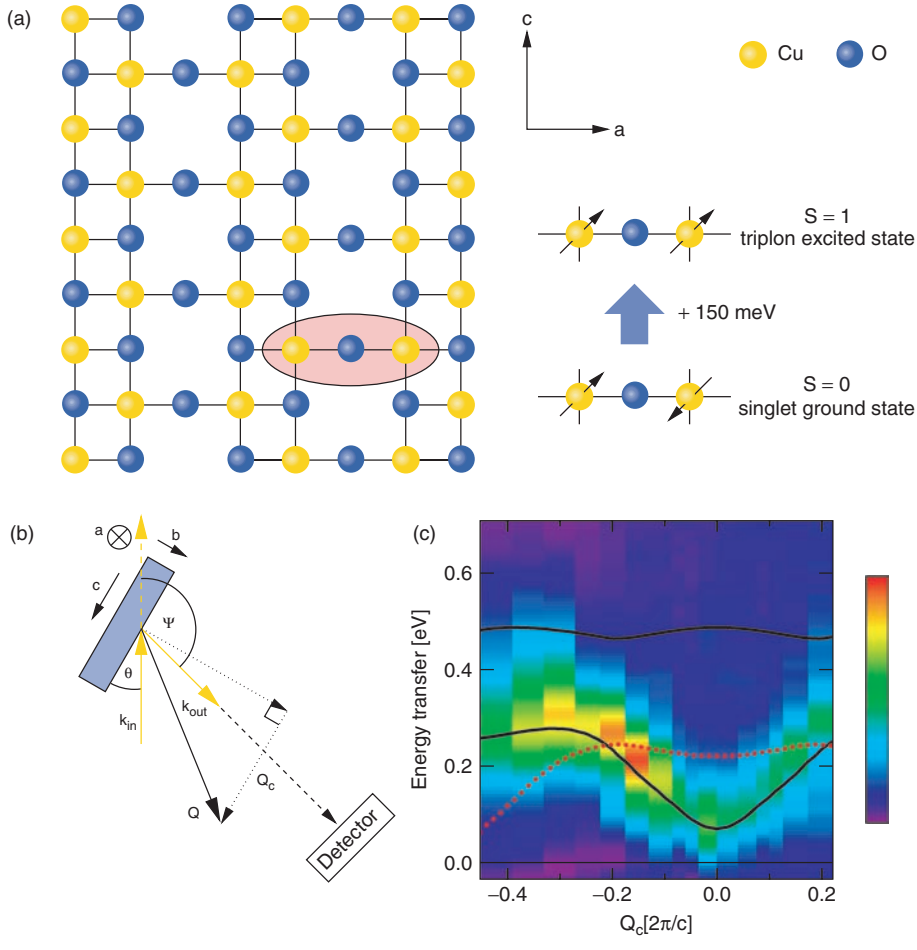

Figure 6.26 *High-resolution and momentum dispersed soft RIXS of the spin ladder compound $Sr_{14}Cu_{24}O_{41}$. (a) The one-dimensional substructure, showing the ladderlike construction of the Cu_2O_3 framework. In its ground state, the copper atoms at the ends of the 'rungs' (highlighted by the ellipse) have opposite spins. An excited triplon state can be accessed with the copper spins parallel. (b) A sketch of the experimental geometry with respect to the sample surface and crystallographic axes. (c) RIXS data shown as an intensity map versus momentum- and energy transfer. The solid and dotted lines are the calculated two-triplon lower boundary (black) and one-triplon dispersion curves (red), respectively. Adapted from [30] with permission of the American Physical Society.*

of the scattering vector **Q** parallel to the c axis could be accessed by fixing the angle between the sample and detector (Ψ) and rotating them together (θ) around the a axis. Different ranges of Q_c were accessed by changing Ψ.

RIXS spectra taken around the Cu L_3 edge with an energy resolution of 120 meV ($E/\Delta E = 7800$) showed a clear maximum centred at approximately 280 meV energy loss, in agreement with previous theoretical studies predicting a two-triplon excitation [31]. The excitation energy was then fixed at 930.9 eV, and the energy loss spectra

were measured for different Q_c-values. The results, after subtraction of the zero-loss Rayleigh peak, are shown in the intensity map of Figure 6.26(c). Comparison of this momentum-resolved dispersion curve with spectral-density calculations (not shown) allowed an unambiguous assignment of the magnetic excitation to a two-triplon channel, while the possibility of there being a dominant one-triplon channel (the red dotted line in Figure 6.26(c)) could be excluded. In addition, the minimum gap for this channel at $Q_c = 0$, recorded when **Q** lies perpendicular to the sample surface, was directly determined to be 100 ± 30 meV. This was a technical breakthrough, as the other primary technique for measuring magnetic excitations, namely inelastic neutron scattering, provides very low signal intensity for small in-plane scattering vectors and is not able to determine this spin gap.

6.6.4 X-ray Standing Waves

The method of structural determination by x-ray standing waves (XSW) depends intimately on details of dynamical diffraction around Bragg peaks [32–36].

Interference between the incident beam and the elastically scattered beam at the Bragg condition produces a standing wave with a periodicity λ_s equal to the interplanar distance d_{hkl} associated with that Bragg peak, that is

$$\lambda_s = d_{hkl} = \lambda/(2\sin\theta). \tag{6.26}$$

Depending on the length scale under investigation, different Bragg maxima can be selected. The smallest length scale for λ_s is given by the energy (wavelength) of the x-ray beam and for $\theta = \pi/2$, in which case $\lambda_{s,min} = \lambda/2$. In diffraction experiments, such a normal geometry cannot be accessed, as this requires that the x-ray detector lies in the path of the incident beam. This poses no problem for XSW, however, as here we are interested in the absorption changes of the standing wave by the atoms in the crystal and not in the properties of the Bragg peak. This begs the question: how do the absorption properties change across a Bragg peak?

According to the dynamical theory of diffraction, the phase of the standing wave produced by interference between the incident and elastically scattered waves changes relative to the scattering centres (the atoms) as one scans across a Bragg peak. Consider Figure 6.27. At the leading edge (lower angles), the phase ν is π radians and the atoms sit in the nodal regions of the standing wave. As one increases the incident angle, the phase of the standing wave shifts vertically until at the opposite edge, the maxima lie exactly above the atoms.

As well as scattering elastically, the atoms of course absorb. Photoabsorption is stronger when the standing wave maxima lie on top of the atoms and the phase is zero. This is the reason why the reflectivity curve is marginally less intense on the high-angle side compared to the low-angle side. The degree of absorption can be measured using fluorescence detectors (such as the lithium-doped semiconductor EDX detectors described in Section 4.6.7) or Auger-electron detectors. The strength of this signal depends also on the choice of photon energy and the absorption edges of the atoms within the crystal.

In the above, we made the simplification that the distance between atomic planes is $\lambda_s = d_{hkl}$ and that there are no other atoms lying between them. This is of course not normally the case – for example, the (200) spacing of GaAs contains two planes of

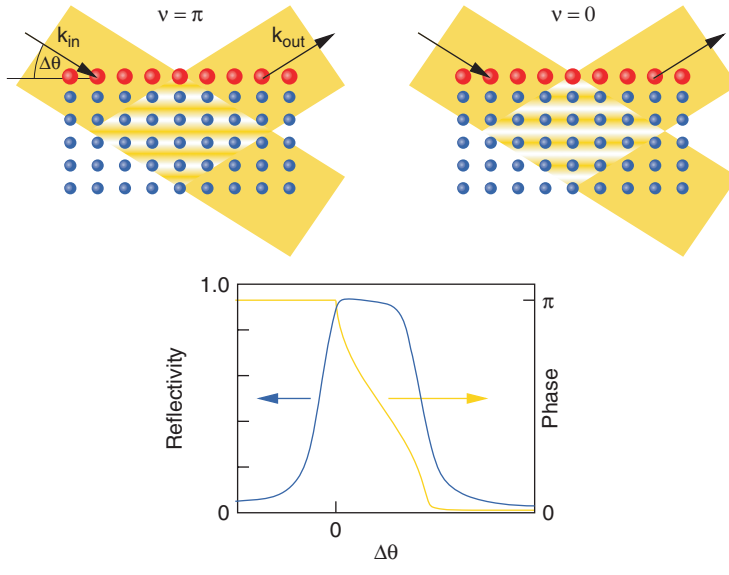

Figure 6.27 *The principle of XSW. As one scans across a Bragg peak, the phase ν between the standing-wave maxima and the scattering atoms flips from being π radians to zero. The sample absorbs differently according to the phase, and can be used to measure the positions of the atoms within the unit cell. Chemical sensitivity is provided by the fluorescence signal, hence the type and positions of absorbates or other surface species (shown here in red) can be determined.*

atoms (see Figure 5.17). So, according to the positions of different atoms within a unit cell, their absorption response (and fluorescence yield) will vary in different manners as one scans the Bragg peak.

Note that, according to the photon energy and the material under investigation, the width of the Bragg peak (the Darwin width) can be very small, of the order of 50 μrad (3×10^{-3} degrees). This sets strict constraints on the accuracy and resolution of the diffractometer, which should be able to move reliably in steps of approximately 10^{-4} degrees. The beam divergence in the scattering plane should also be considerably smaller than the Darwin width, which usually means using beam defining slits and loss of beam intensity.

XSW is also an important technique for investigating the configuration of surface absorbates, particularly if they contain atoms not found elsewhere in the bulk crystal [37].

6.7 Scanning Transmission X-ray Microscopy, STXM

6.7.1 Introduction

Since the introduction of high-brilliance undulator radiation in third-generation synchrotron facilities, x-ray focal spot sizes of the order of a micron or even smaller have become routinely achievable. With this, so-called 'microspectroscopy' methods were

developed, in which the chemistry of heterogeneous samples could be mapped out with high spatial resolution. One of the most important of these to emerge was scanning transmission x-ray microscopy (STXM), particularly in the field of polymer chemistry and physics, where phenomena such as segregation and heterogeneous morphology must be understood at a submicron scale. STXM can be thought of raster-scanning XANES operated in the transmission mode.

6.7.2 The Water Window

There are many problems in biology, organic chemistry and polymer physics which require detailed chemical analysis at a submicron scale but mapped over macroscopic areas. Although traditional methods such as infrared spectroscopy and nuclear magnetic resonance can differentiate chemical species (by observing subtle differences in bond strengths caused by the local chemical environment), their spatial resolution is limited to the millimetre-scale.

STXM is a technique for spatially mapping with submicron resolution the chemical contents of organic samples. This is particularly important for experiments performed within the energy range defined by the so-called 'water window', which lies between the K-edge of carbon at 284 eV and 543 eV, the K-edge of oxygen (Figure 6.28). In this energy range, biological and *in-vivo* specimens absorb approximately an order of magnitude more strongly than water and can therefore reside in an aqueous environment and need not be stained with heavy metals or sliced (both of which are common procedures in the preparation of organic specimens in transmission electron microscopy). Note also that the L-edges of several biologically relevant elements, in particular potassium, sulfur, and calcium, lie within the water window, while phosphorus's L_1-edge is sometimes also accessible at 189 eV. STXM is also frequently used to characterize polymer films

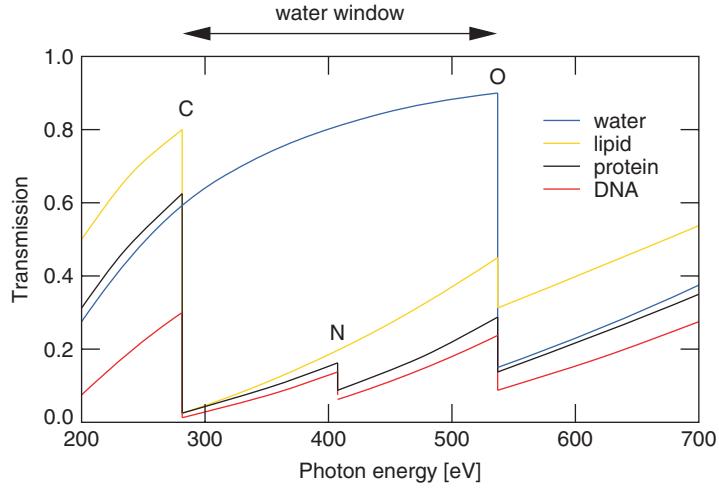

Figure 6.28 *Transmission curves of 1 µm-thick biological materials in the water window between approximately 284 and 540 eV.*

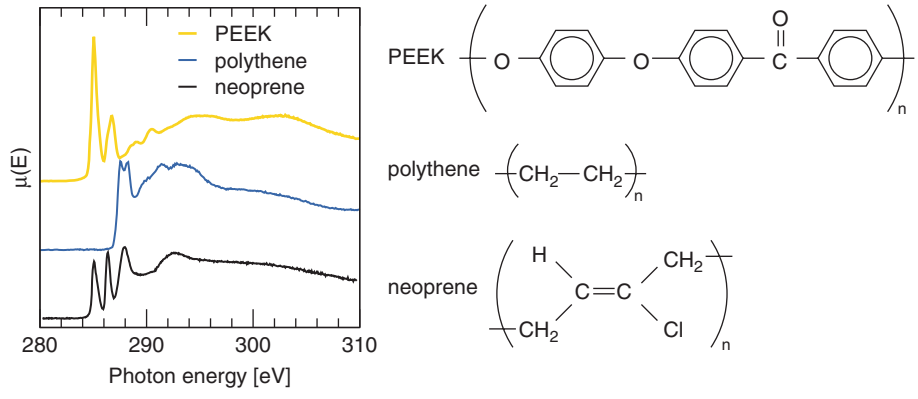

Figure 6.29 *Examples of XANES spectra for three common polymers: the aromatic polymer PEEK, the saturated polymer polythene, and the unsaturated polymer neoprene. The spectra were downloaded from the database described in [38].*

[38], because of the very rich and informative C 1*s* XANES spectra of polymers (see Figure 6.29).

6.7.3 Modes in STXM

A typical STXM setup is shown in Figure 6.30. STXM is performed in one of three modes. In the mapping mode, the sample is rastered across a focused monochromatic x-ray beam. This yields a transmission map of the sample which, depending on the photon energy relative to absorption edges, provides the distribution of a given element or, if the energy resolution is sufficient, even of a chemical group. The rastering resolution and reproducibility should be finer than the size of the x-ray focus, which can set stringent specifications on the translation table and suppression of unwanted sources of vibration (see also Section 4.4).

In the spectrum mode, the sample remains fixed, the photon energy is scanned and the transmitted signal I_{tr} recorded to yield the absorption spectrum

$$\mu(E) = \ln\left(\frac{I_{tr}(E)}{I_0(E)}\right), \tag{6.27}$$

also known as the optical density. Sample drift during acquisition in this mode, which can be significantly larger than the theoretical ultimate spatial resolution, can result in the x-ray spot moving off the region of interest. This problem is resolved in the so-called z-stack mode, in which a sequence of maps at different energies are recorded. These images can then be aligned with one another using reference 'anchor points', either small grains of material that absorb strongly over the entire energy range of interest, or holes where no material exists. Spectrum-mode energy-scans can thereby be retrieved by 'looking down' at a particular coordinate through the aligned stack (Figure 6.31).

Finally, STXM can be operated in a surface-sensitive mode. In the normal mode of operation, the transmitted signal is measured using a photon detector and as such STXM is usually considered to be a bulk technique, though limited to samples thin

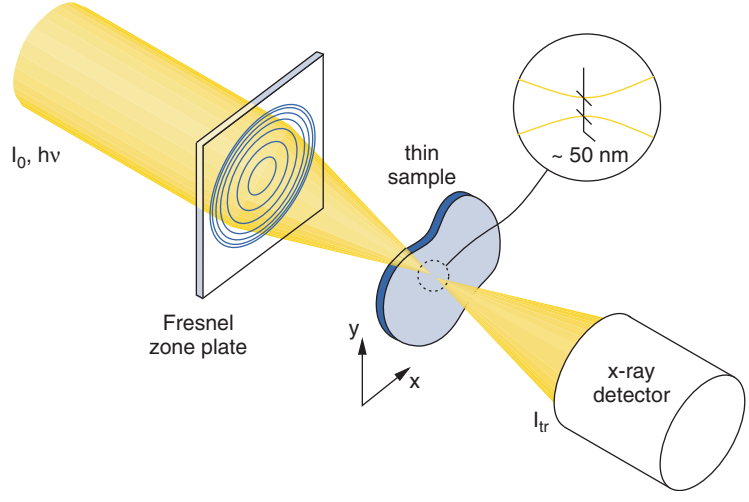

Figure 6.30 *A schematic diagram of a scanning transmission x-ray microscope (STXM). The x-rays are focused to as little as a few tens of nanometres using a Fresnel zone plate. To generate an image, the energy-dispersed transmitted x-ray signal is recorded as the thin specimen is rastered in the x- and y-directions.*

enough to provide a detectable transmission signal. However, one can instead measure the photoelectron yield using detectors such as channeltrons. Typical escape depths of photoelectrons lie in the range of one or two nanometres, and hence STXM operating in this mode becomes surface sensitive [39].

6.7.4 Worked Example – Extraterrestrial Origins of Life

In March 2010, the HiFi infrared instrument aboard the European Space Agency's Herschel telescope sent back spectra from the stellar nursery of the Orion nebula, some 1350 light years away from us. The spectra showed that the nebula is awash with organic compounds and prebiotic matter. This is the first unambiguous confirmation that space is filled with precursors for life.

Back here on Earth, some 40 million kilograms of extraterrestrial matter rains down on the earth each year [40]. This ranges from submicron sized dust particles to catastrophic meteors weighing thousands of kilograms. The mass distribution is sharply bimodal, with the majority of material associated with small, so-called interplanetary dust particles (IDPs), weighing a few micrograms, and thankfully rare 'huge impactors', weighing of the order of a billion tons (see Figure 6.32).

IDPs of typical sizes of $\sim 5 - 35\,\mu m$ (on the rising flank of the left-hand peak of Figure 6.32, typically weighing a few to a few tens of nanograms) are small enough to have a sufficient surface area-to-volume ratio that they can efficiently radiate heat caused by friction as they enter the Earth's upper atmosphere and so do not evaporate or indeed become severely heated (and cause any organic material within them to burn away as CO_2 or become graphitic, i.e. non-organic). Such matter might therefore deliver

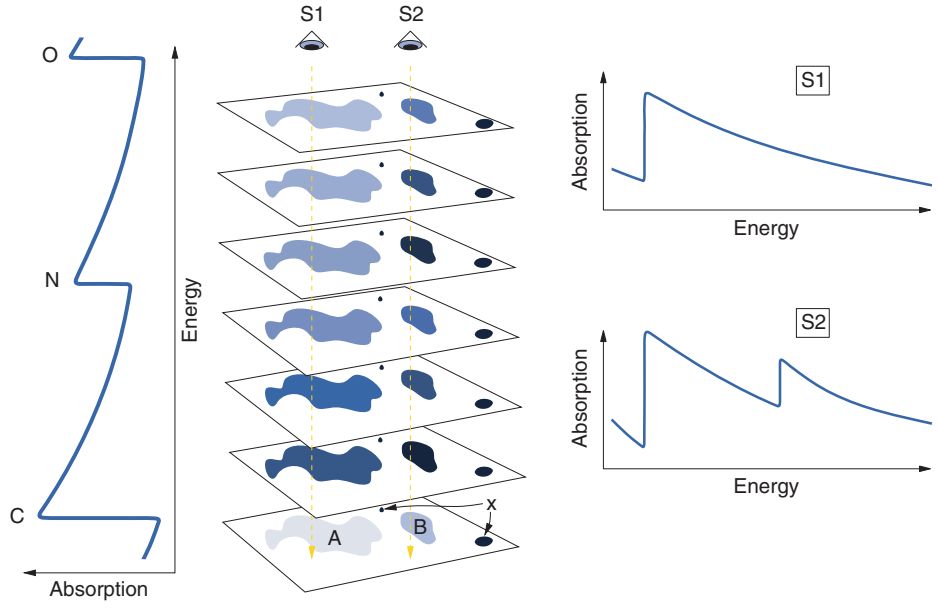

Figure 6.31 *Schematic diagram of STXM in the z-stack mode. A specimen consists of two features, A and B. A is carbon-rich, while B also contains high concentrations of nitrogen-containing compounds. Spatially resolved absorption maps of the sample are recorded over a range of energies. Two small features, labelled 'x', absorb strongly over the entire energy range of interest (they may, for example, be high-Z inclusions such as grains of silicate rock) and are used as reference points to accurately align the images on top of one another and thereby circumvent problems associated with sample drift. Spectra can therefore be extracted from this 'z-stack', by viewing any particular coordinate down through the stack (S1 and S2 in the figure).*

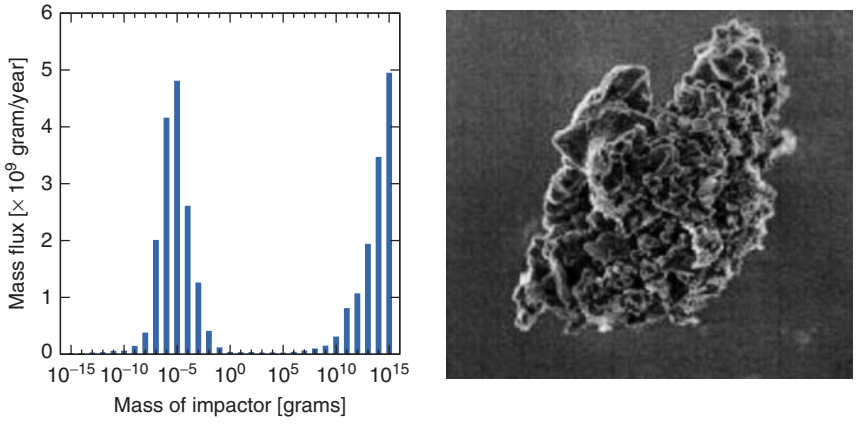

Figure 6.32 *Left: Mass flux for particles impinging on the Earth showing the bimodal distribution of small meteors and IDPs on the left and the setting for a Hollywood disaster movie on the right. Adapted from [40] with permission of Macmillan Publishers Ltd. Right: Scanning electron microscope image of an ∼ 10 μm interplanetary dust particle. Reprinted from [41] with permission of Elsevier.*

significant quantities of prebiotic organic matter, which in the early history of the Earth (when the accretion rate of such material was also likely to have been much higher) could have been potentially important for providing a reservoir for the development of life. A crucial question surrounding such a hypothesis, however, is an accurate knowledge of the abundance and types of organic compounds in the IDPs. Is or was there ever enough extraterrestrial carbon accreting on the Earth's surface for it to have contributed significantly to an organic reservoir? The answer to this question is tackled in this example.

Typical IDPs, as shown on the right-hand side of Figure 6.32, were collected using a fine sieve attached to a NASA aircraft flying at an altitude of 20 km. The particles are generally aggregates of submicron grains, dominated by oxide minerals, but are also thought to contain significant quantities of elemental carbon, carbonates and carbonaceous material, of the order of 12%. Hence, because of their small mass, chemical analysis of IDPs using traditional laboratory techniques is impractical, if not impossible. STXM, however, with its ability to focus down to submicron dimensions, is an ideal tool for such investigations. Experiments were carried out at the X1 undulator beamline of the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratories, which has a spatial resolution in this energy range of 50 nm [41].

STXM carbon mapping just below (282 eV) and above (292 eV) the carbon *K*-edge show the fraction of a typical IDP containing carbon-rich material (Figure 6.33). The volume fraction of carbon-rich material was determined as the ratio of the number of pixels showing a significant increase in absorption to the total number of pixels. Analysis of 12 such STXM maps showed highly variable carbon content, ranging from a few volume per cent to about 90 vol.%.

Once the C-rich areas were located, XANES around the *C*-edge was performed to identify which carbon functional groups were present. Nine of the twelve samples showed clear features at 285 eV and at 288.5 eV, characteristic of the C-ring structure and C=O carbonyl group, respectively. It is the presence of this carbonyl signature that indicates that a substantial fraction of the carbon in the IDPs is organic rather than elemental

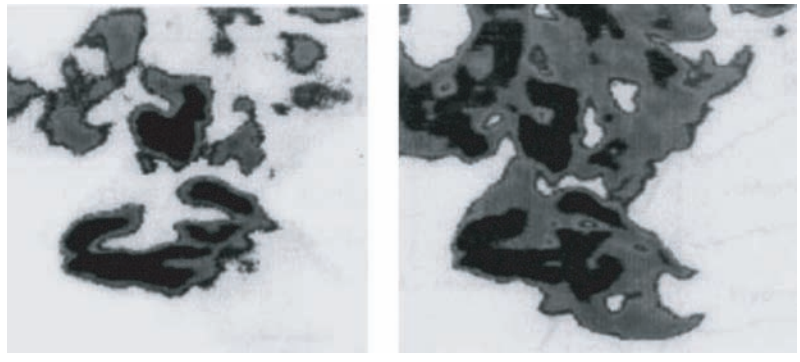

Figure 6.33 *STXM images of an* $\sim 2.5 \times 2.5\,\mu m^2$ *area of an IDP just below the carbon K-absorption edge at 282 eV (left) and just above it at 292 eV (right). Areas which have become dark in the right-hand image are carbon-rich. Reprinted from [41] with permission of Elsevier.*

(i.e. graphitic) in nature. Also, graphite has a strong, narrow, feature at 290 eV, due to exciton excitation, which was completely missing from the XANES spectra, indicating that it does not significantly contribute to the C-ring absorption line at 285 eV, and that therefore the carbon is primarily organic in nature.

The carbonyl content in the IDPs was further assessed as follows. The height of the carbon edge at $\approx$ 290 eV is a measure of the total amount of carbon in the analysis beam, while the area under the 285 and 288.5 eV features is proportional to the abundance of the C-ring and C=O functional group, respectively. In this manner, it was estimated that approximately 2% of the total mass of IDPs is bound in C=O. Additional XANES spectra around the oxygen absorption edge showed a strong absorption in the pre-edge region at 531 eV, consistent with the C=O bond.

In conclusion, of the approximately 300 tons of material that accrete each year on the Earth as IDPs, the STXM and XANES measurements indicate that 15 tons of this is organic. From the evidence of large lunar craters, it appears that the meteor flux in the early history of the Earth was very much higher, and that billions of tons of organic material are likely to have been deposited within a few millions of years of the Earth cooling and the seas forming, contributing significantly to the reservoir of potential life-building material. Is all life alien?

6.8 Photoemission Electron Microscopy

6.8.1 Basics of PEEM

Photoemission electron microscopy (PEEM) is a technique that images the spatial distribution of electrons emitted from a sample by x-ray absorption in the XANES region [42]. Although the technique uses photoelectrons for its signal, it is not an electron-spectroscopy technique, as the energies of the electrons are not distinguished. Instead, the absorption coefficient is indirectly measured via the yield of *secondary* electrons – not the photoelectrons directly emitted after photoabsorption, but the electrons released by the system after multiple scattering events in a cascade process originating with the directly produced photoelectron from the absorbing atom (see Figure 6.34(a)). This initial photoelectron may derive from an atom relatively deep in the material, as the x-rays can penetrate to depths of several tens of nanometres. The initial photoelectron can only travel a few Angstroms, however, before interacting inelastically with its surroundings via one of several scattering phenomena, which produces a cascade of 'daughter' electrons with ever decreasing energy. As the electrons' kinetic energy approaches only a few eV, the mean free path increases rapidly (see Figure 6.7) and they can escape the surface with higher probability – a typical graph of yield versus energy shows a maximum just above the work function, which then drops off with energy (Figure 6.34(b)). The majority of emitted secondary electrons therefore have energies of the order of an electronvolt. Although their detection does not yield absolute results for the absorption coefficient, the spectral dependence is adequately reflected (see Section 6.2).

A PEEM experiment essentially consists of tuning synchrotron radiation illuminating an entire sample in vacuum and imaging the spatial variation of the subsequent

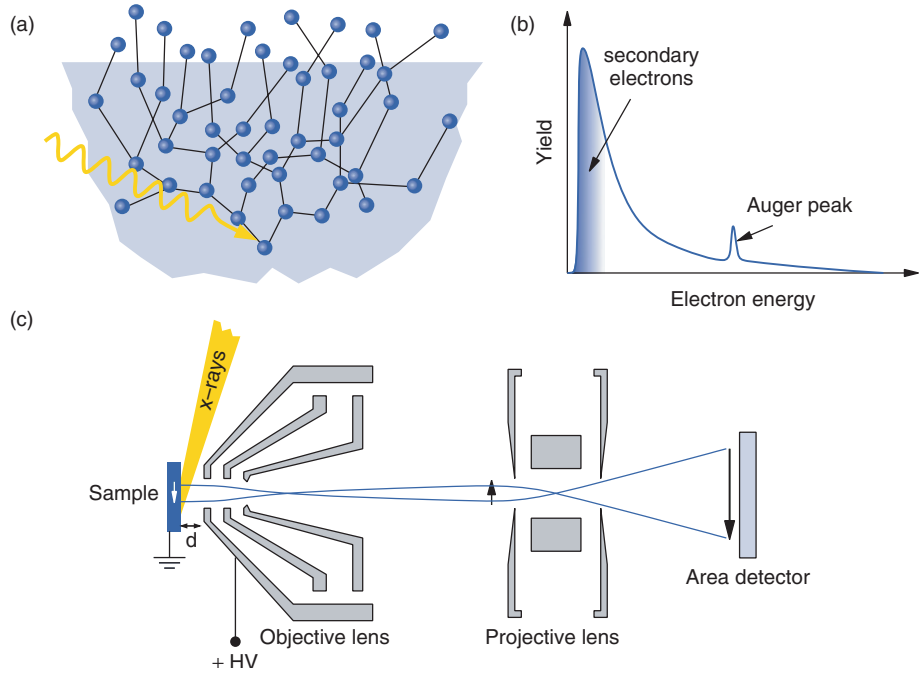

Figure 6.34 *Secondary electron detection in PEEM. (a) Secondary electrons are produced by a cascade of events seeded by the production of a photoelectron. As the electrons' kinetic energy decreases, the mean free path increases dramatically (note the longer paths between collisions closer to the surface), resulting in (b) there being an intense peak in the yield of the lowest energy electrons compared to Auger electron intensities at higher energies. (c) The secondary electrons are imaged in PEEM using electron optics.*

secondary-electron yield using an electrostatic and/or electromagnet lens system very similar to that used in electron microscopes (Figure 6.34(c)). As such, PEEM is a form of 'spectromicroscopy' – full-field images are recorded at different photon energies, in contrast to STXM, a microspectroscopy technique in which the focused x-ray beam is rastered across the sample.

The secondary electrons are imaged using a 2-D detector. In contrast to scanning electron microscopy, PEEM does not use a scanned probe electron beam, but instead the sample surface is uniformly illuminated by the x-rays. The magnified image of the surface can therefore be observed directly and in real-time.

Since electrons are used for imaging in PEEM, the resolution is not limited by the wavelength of the x-ray photon beam. Instead, a high electrostatic field between the sample and the objective lens accelerates the released electrons to energies of typically $eV_L = 10$ to 20 keV across a distance d of the order of 2 mm. This accelerating field acts as a lens – the trajectory of the electrons as they leave the surface form a set of parabolas. The tangents to the parabolas at the point where the electrons enter the objective lens extrapolate back to form a virtual image with unity lateral magnification (see Figure 6.35).

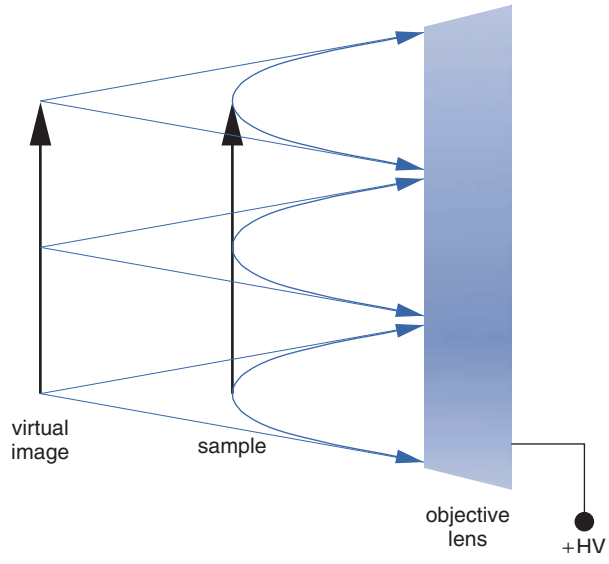

Figure 6.35 *Secondary electrons emitted from the sample with nonzero lateral kinetic energy execute a parabolic trajectory as they are accelerated to the objective lens. Extrapolation of the tangent to these paths where the electrons enter the objective for a virtual image with the same size. The accelerating field therefore acts as a lens with unitary magnification.*

The accelerating field is critical in determining the ultimate resolution. Because the electrons have a range of energies and emission angles, the virtual image will become blurred to a greater or lesser extent. The range of electron energies $\Delta\mathcal{E}_e$ can be reduced by introducing an aperture between the sample and objective. This limits the largest parabola width that can enter the electron microscope and therefore acts as a low-pass energy filter, though obviously at the expense of signal intensity.

To a first approximation, the resolution Δx is given by

$$\Delta x = \frac{d\,\Delta\mathcal{E}_e}{eV_L}. \tag{6.28}$$

The best achievable lateral resolution is therefore about 20 nm.

PEEM combines the techniques of x-ray absorption spectroscopy and electron microscopy. Chemical contrast is obtained by tuning the synchrotron radiation to a strong absorption edge and recording an image. Alternatively, in the spectromicroscopy mode, image 'stacks' can be recorded as a function of photon energy.

6.8.2 Worked Example – The Sea Urchin's Tooth

The Mediterranean purple sea urchin *Paracentrotus lividus* (Figure 6.36(a)) has the remarkable ability to bore into limestone rock ($CaCO_3$, calcite) in order to provide a secure purchase, despite the fact that the implement with which this is done is itself composed of calcite, a relatively soft and crumbly material. The grinding instrument consists of five teeth protruding through the so-called 'Aristotle's lantern' at the bottom

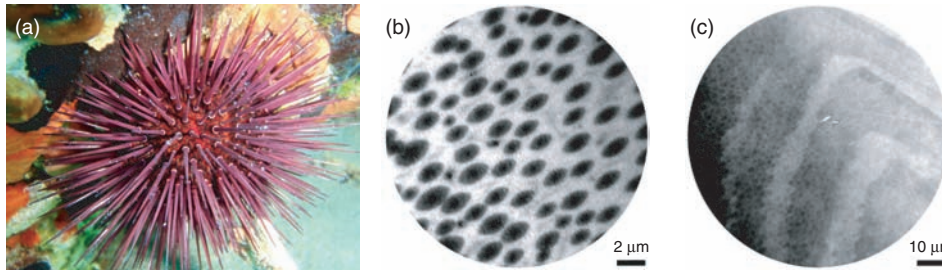

Figure 6.36 *PEEM imaging of a sea urchin tooth. (a) The Mediterranean purple sea urchin* Paracentrotus lividus. *Courtesy Peter Nicolaides, Aegean Institute. (b) Mg K-edge image of the transverse cross-section of the tooth tip. (c) Digital ratio image between the π^* and σ^* peak energies of the carbon XANES spectrum. Reprinted from [43] with permission of Proceedings of the National Academy of Sciences.*

orifice of the sea urchin's shell. This marvellous example of evolutionary engineering was recently investigated using several analytical techniques, of which PEEM provided the most direct and revealing information [43].

It is known that the tip of the sea urchin's tooth is formed by microscopic plates and needlelike structures made of calcite with a low magnesium enrichment ($Ca_{1-x}Mg_xCO_3$, $x \approx 0.1$), within a matrix of Mg-rich calcite ($x \approx 0.4$). The tooth develops first by the formation of the framework of needles and plates, each being a single crystal, followed by a filling-in of the spaces by the Mg-enriched matrix later in the development. Previous studies showed that the needles and plates were either all crystallographically aligned, or formed two subsets with a small angular offset.

The PEEM experiments were performed at the Mg K-edge at 1303 eV and at two carefully chosen positions of the XANES signal of carbon from 285 to 305 eV. The results are shown in Figure 6.36(b) and (c). The needles and plates in Figure 6.36(b) are dark because of their low Mg-content relative to that of the surrounding matrix. Indeed, the Mg-concentration increases towards the centre of the tip of the tooth. Substitution of Ca with Mg increases the hardness of the compound, hence this gradient in Mg-concentration is consistent with a hardening of the tooth tip.

The image in Figure 6.36(c) was generated by taking the ratio at each point between the XANES peaks associated with the σ^* (302 eV) and π^* (290.3 eV) unoccupied orbitals. As we have already seen from the example given in Section 6.4.2 about formate on copper, the orientation of the polarization that induces the maximum absorption for these two orbitals are perpendicular to one another. Therefore, the observation in this so-called polarization-dependent imaging contrast (PIC) display of bands in Figure 6.36(c) implies domains with different crystallographic orientations. The difference in orientation could not be quantitatively determined using the PIC images, but complementary XRD images showed them to be between 1 and 6°, probably close to 5°. Interestingly, within a given band, both the needles and plates *and* the surrounding matrix are aligned. It was therefore proposed that within a band (or 'block') the plate–needle complex is formed as a continuous crystalline structure, after which the high Mg-content matrix is seeded by secondary nucleation on their surface.

The alternating blocks were found to interdigitate, or interleaf, at the grinding tip. This may provide a functional advantage in which the plates at the surface of the tip can break off, or 'cleave' along a preferred crystal plane, exposing a fresh and sharp new surface. The slight misalignment of alternating blocks also means that a crack in one block will terminate at the boundary to the adjacent block and not propagate throughout the entire tip [43]. It seems that nature has discovered a similar solution to that of materials scientists working on tribological coatings for cutting tools, for which their performance is enhanced by depositing multilayers, which hinder crack propagation.

6.8.3 PEEM and Magnetic Dichroism

6.8.3.1 Introduction

Dichroism, meaning 'two-coloured', is the phenomenon of a material having an absorption spectrum which changes according to the polarization of the electromagnetic radiation used. Magnetic dichroism describes the dependence of the absorption of a magnetic material on the polarization *and* the relative orientation of the magnetic field. The discovery of magnetic dichroism in core-level x-ray spectroscopy has opened the route to magnetic domain imaging [44].

Methods utilizing magnetic dichroism in the x-ray regime have received renewed interest since core-levels can be probed with polarized soft or hard x-rays at synchrotron sources, which offer the unique feature of chemical specificity. Magnetic features down to below the micronscale can be imaged by exploiting the dependence of the absorption on the polarization of x-rays in ferromagnetic and antiferromagnetic materials. The methods are called x-ray magnetic circular dichroism (XMCD) and x-ray magnetic linear dichroism, respectively. Imaging down to the nanometre range is very important for magnetic structures, as at this scale, the influence of domain boundaries between one magnetic direction and another becomes significant and new phenomena can occur which would be negligible in larger structures. An understanding of the energetics of nanomagnetism is therefore essential in the drive to further miniaturize magnetic memory storage devices.

6.8.3.2 Spin and Magnetic Materials

Here we first describe the influence of spin on the energetics of core levels and valence band electrons in magnetic materials, then describe how dichroism in such systems can arise.

The intrinsic magnetic moment is produced by the spin of an electron and can be thought of as being a tiny bar magnet.[11] Note that the direction of the magnetic moment is antiparallel to that of the spin. Electrons always possess spin angular momentum of quantum number $s = 1/2$ and magnitude $\sqrt{3/4}\hbar$. The orientation of the spin relative to a magnetic field (either $+1/2$, 'up', or $-1/2$, 'down') determines its only two possible orientations.

Bound electrons in atoms which have non-spherically symmetric orbitals (i.e. not s-type) also have an orbital angular momentum and therefore generate a magnetic field that can be thought of classically as being produced by the 'current' of the electron as it

[11] This analogy should be taken with a pinch of salt – spin is one of only very few quantum-mechanical phenomena that has no classical analogy.

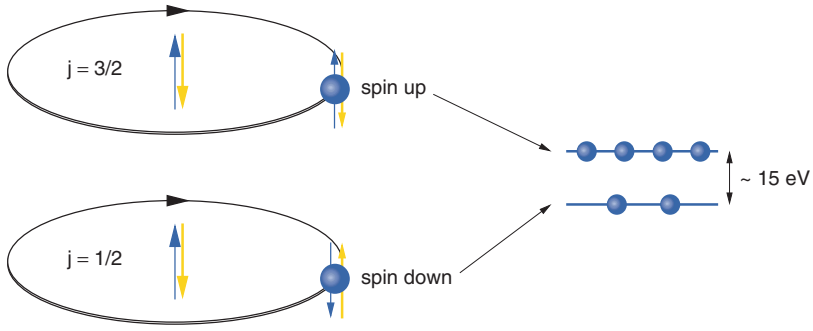

Figure 6.37 *An example of spin-orbit coupling. The 2p-orbital has an orbital angular momentum $l = 1$ and a magnetic moment which couples to that of the spin of the electron, causing the level to split into two distinct eigenstates $j = 1 + 1/2 = 3/2$ and $j = 1 - 1/2 = 1/2$. When these angular momenta are parallel, the magnetic moments are aligned unfavourably compared to when they lie antiparallel to each other. Hence $j = 3/2$ has a higher energy than $j = 1/2$. The maximum allowed occupation of each state is determined by the number of quantum-mechanically allowed projections m_j of j relative to the magnetic axis and is equal to $2j + 1$. The magnetic moments, shown in yellow, are antiparallel to their associated angular momenta, shown in blue.*

orbits the atom. These magnetic moments of the spin and the orbital angular momenta couple with each other just as two magnets are influenced by each others' fields, in so-called 'spin-orbit coupling'. Depending on whether the spin is oriented up or down relative to the axis of the orbital magnetic moment, the energy of the electron is higher or lower (see Figure 6.37).

Magnetic materials are characterized by having a partially filled valence shell. A magnetic moment in ferromagnetic materials is produced by there being an imbalance between spin-up and spin-down electrons. This normally arises because the spin-up and spin-down sub-bands have different energies, caused by the so-called 'exchange interaction'. We can qualitatively explain exchange interaction by considering Pauli's exclusion principle – each electron within an atom must have its own unique set of quantum numbers. The part of the quantum-mechanical wavefunction that describes the spin is special insofar that it does not affect the spatial distribution of the electron. Hence, those electrons that only differ in their spin quantum numbers overlap significantly and are therefore repelled from each other by Coulomb forces. This is the cause of the splitting both of the core levels, but also of the valence band levels, which, despite being extended wavefunctions, locally maintain much of their properties as isolated atoms. The splitting of the 3*d*-band caused by the exchange interaction is of the order of an electronvolt, while the bandwidth is typically a few eV.

6.8.3.3 XMCD

Consider Figure 6.38(a). A ferromagnetic material such as cobalt with 3*d* valence electrons forms domains with different magnetic directions. In region 1, there are more occupied 3*d* states with their spin down than with spin up, hence the magnetization of this domain is upwards (i.e. the opposite direction to the imbalance in spin states because,

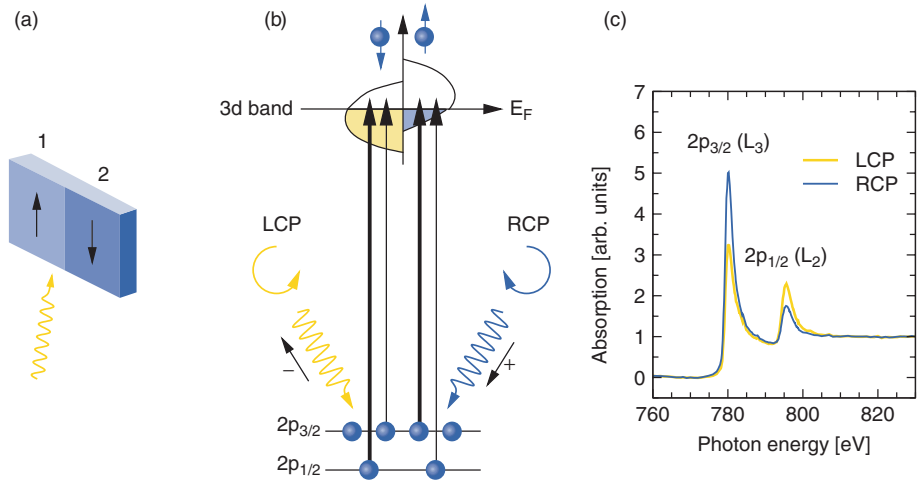

Figure 6.38 *The principle of x-ray magnetic circular dichroism. (a) An experiment where circularly polarized x-rays tuned to the L-edge of a magnetic material is shone at glancing incidence on a sample with different magnetic domain directions. (b) Absorption of RCP photons mainly excites spin-up electrons, while LCP light mainly excites spin-down electrons, resulting in a dichroic absorption spectrum, shown in (c) for the L-edge of cobalt. Courtesy Frithjof Nolting, Paul Scherrer Institut. Adapted from [44] with permission of Elsevier for (a) and (b).*

remember, the spin direction is opposite to the magnetization direction). In contrast, the number and density of *unoccupied* spin-up states above the Fermi level E_F in the $3d$ band is greater than spin-down states for region 1 (Figure 6.38(b)) and the transition probability for absorption of a photon will be greater. In addition, however, one must take a further selection rule into account which states that for dipole transitions, spin flips are forbidden. This rule is not entirely strict for condensed matter, but is sufficiently observed to induce dichroism. To obtain this so-called x-ray magnetic circular dichroism (XMCD), one uses circularly polarized light. So in the case of region 1, an incident RCP photon has its orbital angular momentum pointing forwards (in the direction of propagation, see Section 3.5.4) and will preferentially excite the $2p_{3/2}$ (L_3) levels over the $2p_{1/2}$ (L_2) levels. The opposite case arises for LCP photons, whereby the $2p_{1/2}$ states are favoured (Figure 6.38(c)). Lastly, if we now focus on region 2, the whole story reverses, as now the spin-down $3d$-sub-band is now shifted to higher energies than the spin-up sub-band.

By recording PEEM images of ferromagnetic domains with RCP and LCP x-rays and dividing the intensities of one image, pixel-for-pixel, by those of the other, the domain structure is revealed with maximum contrast. An example is given in Figure 6.39 for magnetic structures fabricated from thin films of Ni-Fe [45].

6.8.3.4 XMLD

X-ray magnetic linear dichroism (XMLD) is a technique used to investigate the domain structure of *anti* ferromagnetic (AFM) materials [46]. AFM materials are characterized by

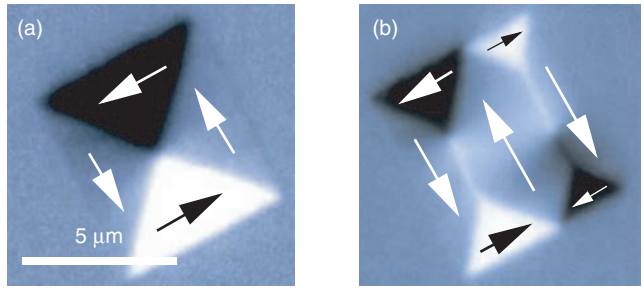

Figure 6.39 *(a) The orientation of microscopic domains in Ni-Fe thin films minimize the stray field energy. (b) Sometimes, such domains can break up into higher-energy metastable configurations with an activation barrier to lower-energy states. Recorded using the PEEM apparatus at the Surfaces and Interfaces Microscopy beamline at the Swiss Light Source. Courtesy Christoph Quitmann, Paul Scherrer Institut.*

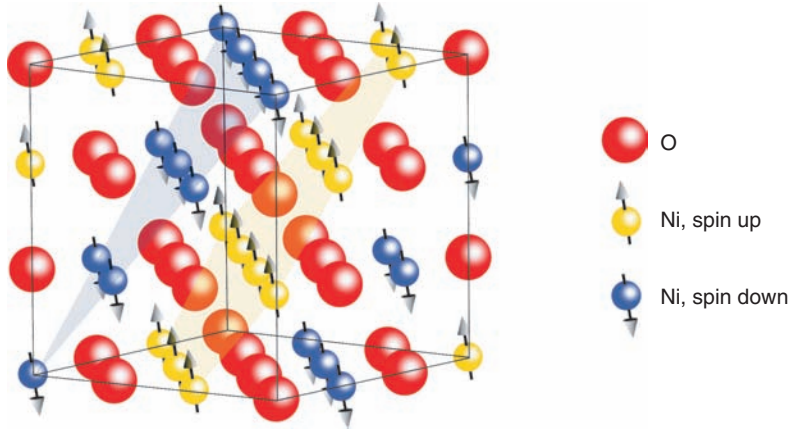

Figure 6.40 *NiO is an antiferromagnetic cubic crystal with a face-centred cubic rocksalt structure. Below the Néel temperature, the spins of the nickel ions in the (111) crystallographic planes (shown here in blue and yellow) alternate between being all spin-up and all spin-down. There is therefore no net magnetic moment, but there does exist a magnetic axis, given by the spin orientations.*

having antiparallel orientations of neighbouring magnetic moments and thus no net spin but a defined magnetic axis below a critical temperature (the Néel temperature). A simple schematic example is shown in Figure 6.40 for the antiferromagnet nickel oxide, NiO.

XMLD signal arises from the fact that, due to spin-orbit coupling, the spatial distribution of the electron density is marginally distorted, providing the necessary dichroism depending on whether the linear polarization vector is parallel to or perpendicular to the magnetic axis. The effect is usually small and detecting XMLD is more challenging than XMCD. It has a $\cos^2\theta$ dependence, where θ is the angle between the polarization vector of the x-ray light and the magnetic axis of the domain. The maximum XMLD

effect therefore occurs when the signal for the polarization parallel to the magnetic axis is compared to that perpendicular to the axis.

6.8.3.5 *Worked Example – Coupling of Ferromagnetic and Antiferromagnetic Domains*

Data recording and storage using magnetic materials is an area of enormous economic potential. Modern recording heads and disks use artificially layered structures containing several magnetic materials. Despite their widespread application in commercial devices, there remains unresolved issues regarding the fundamental physics that leads to the writing of magnetic bits in these modern structures.[12] The need for ever denser storage devices drives continued miniaturization of both bit sizes and recording heads. Techniques such as XMCD and XMLD which can image magnetic domains laterally, as a function of depth, and in a chemically sensitive manner are hugely important in unravelling the physics behind these phenomena.

The relationship between the orientation of ferromagnetic and antiferromagnetic domains was investigated on a system in which a 40-nm thick $LaFeO_3$ (LFO) antiferromagnetic film was grown on a $SrTiO_3$ substrate, followed by the deposition of a ferromagnetic Co-layer, which was capped with a 1-nm Pt-layer to prevent its oxidation and the formation of antiferromagnetic CoO [47]. This sample was then studied using the PEEM2 facility of the Advanced Light Source, Berkeley. The polarization of the x-rays could be changed from linear to right- or left-circular. The resolution was 20 nm. XMCD and XMLD images were recorded around the *L*-edges of Co (780 eV) and Fe (710 eV), respectively (see Figure 6.41). It can be immediately seen from the images that the orientation of the ferromagnetic domains was determined ('pinned') by the underlying antiferromagnetic domain structure. This is a surprising result, as it seems as if AFM domains, with no net magnetic axes can dictate the orientation of ferromagnetic domains. The physics of this fascinating system remains the subject of intense research.

6.9 Photoemission Spectroscopy

6.9.1 Introduction

The goal of photoemission spectroscopy is to determine the kinetic energy and, in the case of angle-resolved studies, also the three orthogonal momentum components of the wavevector of the emitted electrons. Armed with this information, one can draw important conclusions about the material's electronic band structure.

In Section 6.2.2, it was argued that the probability was very high for the ejection of a core electron into an unbound state if a photon with an energy higher than the binding energy of that electron was absorbed. The kinetic energies of these directly produced photoelectrons can be measured, which yields information on the elemental and

[12] The crux of the problem lies in the phenomenon of 'exchange bias'. The magnetic coupling of ferromagnetic and antiferromagnetic films across their common interface causes a shift in the hysteresis loop of the ferromagnet – this is the exchange bias. A practical application of the exchange bias effect is the pinning of the direction of magnetization in a ferromagnetic layer by an adjacent antiferromagnetic layer, so that the former does not reverse its direction when exposed to an external magnetic field. Read heads based on this effect are used in magnetic disk storage. The physics of the coupling, however, is poorly understood.

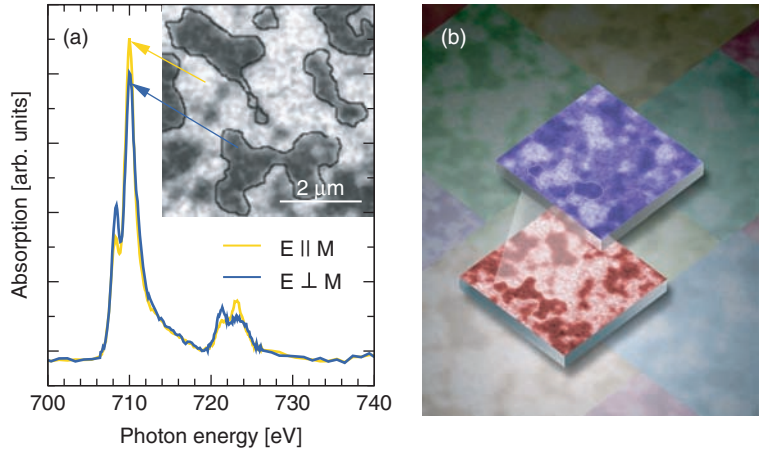

Figure 6.41 *Pinning of antiferromagnetic and ferromagnetic domains. (a) The antiferromagnetic axis M can change from domain to domain. These are dichroic with respect to the relative orientation of M with the electric field vector E of the linearly polarized x-rays. (b) Comparison of the XMLD map recorded at the Fe L-edge and the XMCD map of the same region recorded at the Co L-edge demonstrates that the exchange bias between antiferromagnetic domains (shown in red) pin the ferromagnetic domains (blue). Adapted from [47] with permission of Macmillan Publishers Ltd.*

chemical composition of materials. Photoemission spectroscopy (PES) exploits Einstein's explanation for the photoelectric effect to draw quantitative conclusions about the electronic and/or chemical nature of materials over a broad range of excitation x-ray energies. The reader is recommended a recent review of photoemission spectroscopies in all their varied guises as an excellent general introduction [48].

On the other hand, the valence electrons of atoms, molecules and condensed matter determine their electronic, optical, mechanical and magnetic properties. Techniques that probe the spatial distribution, energy, and momentum of valence electrons are, therefore, indispensable tools in understanding the properties of matter. These include on the one hand optical techniques in the visible or near-visible regime such as ellipsometry, reflectivity and transmission spectroscopy, in which the response is investigated of electrons which remain bound to the ionic cores. In the ultraviolet and soft x-ray range, on the other hand, the photon energy is sufficient to eject valence electrons from bound states where they can be detected and analysed for their energy and momentum. This information can then be interpolated back to the electrons' original properties within the bound system.

6.9.1.1 The Work Function

The minimum photon energy required to produce direct photoelectrons is given by the work function $e\phi$ of a given material, as famously described by Einstein in his Nobel-prize winning explanation of the photoelectric effect in 1905. This lies around 5 eV, hence photoelectron spectroscopy (PES) requires photons in the ultraviolet and x-ray regions. The work function arises because, as an electron originally residing in a

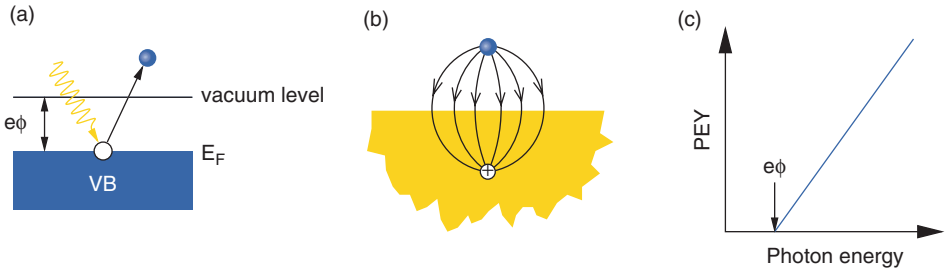

Figure 6.42 *(a) An electron in the highest-energy occupied state of the valence band (VB) still requires an energy $e\phi$ in order to reach the vacuum level and free itself of the system. (b) The reason for this is that, as it leaves the surface of a material, the electron is attracted to its positively charged mirror image produced in the near-surface region caused by screening effects due to the electron's removal. (c) The photoelectric yield (PEY) therefore increases with photon energy above the threshold energy $e\phi$. Below $e\phi$, no photoelectrons are produced, no matter what the beam intensity (or rate of impinging photons) is, a result of the corpuscular (quantum) nature of light.*

state at the top of the valence band (the Fermi level) is removed from the surface of a material, it is attracted by its positively charged mirror image produced by a rearrangement of the surface charge resulting from its departure (see Figure 6.42).

Only at distances of more than about 100 Å does this electrostatic image force become negligible. This image potential is the reason why electricity does not 'leak' out from wires and electrons do not shoot out of kinks in current-carrying wires; for typical work functions of 5 eV, one requires electric field strengths of the order of 100 MeV/m to electrostatically remove electrons from metal surfaces. On the other hand, the phenomenon of thermionic emission results from valence electrons in a resistively heated metal obtaining sufficient thermal energy that they can overcome the work function barrier and 'boil off'.

We now include the work function $e\phi$ in describing the kinetic energy of an emitted photoelectron, by modifying our definition of the binding energy E_B of electrons in isolated atoms given in Equation (6.8). From our explanation of the origin of the work function, it should be clear that it is a manifestation of condensed matter. Hence, now, E_B is the energy needed to promote the electron to the continuum (or 'vacuum level') but is insufficient to pull it away from its positively charged mirror image, for which we still need the work function $e\phi$, that is

$$\mathcal{E}_e = h\nu - E_B - e\phi. \tag{6.29}$$

It is noted that the binding energies of energy levels in condensed matter are conventionally measured with respect to the Fermi level, rather than the vacuum level.

6.9.1.2 Energy Regimes of Photoelectron Spectroscopy

Here, we discuss only photoelectron spectroscopy of condensed matter. One can divide PES into three broad categories according to the energy range of the photons that are used (see Figure 6.43).

Techniques such as ultraviolet photoelectron spectroscopy (UPS) and the more sophisticated angular-resolved photoelectron spectroscopy (ARPES) probe the electronic

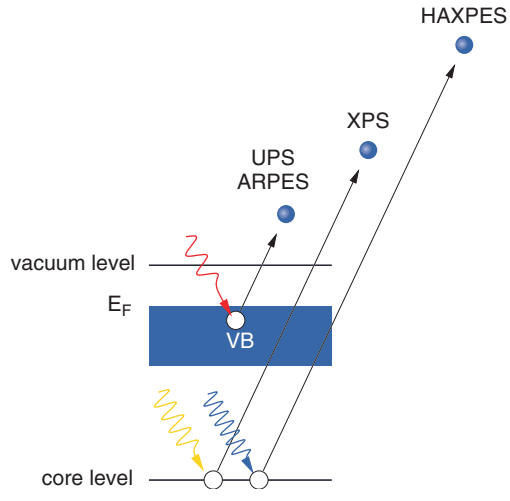

Figure 6.43 *Low-energy ultraviolet and soft x-rays are used in UPS and ARPES to investigate the electronic structure of matter by probing the valence electrons. In XPS, elemental and chemical information is gleaned from the energy and signal intensity of photoelectrons originating from core levels. HAXPES also provides similar information, but from deeper within the sample using hard x-rays. The reason for this is not that hard x-rays penetrate deeper, but that high-energy photoelectrons have a longer mean-free path (Figure 6.7).*

structure of matter using low-energy photons from a few eV to a few hundred eV. Elemental and chemical analysis is primarily carried out in x-ray photoelectron spectroscopy (XPS, also called electron spectroscopy for chemical analysis, ESCA) by probing photoelectrons originating from core levels, and typically spans the energy range of a few hundred to one or two thousand electronvolts. Finally, hard x-ray photoelectron spectroscopy (HAXPES) is a variant of XPS in which the incident photon energy far exceeds the electron binding energy, hence the photoelectron energy is also high. This has the advantage that the escape depth of high-energy electrons is large (see Figure 6.7) and so the properties of the bulk material or of buried interfaces far below the sample surface can be investigated.

We now discuss each of these techniques in turn.

6.9.2 Ultraviolet Photoemission Spectroscopy

Ultraviolet photoelectron spectroscopy (UPS) uses photons in the vacuum ultraviolet and soft x-ray regimes from approximately 10 to 100 eV and utilizes energy-dispersive analysis of the emitted photoelectrons to study the electronic states of the near-surface region of a sample, down to a few nanometers.

In laboratory-based UPS equipments, the choice of radiation is limited to fixed energy line sources such as the He I (21.1 eV) and He II radiation (40.8 eV). Importantly, UPS performed at synchrotron sources has the added experimental degrees of freedom that the photon energy can be tuned and the polarization changed. Hence, by using a synchrotron, the relative change in photoemission cross-section for various electron states can be used to determine the partial density of states (PDOS).

In addition to analysing the energies of the photoelectrons, it is also possible to probe their angular distribution. This enables the energy versus wavevector of the electronic states to be measured in so-called 'energy dispersion curves' (EDCs, described below). When the angular distribution is measured, UPS is termed angle-resolved photoelectron spectroscopy (ARPES), otherwise the UPS measurement is angle-integrated.

We now proceed first by introducing concepts related to valence bands in a qualitative manner, after which we discuss how these features can be probed in a UPS setup. The purpose here is not to provide a rigorous description of band theory, but to convey the relevant physical phenomena that can be probed using UPS. For rigorous derivations, there are many classic texts (see for example [49, 50]), or specialist reviews [51, 52].

6.9.2.1 Valence Band Structure

As we have already discussed at the beginning of this chapter, the valence electrons in a solid occupy very closely packed levels within bands. We begin with the simplest case of a good metal. To a first approximation, one can assume that electrons in metals are essentially free to move and, because of efficient screening by the other surrounding electrons (the 'electron gas'), are not influenced by the periodic electrostatic field of the ionic cores. In this case, the allowed electron states are determined by the permitted de Broglie wavelengths of the electrons, which in turn are determined only by the boundary condition that the electric field must be zero at the surface of the metal.

The energy of an electron in these metallic bands depends on its momentum $\hbar k = h/\lambda$. More precisely,

$$E_k = \frac{p^2}{2m_{\text{eff}}} = \frac{\hbar^2 k^2}{2m_{\text{eff}}}, \tag{6.30}$$

where $k = 2\pi/\lambda$ is the wavevector of the electron in the crystalline solid and m_{eff} is its effective mass. This last quantity describes the apparent mass of an electron if one were to assume it were completely free – a large effective mass indicates a high degree of localization. It is different from the rest mass of a free electron because in a crystal, an electron interacts with the surrounding ions and other electrons. Hence our assumption that an electron in a metal is free is only approximately true – it is perhaps better to focus on the statement that electrons in a metal do not feel the *periodic* potential of the crystal field. This potential is, as we will shortly see, responsible for the formation of bandgaps–energy regions which contain no electron states. In spite of these, the bands may still be thought of as being *locally* parabolic. Therefore, from Equation (6.30), the parabola describing energy versus wavevector is shallow and wide for an electron with a large electron mass, while a low effective mass results in a steep and narrow parabola.

The energy of the electron within a solid therefore increases parabolically with its wavevector. Note that E_k is the energy of the electron *within* the material and should not be confused with $\mathcal{E}_e$, the kinetic energy of an escaped photoelectron, although we will see that by measuring the latter, the former can be deduced.

How closely packed *are* adjacent valence-band levels within a macroscopic solid? Let us first consider the simplest case of the allowed wavelengths along a linear object of length L. Each level has its own wavelength and is associated with two states with opposite spins. The electron wavelength can only assume values $2L/n$, where n is a positive integer, because the boundary conditions dictate that the amplitude must be

zero at the ends (or surfaces, in 3-D). Hence

$$k = n\pi/L. \tag{6.31}$$

The difference Δk between adjacent states n and $n+1$ is π/L and so the density of states is therefore linear with respect to k. From Equation (6.30), we obtain

$$E_k^n = \frac{\hbar^2}{2m_{\text{eff}}}\left(\frac{n\pi}{L}\right)^2. \tag{6.32}$$

and the energy separation ΔE_k between adjacent states is

$$\Delta E_k = \frac{n}{m_{\text{eff}}}\left(\frac{\hbar\pi}{L}\right)^2. \tag{6.33}$$

Depending on n and considering macroscopic objects, the highest occupied level, ΔE_k is many orders of magnitude smaller than an electronvolt, and is much smaller still for three-dimensional objects.

What values do the wavevector and energy take on at the highest occupied level at absolute zero, the Fermi level? Let us assume that in the linear model introduced above that each unit 'cell' of length a contributes a single valence electron to the band structure. The total number of electrons is therefore $N = L/a$. At the Fermi energy, the total number of valence electrons N equals $2n = 2n_F$, as each level can accommodate a spin-up and spin-down state. Hence

$$k_F = \frac{N\pi}{2L}, \tag{6.34}$$

$$E_F = E_k^N = \frac{\hbar^2}{2m_{\text{eff}}}\left(\frac{N\pi}{2L}\right)^2. \tag{6.35}$$

As we have mentioned, this 'free-electron model' ignores the periodic electrostatic potential of the core ions. In reality, the valence electrons are affected by this potential energy field U, and interact with it most strongly when

$$k = \frac{m\pi}{a}, \tag{6.36}$$

where m is an integer. This can happen either in an attractive or repulsive manner, depending on whether the electron density (which is proportional to the square of the electron wavefunction) is highest or lowest above the ion cores, respectively (see Figure 6.44). The result of this interaction is to lower or raise the energy of those electrons depending on the phase of their wavefunctions relative to the ion core array.

At which energy does the band gap appear? If we equate Equations (6.31) and (6.36), we see that

$$\frac{n\pi}{L} = \frac{m\pi}{a}. \tag{6.37}$$

Let us consider the first band gap, for $m = 1$, for which therefore $L = na$, which we have argued above is true for $n = N$. In other words, the band gap opens up around the Fermi energy.

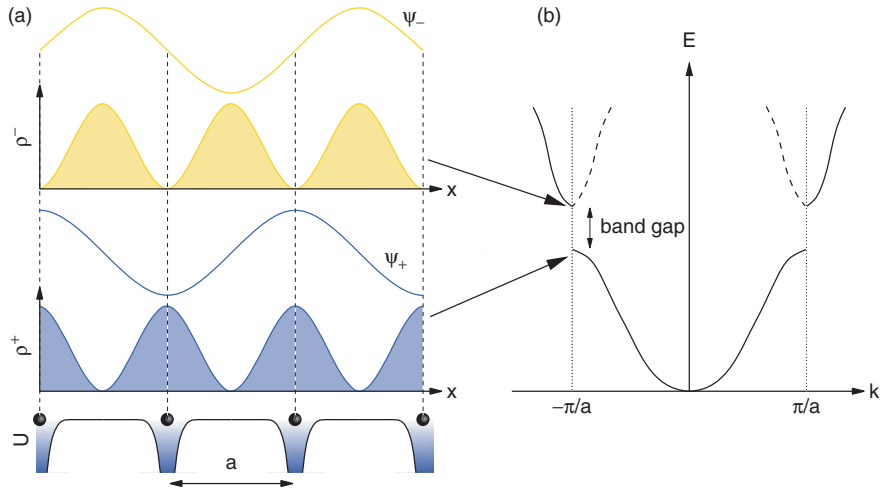

Figure 6.44 *The formation of band gaps. (a) Two standing waves ψ_+ and ψ_- with wavevectors $n\pi/a$ are possible solutions of electrons moving back and forth within a material of periodicity a. The probability ρ of finding an electron at any one position is proportional to the square of the wavefunction. ρ_+ has maxima that lie exactly above the positively charged ionic cores and is therefore more strongly bound than ρ_-, which exhibits minima at the ion cores. The result of this is shown in the plot of energy versus wavevector in (b) whereby at these special wavevector values, an energy band gap is produced. One can also represent the bandstructure in only one zone between $-\pi/a$ and $+\pi/a$ by translation by a reciprocal lattice vector $2\pi/a$ (shown here as dashed lines).*

Lastly, because of the periodic nature of crystals, the reciprocal space is also periodic, with the periodic unit in one dimension being $2\pi/a$. One can therefore 'fold' the bandstructure within the central region bounded by $\pm\pi/a$ by translation by a reciprocal lattice vector. This is shown in Figure 6.44 by the dashed lines reflecting the bandstructure above the band gap.

6.9.2.2 The Fermi Surface

In our one-dimensional model, the highest occupied state is a point along the k-axis with energy E_F and wavevector k_F, which we call the 'Fermi point'. In three-dimensional k-space, we speak of the 'Fermi surface', which describes the wavevectors in every possible direction below which the electron states are filled at absolute zero. In the case of a metal, the Fermi level lies within a continuous band, *at least along a crystallographic axis*. In some arbitrary direction, however, the magnitude of the wavevector might assume some other value, depending on the interaction between the electron and the ionic electrostatic potential in that direction, and a band gap might be formed. In other words, the form of the Fermi surface is a function of the crystallographic directions, and even for the most simple elemental metals such as copper, the Fermi surface can assume complex forms, as shown in Figure 6.45.

In this figure, the Fermi surface is seen to be encased within a polyhedron which is called the 'Brillouin zone'. Due to invariance of the bandstructure with respect to

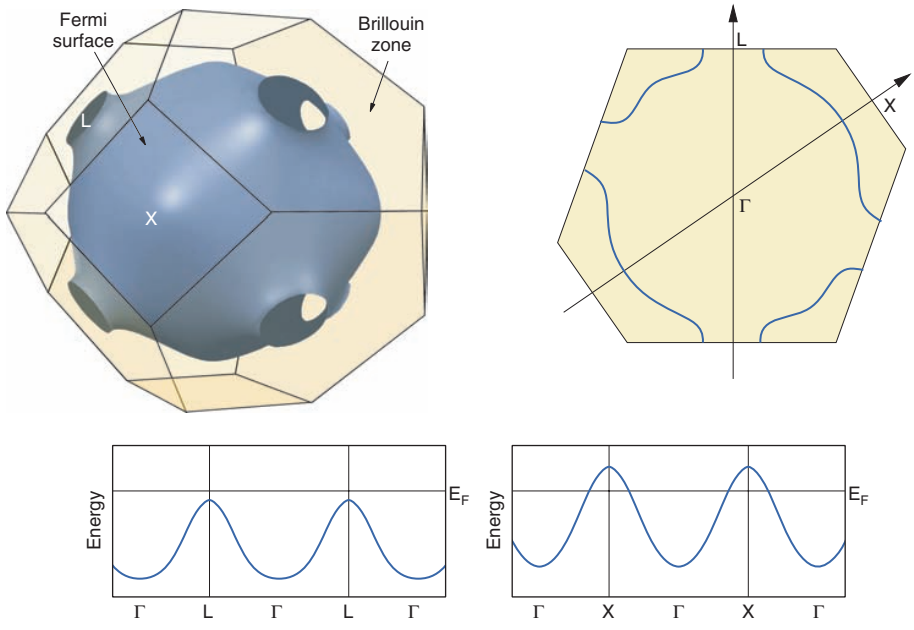

Figure 6.45 *The Fermi surface of copper within the Brillouin zone, which, for the face-centred cubic Cu-lattice assumes the form of a truncated octahedron. A slice through the Fermi surface and Brillouin zone is shown on the right. The centre of the Brillouin zone is called the* Γ*-point, while the two crystallographic directions (111) (denoted 'L') and (100) ('X') are also shown. Note that the Fermi surface never crosses the* Γ*–L line and in this direction all the states are below the Fermi energy and therefore a gap opens up. In contrast, in the* Γ*–X direction, the Fermi surface lies within the Brillouin zone and copper is conducting in this direction.*

translations by integral multiples of the reciprocal lattice vector (a consequence of the periodicity of the crystal), the entire bandstructure is contained within the Brillouin zone. We see from Figure 6.44 that the boundaries of the Fermi surface (or 'Fermi point' in one dimension) were at $\pm\pi/a$, in other words exactly half way in between adjacent reciprocal lattice points at $\pm 2\pi/a$. If we extend this concept to three dimensions in k-space, what we obtain is the Brillouin zone. This is therefore the polyhedron defined by the smallest volume to be enclosed by planes that bisect at right-angles lines connecting adjacent reciprocal lattice points. We have come across this construct already in our discussion of Wigner–Seitz cells in Section 5.3.1. Indeed, the Brillouin zone is the Wigner–Seitz cell of the reciprocal lattice of the crystal type being investigated. In the example shown in Figure 6.45, the reciprocal lattice of the face-centred cubic lattice of copper is a body-centred cubic lattice, which has the shown truncated octahedral Brillouin zone.

Hence, the Fermi surface of a material reflects the interaction of the valence electrons with their crystallographic environment, which in turn determines the electronic properties (conductivity) and optical properties (reflectivity and colour). A detailed knowledge of the Fermi surface is therefore of great importance. This, and related information, are the primary goals of UPS and in particular ARPES experiments.

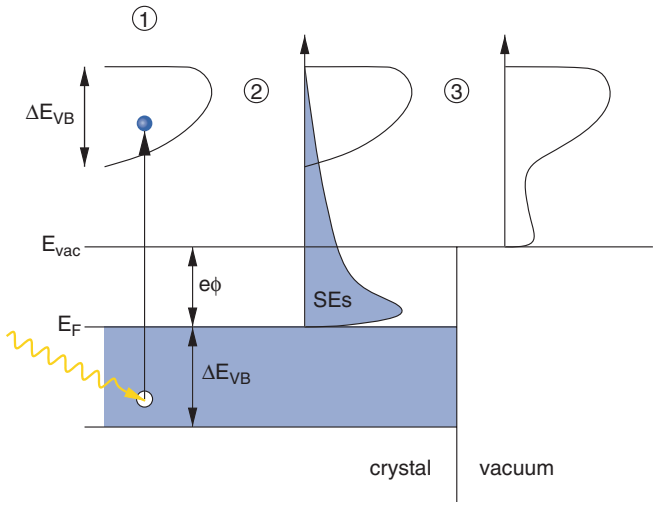

Figure 6.46 *The three-step model. First, a photoelectron is produced by absorption of an x-ray photon at some finite depth in the crystal. Next, the photoelectron travels to the surface. Because the escape depth is in general significantly smaller than the x-ray absorption depth, the photoelectron can be scattered, producing secondary electrons (SEs). The electrons then penetrate the surface. The spectral form consists of the direct (unscattered) photoelectrons, plus a low-energy tail of those secondary electrons which have energies above eϕ, the work function.*

6.9.2.3 The Three-Step Model

It is convenient to discuss the physical processes in photoemission by breaking down the ejection of a photoelectron from matter into successive events in the so-called 'three-step model' (see Figure 6.46):

- photo-induced excitation of an electron from a filled state to an empty conduction band state within the crystal;
- ballistic (i.e. without scattering) transport of the electron to the surface. This might be also contaminated by signal originating from inelastically scattered secondary electrons;
- transmission of the electron through the surface, resulting in its emission from the solid.

We begin with the first process. We have already discussed the conservation of energy (Equation (6.29)) with regards to photoemission. Let us next consider momentum conservation. The momentum of the electron is $\hbar\mathbf{k}$ with $|k| \sim 2$ Å. On the other hand, the momentum of a 100 eV photon is only of the order of 0.05 Å. Hence photoabsorption imparts insignificant momentum transfer and the transition from the initial to final electronic state is 'vertical', that is, they both have essentially the same wavevector $\mathbf{k}$. This is an important result – a truly free electron has an open, parabolic energy-versus-momentum curve and for a given energy, the wavevector k is uniquely determined by Equation (6.30). Changing the kinetic energy of a free electron is

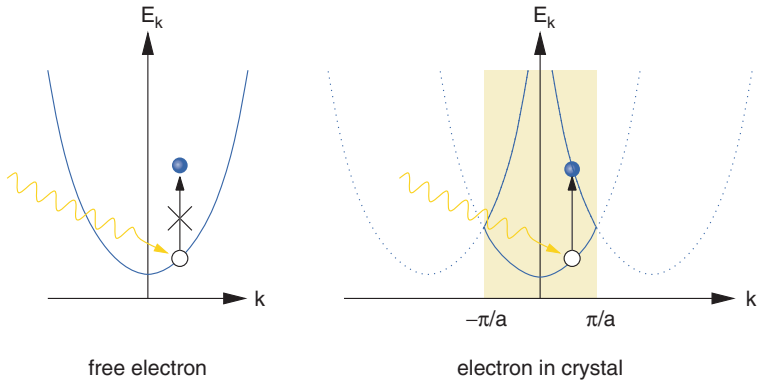

Figure 6.47 *A truly free electron cannot absorb a photon, as the gain in energy would put it into a state of higher energy, but almost equal momentum. Such a state does not exist for free electrons, where $E(k)$ has unique solutions. Only in bound or periodic structures can states arise with different energies but the same wavevectors, allowed by the folding of higher-energy band structures back into the first Brillouin zone.*

associated with a corresponding change in k, but this cannot occur, as the photon momentum is insignificant. Therefore incompatibility between energy- and momentum conservation means that photon-stimulated transitions are forbidden for truly free electrons. On the other hand, photoemission transitions *are* allowed in periodic structures, thanks to the fact that the bandstructure can be folded back on itself within the Brillouin zone (see Figure 6.47). The folding process means that the momentum of the final (upper) state differs from that of the initial state by an integer multiple of $2\pi/a$.

The depth to which photoelectrons are generated is determined by the absorption depth of the x-rays being used in the material under investigation, and is of the order of a few tens of nanometres for soft x-rays between 50 and 200 eV. However, the escape depth of the photoelectrons, given by the universal curve (Figure 6.7) is an order of magnitude smaller, hence valence bandstructure measurements are sensitive only to the first few atomic layers. In addition, spectral peaks have 'loss tails' towards low kinetic energies.

Lastly, we must consider what happens when the electron passes through the surface and into vacuum. Because of the work function the electron slows down as it passes through the surface (Figure 6.48(a)). Hence, in this third process only the component of **k** parallel to the surface is preserved, that is

$$k_{out,\parallel} = k_{in,\parallel} \equiv k_{\parallel}.$$

The potential step at the surface breaks the translational invariance perpendicular to that surface and so, unlike $k_{\parallel}$, $k_{\perp}$ is not conserved. From Equation (6.30),

$$\begin{aligned} k_{out} &= \sqrt{\frac{2m}{\hbar^2}\mathcal{E}_e}, \\ k_{in} &= \sqrt{\frac{2m}{\hbar^2}(\mathcal{E}_e + V_0)}, \end{aligned}$$

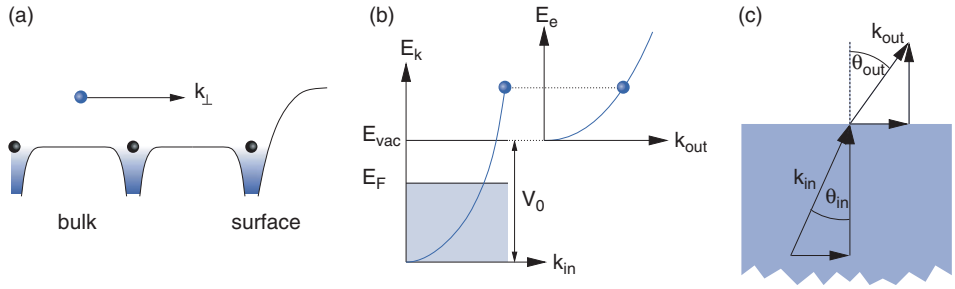

Figure 6.48 *Transmission of photoelectrons through a surface. (a) Because of the work function, there is a potential barrier for the photoelectron leaving the surface, which slows it down in the direction normal to the surface. (b) In order to determine the wavevector k_{in} of the photoelectron within the solid, the free-electron parabolas must be matched, which requires a measurement of $\mathcal{E}_e$, the kinetic energy of the photoelectron after leaving the surface and thereby k_{out}, its wavevector, plus a knowledge of V_0, the inner potential. (c) For a given kinetic energy of the photoelectron $\mathcal{E}_e$, there is a maximum angle θ_{in}^{max} that produces a photoemission signal. For larger angles, the component of the energy of the photoelectron within the solid perpendicular to the surface is insufficient to overcome V_0.*

where V_0 is the so-called 'inner potential' the energy required to remove an electron from the bottom of the parabolic potential within the solid to vacuum. From Figure 6.48, we see that

$$\begin{aligned} \sin\theta_{in} &= \frac{k_{\parallel}}{k_{in}} \\ &= \frac{\sin\theta_{out}\sqrt{\frac{2m}{\hbar^2}\mathcal{E}_e}}{\sqrt{\frac{2m}{\hbar^2}(\mathcal{E}_e + V_0)}}, \end{aligned}$$

which assumes its maximum value when $\theta_{out} = 90^\circ$, that is

$$\sin\theta_{in}|_{max} = \sqrt{\frac{\mathcal{E}_e}{\mathcal{E}_e + V_0}}. \tag{6.38}$$

Hence, by detecting photoelectrons emitted from a surface at different emission angles, the energy of the electrons as a function of the momentum vector may be determined. This process is known as 'band mapping' and is a powerful probe of the electronic structure of crystalline materials. The measurements can usually be compared with theoretical predictions.

6.9.2.4 Presentation of ARPES Data

One can map out the density of occupied states in the valence and conduction bands of a material using ultraviolet photons. It is tempting to assume that the energy distribution of the electrons measured by UPS is a direct map of the band density of states. That is only true if the photoemission probability is the same for all valence electrons. From our discussions of photoabsorption in Section 6.2, in particular Equations (6.4) and (6.14),

it was argued that both the initial and final states are of importance in the photoemission process and the photoemission probability may vary across the range of binding energies of the valence band. Therefore, the intensities of the UPS spectra are weighted by the transition matrix element representing the coupling of the initial and final states.

If we assume a constant transition probability for all valence band states, the resulting intensity distribution for a fixed incident photon energy as a function of photoelectron energy is therefore equal to the valence band density of occupied states multiplied by the Fermi–Dirac function (which accounts for thermal promotion of valence-band electrons into the conduction band) and is shown schematically in Figure 6.49 for an UPS experiment at a fixed emission angle. Such plots are referred to as 'energy distribution curves' (EDCs).

We argued above that a knowledge of the Fermi surface of a crystalline material is indispensable in predicting the optical and electronic properties of that material. The binding energy E_B at the Fermi surface is zero. We can therefore map out the Fermi surface by setting the kinetic energy of our detector (the CHA, Section 4.6.7) to $\mathcal{E}_e = h\nu - e\phi$ and varying the probed exit angles of the photoelectron. In this manner, the Fermi surface can be plotted as a function of the orthogonal in-plane wavevectors k_x and k_y, whereby $k_{\parallel}^2 = k_x^2 + k_y^2$. An example of the Fermi surface map (FSM) of Ni(110) is shown in Figure 6.50.

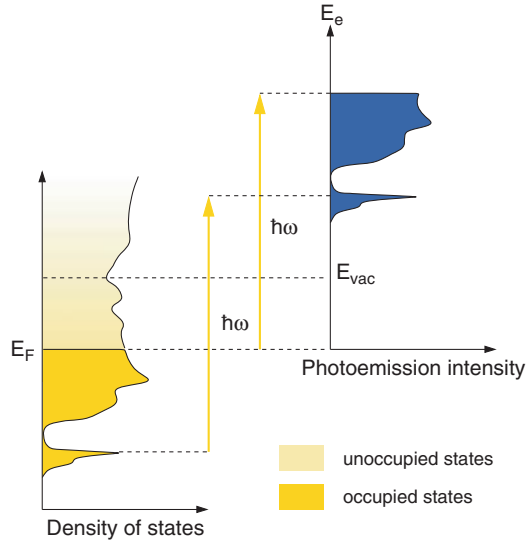

Figure 6.49 *Energy distribution curves. EDCs are generated by recording the photoelectron intensity as a function of photoelectron kinetic energy for a fixed photon energy and detector angle. Assuming a constant transition probability, the intensity at a kinetic energy* $\mathcal{E}_e = h\nu - e\phi - E_B$ *is proportional to the local density of states at* E_B *having the in-plane wavevector* $k_{\parallel}$ *selected by the detector orientation. In this schematic, the photoemission intensity has been plotted as a function of* $\mathcal{E}_e$ *for didactical reasons. More commonly, however, it is plotted as a function of* E_B.

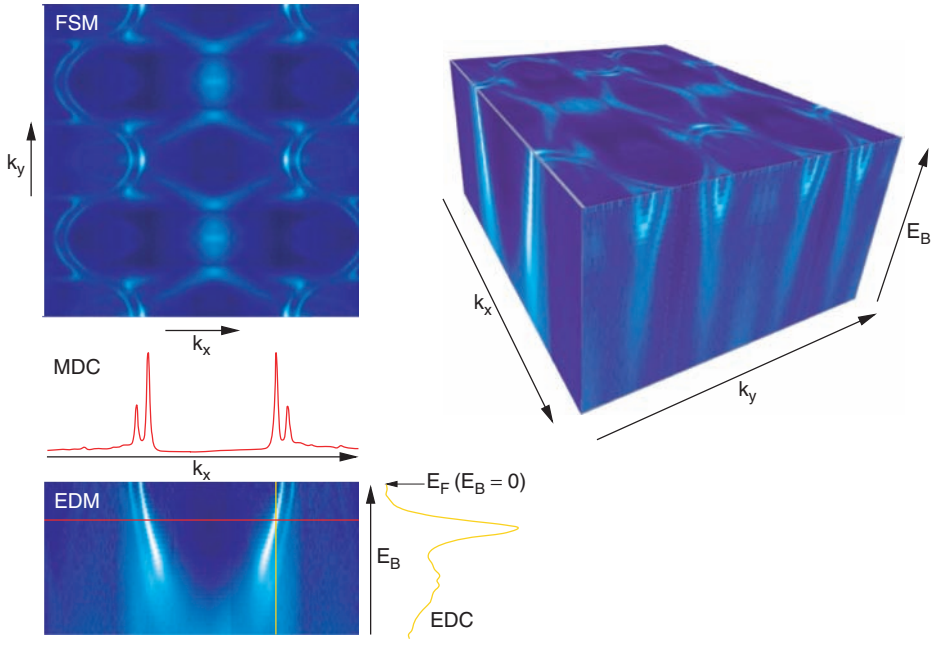

Figure 6.50 *Representations of ARPES data. A Fermi surface map (FSM) is generated by recording the intensity at $\mathcal{E}_e = h\nu - e\phi$ for different k_x and k_y. Momentum distribution curves (MDCs) track the change in intensity of the ARPES signal for a fixed binding energy as one navigates in a fixed in-plane direction (e.g. along k_x). Energy distribution maps can either be generated by stacking MDCs for a range of binding energies, or by stacking EDCs for different positions along a certain in-plane momentum axis. Finally, a full three-dimensional representation of ARPES data is possible by plotting the intensity as a function of E_B and the in-plane wavevectors k_x and k_y. The results shown here are for Ni(110). Adapted from [53] with permission of the American Physical Society, courtesy Luc Patthey, Paul Scherrer Institut.*

The FSM of Ni(110) shown in Figure 6.50 is for $E_B = 0$, that is, at the Fermi energy. Similar maps can be generated for electrons originating from occupied states below the Fermi energy by adjusting the CHA. Vertical slices through the resulting three-dimensional intensity distribution, shown on the right-hand side of Figure 6.50, expose energy distribution maps (EDMs), also shown in Figure 6.50. These reveal some of the most important information on the electronic properties of crystalline materials, in particular on the presence and directions of band gaps, and details of the overall bandstructure.

6.9.2.5 Surface States

Surface states are associated with phenomena which are localized normal to the surface. This can be due to a rearrangement of bonds and change in bond strengths at a surface, often seen in semiconductors, but also occurs in quasi-two-dimensional systems such as in the conducting planes of graphite or in the superconducting CuO_2 sheets of high-temperature superconductors.

How can one distinguish between 'normal' bulk states and surface states? To answer this, the most important aspect to recognize is the *localization* of surface states to a two-dimensional plane. If a state is physically localized in one direction, the corresponding periodicity is lost, which can be mathematically formulated as an *infinite* periodicity in this direction. In contrast, as the wavevector is inversely proportional to the periodicity, this becomes completely *de*localized perpendicular to the two-dimensional feature, and can take on any value. This is shown schematically in Figure 6.51.

In reciprocal space, therefore, surface states are present for all possible values of $k_\perp$, while the in-plane wavevectors are well-defined and fixed. Surface states hence have a fixed binding energy E_B, independent of $k_\perp$.

Now let us imagine an experiment in which we probe only those electrons emerging perpendicular from the surface, for which $k_{in,\parallel} = k_{out,\parallel} = 0$. From Figure 6.52, it is immediately clear that, in contrast to bulk states, the position of surface states does not disperse if one changes the photon energy.

Surface states are in general located in bandgap regions for which there are no bulk-like states. The reason for this is that if a surface state were degenerate with a bulk state, they could couple and the surface state would take on some of the character of the bulk state, thereby losing its surface localization.

A surface state will be strongly affected by adsorption and the introduction of contaminants to the surface of the sample under investigation can kill surface-state signal. Because, however, the influence of some contaminants can extend several monolayers into a sample, states which are normally considered to be bulk-like can be affected, and this test is less stringent than the first.

An example of a surface state in the Fermi surface map of Cu(111) is shown in Figure 6.53. Normal to the surface, the bulk density of states at the Fermi energy is zero and copper has a band gap in this direction. However, a localized signal is indeed found in this direction, associated with a surface state [51].

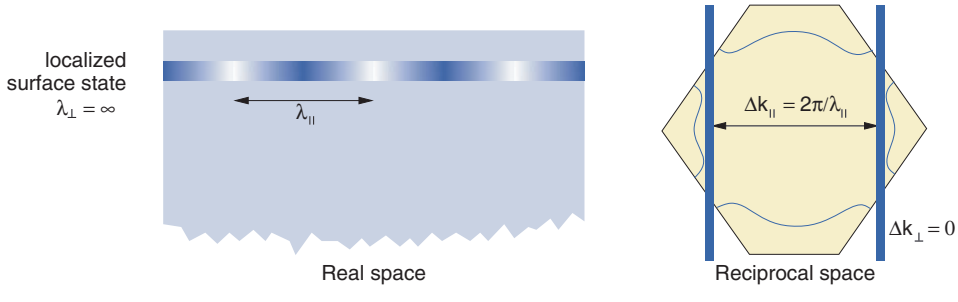

Figure 6.51 *A localized or isolated two-dimensional electronic feature contains so-called 'surface states', in which there is a well-defined periodicity in a plane but not out of that plane. One can therefore assign an out-of-plane wavelength to such states of $\lambda_\perp = \infty$, as the states are never repeated in this direction. In reciprocal space, this translates to the out-of-plane wavevector $\Delta k_\perp = 2\pi/\lambda_\perp = 0$ and in this direction, the surface state is entirely* de*localized.*

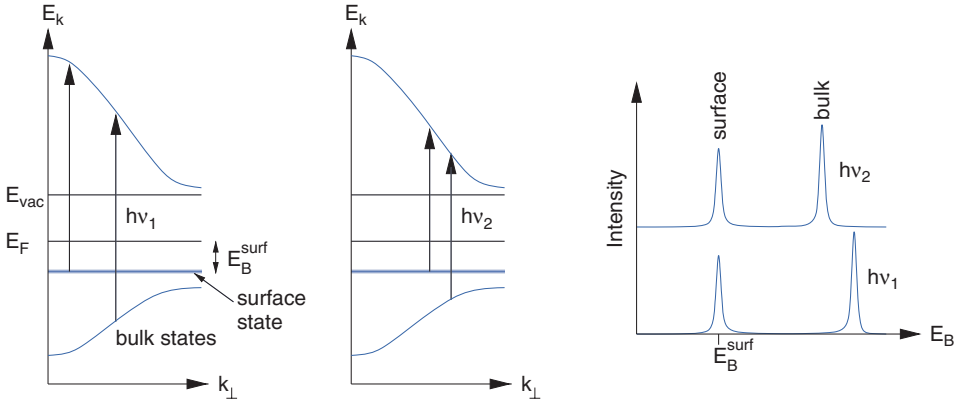

Figure 6.52 *Identification of surface states. Consider two binding energy spectra taken with two different photon energies $h\nu_1$ and $h\nu_2$ and for normal emission of the photoelectrons. Different photon energies probe different bulk states according to which $k_\perp$ corresponds to the energy difference between the occupied and unoccupied states. Because surface states have a constant binding energy E_B^{surf} as a function of $k_\perp$, their positions in the binding-energy spectra remain invariant and can thus be distinguished from bulk states.*

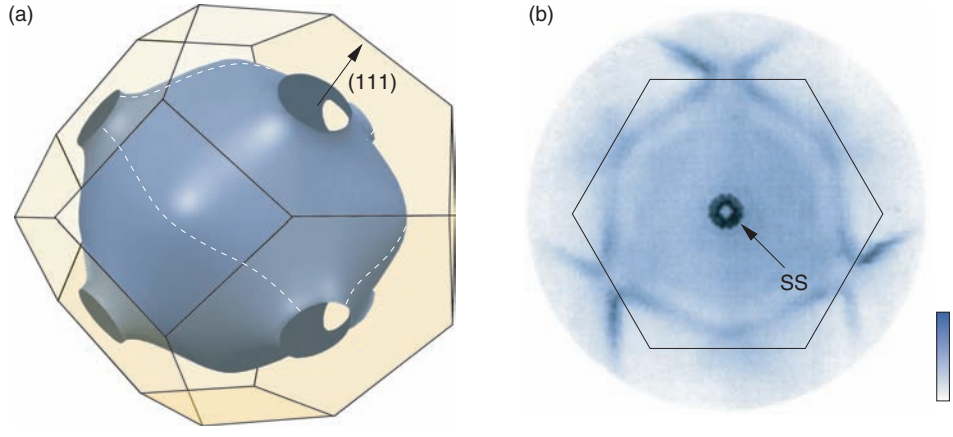

Figure 6.53 *Surface states in Cu(111). (a) The (111) crystallographic direction penetrates the Brillouin zone but not the Fermi surface – in this direction, copper has a band gap. (b) A Fermi surface map of the Cu(111) surface. The projection of the Brillouin zone on the (111) surface is shown partially as the hexagon. A high partial density of states at the Fermi energy can be recognized by the arcs of darker bands, also highlighted by the dashed white lines on the Fermi surface in (a). In the normal direction, ($k_x = k_y = 0$), a localized high-intensity signal is detected, associated with a surface state. Adapted from [51] with permission of Elsevier.*

6.9.2.6 Experimental Considerations

From the above description of the three-step model, the basic idea of ARPES should be clear–by determining both the kinetic energy and three orthogonal momentum components of the wavevector of the emitted electrons, one can infer the energy as a function of momentum of the occupied states within the solid, in other words, the material's electronic band structure.

How does one achieve angular resolution in ARPES? For electron-energy analysers with small entrance apertures (shown as S in Figure 4.43), this is achieved either by moving the orientation of the sample relative to the stationary analyser (normally a concentric hemispherical analyser, CHA, see Section 4.6.7), or alternatively by mounting the detector on a two-circle goniometer. The problem with the first method is designing a sample holder with sufficient angular precision. Presently, the highest resolution sample goniometer is the CARVING manipulator, a collaborative development between the Paul Scherrer Institut, Switzerland, and the University of Amsterdam, Holland.

The second possibility, of keeping the sample fixed and moving the electron analyser, poses the problem of maintaining ultra-high vacuum conditions when repositioning the analyser. The basic geometry of an angular-resolved photoelectron experiment is shown in Figure 6.54.

In modern CHAs, data acquisition is accelerated by using area electron detectors such as multichannel plates (shown in Figure 4.44), such that a range of angles (and thereby momenta) can be recorded simultaneously.

6.9.2.7 Worked Example – The Electronic Bandstructure of Cuprate Superconductors

It was long well known and theoretically explained that upon cooling down conventional superconductors to below their superconducting transition temperature, an energy

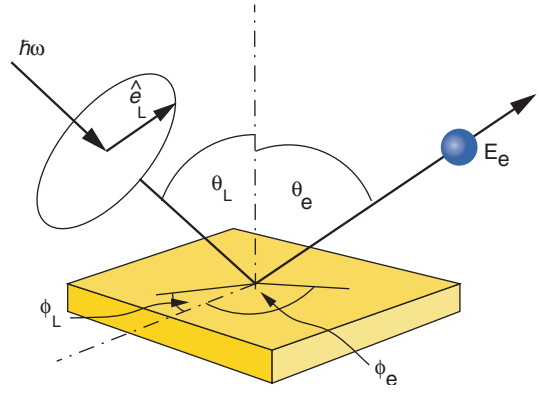

Figure 6.54 *Schematic diagram of an angular-resolved photoelectron experiment, including the possible experimental variables. Those relating to the incident beam are suffixes with L, while those relating to the electron are suffixed with e. The unit vector* ***e*** *is a polarization vector. The angles are normally referred to a high symmetry axis of the crystal, given here by the dash-dot line on its surface.*

gap appears in the electronic excitation spectra around the Fermi level. Soon after the discovery of high-temperature (high T_c) superconductivity (HTSC) in cuprate compounds in 1986, evidence started to grow for the appearance in these materials of a so-called 'pseudogap' above the transition temperature.

In 1996, the first direct spectroscopic evidence for the pseudogap was reported using ARPES on the underdoped material $Bi_2Sr_2CaCu_2O_{8+\delta}$ (Bi2212) [54].

Samples with different doping levels (determined by the oxygen content δ) were introduced into an ultra-high vacuum ARPES chamber and cleaved in-situ along the weakly bound BiO planes. The photon energy was either 19 or 22 eV, and the resolution $E/\Delta E \approx 1000$.

Because the pseudogap was found to have a size similar to the spectroscopic resolution, the inflexion points of the EDCs were taken as reference points, while an absolute scaling was made possible by recording an EDC for a polycrystalline metallic Pt-foil in electrical contact with the sample. According to the probed position in momentum space, the EDC of the sample shows a different response. In the superconducting phase below T_c, the material exhibits a single, precisely defined superconducting state associated with a $d_{x^2-y^2}$ superconducting wave. For all other k-space directions, there is a gap at the Fermi level. However, above T_c, arc segments begin to grow on the Fermi surface, though remain incomplete until a temperature T^*, when the material becomes metallic. Interestingly, increasing the underdoping, while reducing the transition temperature to superconductivity, *increases* the value of T^*. Based on these studies, a phase diagram for Bi2212 could be proposed (see Figure 6.55).

This ARPES study was the first direct evidence of the pseudogap phase between T_c and T^* in HTSCs. An understanding of this is thought to be critical in a full description of the enigma that remains high-temperature superconductivity.

6.9.3 X-ray Photoelectron Spectroscopy

The chemical composition of the surface region of a material can be determined down to a fraction of a per cent using x-ray photoelectron spectroscopy (XPS). In addition to providing information on the elemental constituents, XPS also yields information on the type of chemical bonds in the material. Indeed, XPS is alternatively called ESCA, or electron spectroscopy for chemical analysis.

6.9.3.1 The Chemical Shift

The chemical bond between atoms or ions arises due to the rearrangement of the spatial distribution of one or more of the valence electrons – for example, in covalent bonds pairs of electrons between atoms are shared, while in ionic bonds there is to a significant extent a transfer of one or more electrons from one atom to another, resulting in the formation of oppositely charged and Coulomb-attracted ions.

As an example, a 1*s* electron from oxygen ejected by an x-ray photon has a binding energy, which depending on the local chemical environment, can differ by as much as 10 eV around approximately 543 eV. As a rule of thumb, elements in a higher oxidation state have electrons with higher binding energies, and vice-versa. This is because atoms (or ions) in higher oxidation states have less electrons within their immediate neighbourhood, hence the electrostatic repulsion between the ejected electron and those electrons

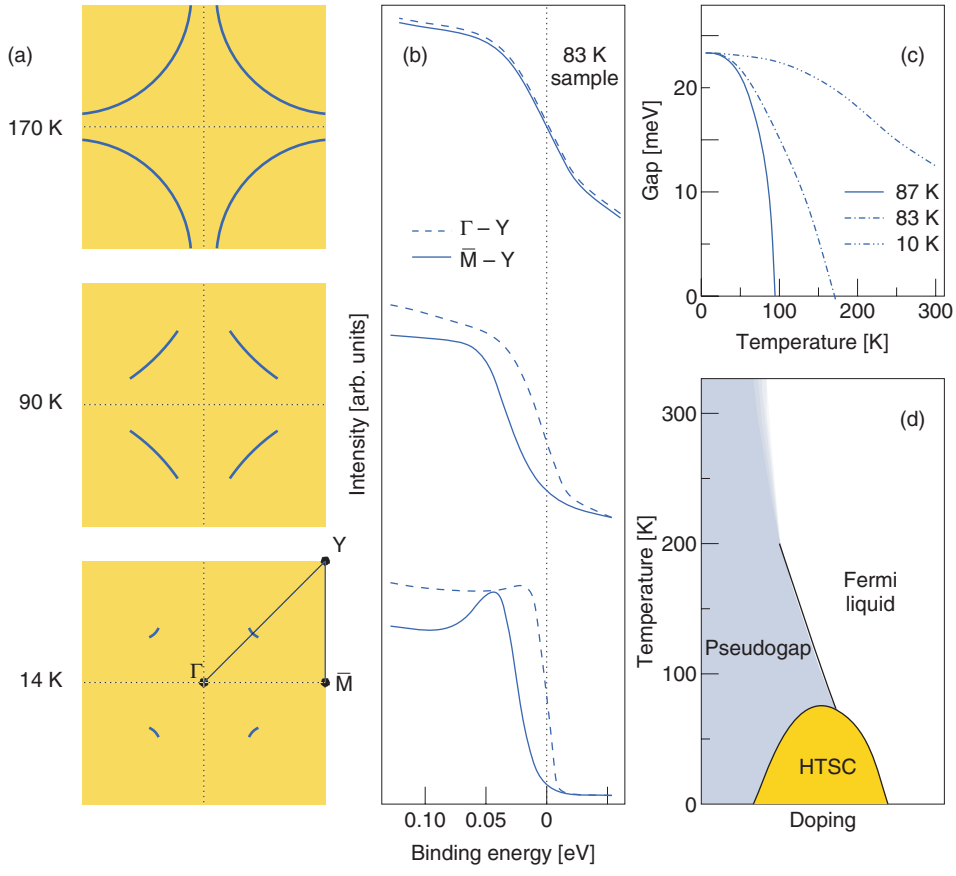

Figure 6.55 *Direct evidence of a pseudogap in the high-temperature superconductor* $Bi_2Sr_2CaCu_2O_{8+\delta}$ *above the superconducting transition temperature. (a) A slightly underdoped sample with a transition temperature* $T_c = 83$ *K was investigated at three different temperatures and in-plane momentum vectors within the Brillouin zone. At 14 K in the superconducting phase, there exists only one state at the Fermi surface (or four equivalent states) associated with so-called d-wave superconductivity. Above* T_c *arc-sections cross the Fermi surface, while in between these, there exists the pseudogap. At still higher temperatures, the arc-sections close to form a metallic-like phase called a 'Fermi liquid'. (b) Signal intensity was seen at the Fermi level when EDCs were recorded along the* $\Gamma - Y$ *line at all temperatures. EDCs along the* $\overline{M} - Y$ *direction, however, show a gap to well above* T_c*, indicative of the pseudogap. Above* T^**, the pseudogap closes as the arc-sections join and the material becomes a Fermi liquid. (c) Samples with increasing underdoping have a lower* T_c*, but a higher* T^**. (d) From the ARPES data, a phase diagram for Bi2212 could be sketched. The boundary between the pseudogap and Fermi liquid phases becomes poorly defined as the width of the EDCs becomes broader with underdoping. Adapted from [54] with permission of Macmillan Publishers Ltd.*

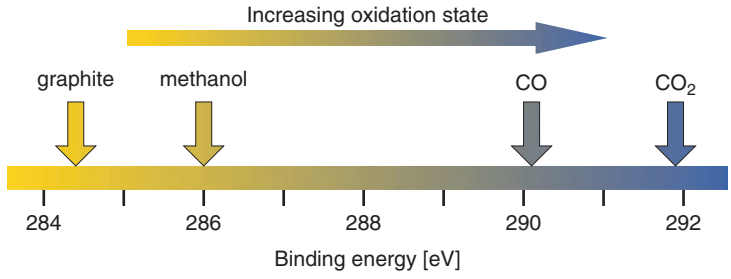

Figure 6.56 *A plot of the shift in binding energies of the C 1s core electron in different carbon-containing compounds.*

remaining is smaller.[13] The 1*s* binding energy of carbon for different oxidation states is shown in Figure 6.56. Chemical shifts can be detected with a resolution of 50 meV or better, using third-generation sources.

Although XPS setups at synchrotrons can be used for high-resolution chemical analysis, this is nowadays normally considered to be a somewhat wasteful use of beamtime. One clear advantage of synchrotron sources, other than the increased brightness and sensitivity, is the fact that the beam can be focused down to approximately 20 nm, with the promise of further improvement of a factor of approximately five in the near future. This allows one to perform two-dimensional maps in 'photoelectron microscopy' [55].

6.9.3.2 Photoelectron Diffraction

X-ray photoelectron diffraction (XPD) is a technique used to obtain direct structural information in the neighbourhood of the photoelectron emitter in single crystals. It depends on the strong forward-scattering nature of photoelectrons by neighbouring atoms, as first identified and reported by Siegbahn *et al.* in 1970 [56] and developed primarily by Fadley, starting in 1971 [57]. XPD is now an established and invaluable tool for investigating surface absorbates and heteroepitaxial thin films [58, 59].

In an XPD experiment the outgoing photoelectron wave can reach an electron detector on a direct path or can be scattered by the atoms surrounding the emitter. The intensity measured at the detector is given by the interference of the direct and the scattered wave parts and changes as a function of emission direction (θ_s in Figure 6.57) or electron kinetic energy. These modulations can be used to determine the structure around the emitter.

Prediction of the angular distribution of the signal in photoelectron diffraction is much simplified by the well-known phenomenon of the enhanced forward elastic scattering of photoelectrons (see Figure 6.57). Generally, emitted electrons are always scattered by the atoms on their way to detector, this scattering being rather complex and depending on both the kinetic energy and angular momentum character of the outgoing photoelectron wave. However, for high-energy electrons (in the range from a few hundred electronvolts to several keV), elastic scattering from the electronic clouds around adjacent atoms is associated with only a very small phase shift. Hence in the forward direction interference

[13] A useful mnemonic is OIL RIG: *O*xidation *I*s *L*oss, *R*eduction *I*s *G*ain (of electrons).

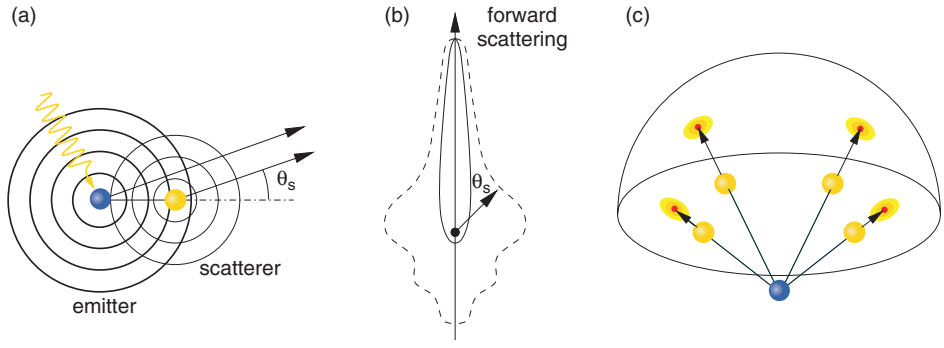

Figure 6.57 *The scattering processes in x-ray photoelectron diffraction. (a) The outgoing wave of an ejected photoelectron can be either direct or be scattered elastically by neighbouring atoms. (b) Scattering has a pronounced forward contribution for electrons with energies of several hundred to several thousand eV. Towards the lower limit of this range (dashed polar distribution), forward scattering is less pronounced than at higher electron energies (solid line). (c) In a crystal, pronounced forward scattering has the effect that scattering along high-symmetry crystal axes is therefore much enhanced.*

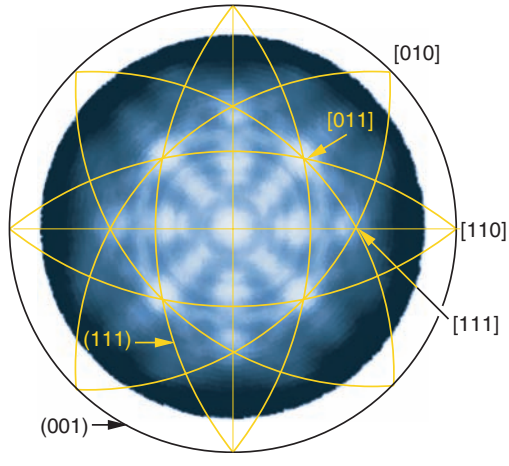

Figure 6.58 *XPD pattern of a heteroepitaxially grown film of TiN on Si(001). Stereographic projection of the angular intensity distribution of the N 1s signal ($E_B = 402.6\,eV$) up to a polar angle of 78°. Also shown are some low-index crystallographic directions (in square brackets) and planes (in round brackets). Adapted from [60] with permission of American Institute of Physics.*

is constructive, while at larger angles partial destructive interference causes the scattering amplitude to drop off sharply. This results in an amplification of the scattering intensity along low-index crystallographic axes (see Figure 6.58).

At first glance, XPD seems similar to EXAFS. There are, however, fundamental differences. While in EXAFS, interference with backscattered wavelets of the emitted photoelectron leads to variations in the final-state wave amplitude and thereby a

modulation of the total absorption cross-section (the EXAFS signal, see Section 6.5), in XPD the emitted photoelectrons are detected and one therefore measures a partial, angle-derivative cross-section associated with elastic scattering only. The strong interference effects in XPD mean that the intensity modulations as a function of emission angle are an order of magnitude more pronounced than in EXAFS modulations. While EXAFS yields information on interatomic distances via Fourier techniques, XPD provides direct real-space information on the *directions* between neighbours. Note also that EXAFS can be performed on bulk and polycrystalline samples, often in ambient conditions, while XPD only works for the near-surface region of single-crystal samples in vacuum.

6.9.3.3 Worked Example – Chiral Recognition of Organic Molecules on Metal Surfaces

Molecular recognition is crucial in nature – in biological processes, a lock-and-key process allows a receptor molecule to discriminate a target molecule in the organic soup of cell nuclei. Chiral recognition of enantiomers (mirror-image molecules that cannot be mapped on to one another by translation and rotation) is required in nature for the separation of, for example, the amino acid D-cysteine from L-cysteine. This essential ability, if it could be controlled artificially in heterogeneous catalysis, would revolutionize synthesis in the pharmaceutical industry.

It should be logically clear that for chiral heterorecognition, both the receptor *and* target must be chiral. If the receptor is a crystal, chirality can be achieved by preparing an asymmetrically cut surface, such as shown in Figure 6.59(a). In such cases, the absorption energies of the enantiomers will in general be different and involve different absorbate (target)–surface (receptor) orientations. In this example, it is shown using XPD

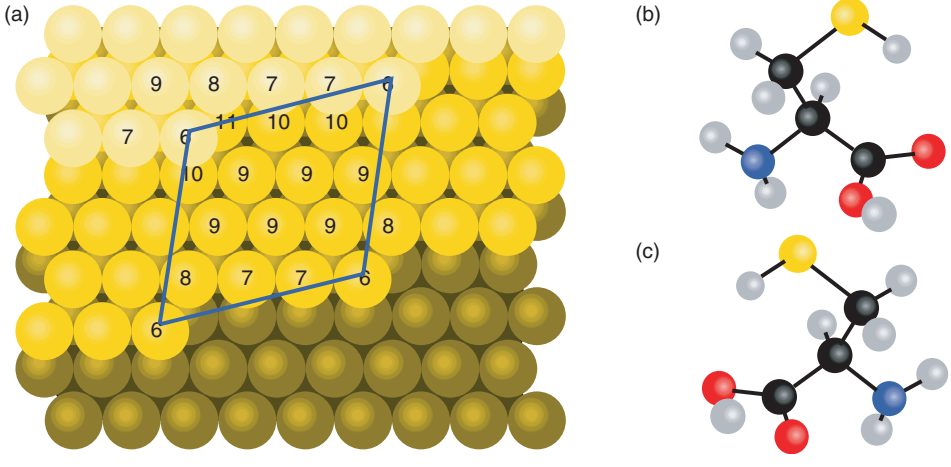

Figure 6.59 *(a) The surface unit cell of Au(17 11 9), which has a miscut angle of 15.4° to the Au(111) direction. Coordination numbers of the surface atoms within the surface cell are marked. The kink atom (c = 6) has the lowest coordination. The structures of D- and L-cysteine are shown in (b) and (c), respectively. Black = C, Red = O, Blue = N, Gold = S, Grey = H. Adapted from [61] with permission of the American Physical Society.*

that L- and D-cysteine absorb in two distinct and non-mirror-symmetric conformations on the chiral Au(17 11 9) surface [61].

Two depositions were performed, one of D-, the other of L-cysteine, each to marginally less than one monolayer, that is, one cysteine molecule per surface unit cell, as shown by the parallelogram in Figure 6.59(a). XPS measurements were then performed to identify how the cysteine binds to the gold surface: The S $2p$ binding energy was determined to be 162 eV, indicating deprotonation of the sulfur atom, which then binds strongly to the gold substrate. In addition, the N $1s$ binding energy of 399.7 ± 0.2 eV coincides with that of NH_2 (compared to that of NH_3 at 1.7 eV higher binding energy), while evidence of a COOH group was provided by O $1s$ signature at 531 and 533 eV.

Angular-resolved diffraction patterns of L- and D-cysteine on Au(17 11 9) were measured at 768 eV photon energy. Measurements at $E_{kin} = 768 - 282 = 486$ eV, that is, that for C $1s$, showed no mirror symmetry, and thus indicated a defined orientation of the absorbates. Figures 6.60(a) and (b) show the N $1s$ emission patterns at $E_{kin} = 368$ eV for D- and L-cysteine, respectively. One single forward scattering peak, labelled X_D and X_L, in Figure 6.60(a) and (b) dominates the patterns. Importantly, the positions of these signals are different, proving that Au(17 11 9) provides an effective template for chiral recognition. The intensity of the peak indicates that the N $1s$-emitter scatterer is the central carbon atom (see Figures 6.59(b) and (c)), which thereby yields the absolute orientation of the two enantiomers on Au(17 11 9).

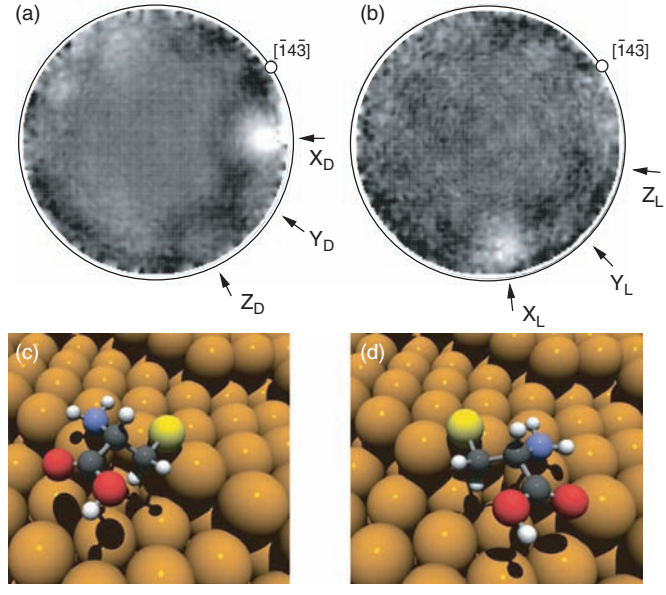

Figure 6.60 *N 1s emission patterns for (a) D-cysteine and (b) L-cysteine. The patterns are dominated by the features labelled X_D and X_L, associated with forward scattering with the central carbon atom nearly in plane. The two weaker features Y and Z are due to scattering of the N 1s wave with the COOH carboxyl group. The absorption structures and orientations of (c) D-cysteine and (d) L-cysteine, determined by XPD results and theoretical calculations. Adapted from [61] with permission of the American Physical Society.*

Theoretical calculations predict absorption structures having the experimentally determined orientations, and are shown schematically in Figures 6.60(c) and (d). The molecules bind to the lowest coordinated Au-atoms of the surface unit cell, in the region of the kink, where binding with potentially available electrons is most probable. Also, the calculations indicate that D-cysteine binds to the chiral gold surface 140 meV more strongly than does L-cysteine.

In conclusion, XPD was able to show that the chiral Au(17 11 9) surface binds differently to the two chiral enantiomers of cysteine. Such selectivity, if optimized, may have great commercial use in both the pharmaceutical industry and in catalysis.

6.9.3.4 Hard X-ray Photoelectron Spectroscopy

Curiously, the first x-ray spectroscopic measurements, by Siegbahn and co-workers in the 1950s, were carried out using photon energies in the range of 5 to 8 keV. However, a desire to improve the spectroscopic resolution (see Equation (4.30)) and scientific interest in the surface properties of materials rapidly led to a shift to lower energies, where the surface sensitivity is highest, of the order of 1 to 2 nm (see Figure 6.7).

As we have already argued in Section 5.13, the surface region of a material, that is, that which structurally or electronically differs from the bulk material, can extend down to several monolayers or in other words, the same depth as the photoelectron escape depth between approximately 30 to 2000 eV. Hence, conclusions drawn about the chemistry or electronic structure of a sample using photoelectron spectroscopy in this energy range may not in fact represent the bulk properties.

Therefore, those experiments which are designed to investigate bulk properties, or embedded regions deeper than a few monolayers, need to use photoelectrons with larger escape depths. This implies a move to higher x-ray energies, as suggested by Lindau *et al*. in 1974 [62]. For a long time, this idea was not seriously pursued, as it was thought that the photoelectron intensities would be too low – if we refer back to Figure 2.16, showing the absorption coefficient as a function of energy, we remember that this drops off as $1/(h\nu)^3$. In addition, the sensitivity of the electron-energy analyser drops off as E^{-1}. Hence, the photoemission signal intensity for experiments using photons with energies far in excess of the binding energy of the quantum state of interest is very small.

With the development of highly brilliant, third-generation synchrotron sources, however, the idea of using hard x-rays (between 5 and 15 keV) in photoelectron spectroscopy (HAXPES) has been recently revived, and now several groups are developing this technique [63]. In this energy range, the escape depth of the photoelectrons increases to approximately 10 nm at 5 keV and to some 25 nm at 15 keV, far larger than the extent of typical surface regions.

In conventional XPS carried out with soft x-rays at synchrotron facilities, the spectral resolution is typically $\Delta E = 0.1$ to 0.2 eV, which is necessary in order to unambiguously distinguish chemical shifts corresponding to different chemical or valence states (see Figure 6.56). The same absolute energy resolution is also desirable in HAXPES, but, as we see from Equation (4.30), this sets stringent specifications on the one hand on the detector geometry and on the accuracy of setting the pass energy, which in HAXPES must be much smaller than the original kinetic energy of the electrons as they emerge from the sample surface. Most commercial electron analysers designed for conventional

XPS are unable to detect energies much above 5 keV, although some do extend as far as 12 keV. At the time of writing, electron energies as high as 15 keV could only be detected using custom-made equipment.

On the other hand, the degree of monochromation of the synchrotron radiation must be of the order of $\Delta\lambda/\lambda = 10^{-5}$. The intrinsic width of the most common crystal, Si(111), used in hard x-ray monochromators is, at around 10^{-4}, one order of magnitude larger. The width can be decreased to approximately 2×10^{-5} by using the (333) or (400) reflections, although this is accompanied by a reduction in flux. In this case, the resolution at 10 keV is approximately 0.2 eV

6.9.3.5 Worked Example – The Interface Between $SrTiO_3$ and $LaAlO_3$

In 2004, it was reported that the interface between the two insulators $SrTiO_3$ (STO) and $LaAlO_3$ (LAO) is a conducting, quasi-two-dimensional electron gas (2DEG). This exciting phenomenon immediately provoked a concerted effort to understand the physics behind it. There soon arose several contentious claims, including charge transfer across the interface allowed by a change in the Ti-ion valency (from 4+ to 3+); intermixing of STO with LAO at the interface; and the proposal that the conductivity is merely due to the formation of oxygen vacancies in the STO substrate. One of the most discussed issues has been the thickness of the 2DEG, and whether it contains mixed-valence Ti-ions.

This system was investigated using HAXPES at the KMC-1 beamline at BESSY in Berlin. The photon energy was 3000 eV, used to excite the Ti 2*p* electrons at approximately 460 eV binding energy. The experimental resolution was 0.5 eV. Although this photon energy is fairly low for HAXPES, the signal intensity is concomitantly higher. This was particularly important here, where a signal originating from perhaps significantly less than one monolayer was being sought after. The corresponding escape depth for photoelectrons with kinetic energies of approximately $3000 - 460 = 2500$ eV is ~ 5 nm. The LAO film was 2 nm, hence the interfacial region could be easily probed at this energy [64].

The effective depth which was being probed could be controlled by changing the angle between the sample normal and the analyser axis (see Figure 6.61(a)), which is proportional to $\cos\theta$. The results are shown in Figure 6.61(b). The Ti 2*p* doublet at 465.5 and 459.5 eV originating from the tetravalent Ti present in most of the STO dominates the signal. The splitting into a doublet is due to spin-orbit coupling between the orbital and spin angular momenta, and is a standard signature for *p*-electrons. However, importantly, there is also a small peak on the low-binding energy flank of the main signal, shifted down in binding energy from the latter by 2.2 eV. This lower energy is indicative of a more reduced state, and agrees very well with the difference between the Ti-signal in Ti^{3+}- and Ti^{4+}-compounds.

Importantly, this weak signal becomes stronger as the emission angle θ increases, which means that most of the Ti^{3+}-species must lie closer to the surface than the maximum escape depth of approximately 5 nm. Indeed, as the uppermost 2 nm is LAO, it can be stated with confidence that the trivalent Ti-ions lie no deeper than one to two nanometers from the interface. The extra electron resides in the otherwise empty 3*d* valence shell. This experiment was the first direct experimental proof of the role of mixed valence in the observed conductivity and its limited extent to the immediate neighbourhood of the interface.

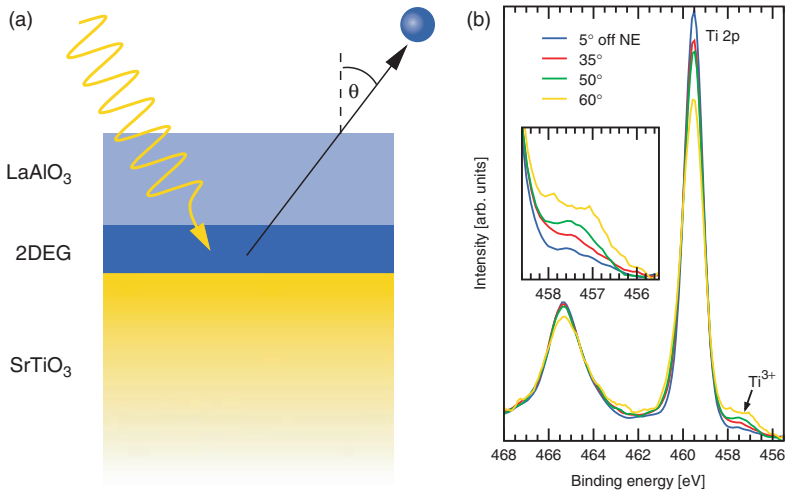

Figure 6.61 *HAXPES of the conducting interface between $SrTiO_3$ and $LaAlO_3$. (a) X-rays of 3 keV photon energy impinge on a thin film of $LaAlO_3$ on a $SrTiO_3$ substrate. The effective escape depth of the photoelectron can be probed by varying the angle θ between the sample normal and the analyser detector entrance axis. (b) The change in the Ti 2p signal as a function of θ. The small signal at 2.2 eV lower binding energy, associated with trivalent titanium is shown in detail in the inset. Adapted from [64] with permission of the American Physical Society.*

6.10 Concluding Remarks

Spectroscopic techniques that use synchrotron radiation have the unique advantage that the incoming x-rays can be tuned and/or scanned in energy. This adds an invaluable new dimension to experiments excluded to those based on laboratory x-ray sources and has been exploited to full advantage in many of the above-described methods, not least in XAS, ARPES and RIXS.

The high brilliance of third-generation sources, coupled with the ability to focus down to the nanometer scale, has opened up new vistas in synchrotron spectroscopy. Enormous progress has also been made in the performance of fast, energy dispersive detectors, and it is expected that the range of experiments available will continue to burgeon as rapidly as has been experienced in the last decade.

Spectroscopy using x-rays will remain the most prominent and sought after general synchrotron technique because of the detailed information that can be gleaned about the chemical and electronic properties of matter. In an era when ever more novel and unexpected materials are being tailored and synthesized, synchrotron spectroscopy will continue to play a central role in materials research.

References

[1] J. J. Rehr and R. C. Albers, "Theoretical approaches to x-ray absorption fine structure," *Rev. Mod. Phys.*, 72, pp. 621–654, 2000.

[2] J. J. Rehr and A. L. Ankudinov, "Progress in the theory and interpretation of XANES," *Coord. Chem. Rev.*, 249, pp. 131–140, 2005.

[3] G. Meitzner, G. H. Via, F. W. Lytle, and J. H. Sinfelt, "Analysis of x-ray absorption-edge data on metal-catalysts," *J. Phys. Chem.*, 96, pp. 4960–4964, 1992.

[4] A. Puschmann, J. Haase, M. D. Crapper, C. E. Riley, and D. P. Woodruff, "Structure determination of the formate intermediate on Cu(110) by use of x-ray-absorption fine-structure measurements," *Phys. Rev. Lett.*, 54, pp. 2250–2252, 1985.

[5] M. Sandstrom, F. Jalilehvand, I. Persson, U. Gelius, P. Frank, and I. Hall-Roth, "Deterioration of the seventeenth-century warship Vasa by internal formation of sulphuric acid," *Nature*, 415, pp. 893–897, 2002.

[6] D. E. Sayers, E. A. Stern, and F. W. Lytle, "New technique for investigating noncrystalline structures – Fourier analysis of extended x-ray absorption fine structure," *Phys. Rev. Lett.*, 27, pp. 1204–1207, 1971.

[7] M. Newville, S. A. Carroll, P. A. O'Day, G. A. Waychunas, and M. Ebert, "A web-based library of XAFS data on model compounds," *J. Synchrotron Rad.*, 6, pp. 276–277, 1999.

[8] B. K. Teo and P. A. Lee, "Ab-initio calculations of amplitude and phase functions for extended x-ray absorption fine-structure spectroscopy," *J. Am. Chem. Soc.*, 101, pp. 2815–2832, 1979.

[9] P. A. Lee, P. H. Citrin, P. Eisenberger, and B. M. Kincaid, "Extended x-ray absorption fine-structure – its strengths and limitations as a structural tool," *Rev. Mod. Phys.*, 53, pp. 769–806, 1981.

[10] Q. Wang, C. X. Li, X. J. Niu, R. Shen, K. Q. Lu, S. Q. Wei, Z. H. Wu, T. Liu, Y. N. Xie, and T. D. Hu, "Structure of liquid krypton under atmospheric pressure: An EXAFS and reverse Monte Carlo study," *Phys. Rev. B*, 72, p. 092202, 2005.

[11] E. A. Stern and S. M. Heald, *Handbook of Synchrotron Radiation*, 1b. Amsterdam: North-Holland, 1983.

[12] B. K. Teo, P. A. Lee, A. L. Simons, P. Eisenberger, and B. M. Kincaid, "EXAFS - approximation, parameterization, and chemical transferability of amplitude functions," *J. Am. Chem. Soc.*, 99, pp. 3854–3856, 1977.

[13] P. A. Lee, B. K. Teo, and A. L. Simons, "EXAFS – new parameterization of phase-shifts," *J. Am. Chem. Soc.*, 99, pp. 3856–3859, 1977.

[14] G. Sarret, A. Manceau, D. Cuny, C. Van Haluwyn, S. Deruelle, J. L. Hazemann, Y. Soldo, L. Eybert-Berard, and J. J. Menthonnex, "Mechanisms of lichen resistance to metallic pollution," *Environ. Sci. Technol.*, 32, pp. 3325–3330, 1998.

[15] T. Matsushita and R. P. Phizackerley, "A fast x-ray absorption spectrometer for use with synchrotron radiation," *Jap. J. Appl. Phys.*, 20, pp. 2223–2228, 1981.

[16] P. G. Allen, S. D. Conradson, and J. E. Pennerhahn, "A 4-point crystal bender for dispersive x-ray absorption spectroscopy," *J. Appl. Crystallogr.*, 26, pp. 172–179, 1993.

[17] R. Frahm, "QEXAFS – x-ray absorption studies in seconds," *Physica B*, 158, pp. 342–343, 1989.

[18] R. Frahm, M. Richwin, and D. Lützenkirchen-Hecht, "Recent advances and new applications of time-resolved x-ray absorption spectroscopy," *Phys. Scripta*, T115, pp. 974–976, 2005.

[19] J. Stötzel, D. Lützenkirchen-Hecht, E. Fonda, N. De Oliviera, V. Briois, and R. Frahm, "Novel angular encoder for a quick-extended x-ray absorption fine structure monochromator," *Rev. Sci. Instrum.*, 79, p. 083107, 2008.

[20] P. Spieker, M. Ando, and N. Kamiya, "A monolithic x-ray monochromator with fixed exit beam position," *Nucl. Instrum. Methods A*, 222, pp. 196–201, 1984.

[21] R. Frahm, M. Nachtegaal, J. Stötzel, M. Harfouche, J. A. van Bokhoven, and J. D. Grunwaldt, "The dedicated QEXAFS facility at the SLS: performance and scientific opportunities," *AIP Conf. Proc.*, 1234, pp. 251–255, 2010.

[22] J. Dik, K. Janssens, G. Van der Snickt, L. van der Loeff, K. Rickers, and M. Cotte, "Visualization of a lost painting by Vincent van Gogh using synchrotron radiation based X-ray fluorescence elemental mapping," *Anal. Chem.*, 80, pp. 6436–6442, 2008.

[23] S. Eisebitt, J. Lüning, J. E. Rubensson, and W. Eberhardt, "Resonant inelastic soft X-ray scattering as a bandstructure probe: A primer," *Phys. Stat. Sol. B*, 215, pp. 803–808, 1999.

[24] A. Kotani and S. Shin, "Resonant inelastic x-ray scattering spectra for electrons in solids," *Rev. Mod. Phys.*, 73, pp. 203–246, 2001.

[25] L. J. P. Ament, M. van Veenendaal, T. P. Devereaux, J. P. Hill, and J. van den Brink, "Resonant inelastic x-ray studies of elementary excitations," *Rev. Mod. Phys.*, 83, 2011 (in press).

[26] M. Z. Hasan, E. D. Isaacs, Z. X. Shen, L. L. Miller, K. Tsutsui, T. Tohyama, and S. Maekawa, "Electronic structure of Mott insulators studied by inelastic X-ray scattering," *Science*, 288, pp. 1811–1814, 2000.

[27] G. Ghiringhelli, A. Piazzalunga, C. Dallera, G. Trezzi, L. Braicovich, T. Schmitt, V. N. Strocov, R. Betemps, L. Patthey, X. Wang, and M. Grioni, "SAXES, a high resolution spectrometer for resonant x-ray emission in the 400–1600 eV energy range," *Rev. Sci. Instrum.*, 77, p. 113108, 2006.

[28] V. N. Strocov, T. Schmitt, U. Flechsig, T. Schmidt, A. Imhof, Q. Chen, J. Raabe, R. Betemps, D. Zimoch, J. Krempasky, X. Wang, M. Grioni, A. Piazzalunga, and L. Patthey, "High-resolution soft x-ray beamline ADRESS at the Swiss Light Source for resonant inelastic x-ray scattering and angle-resolved photoelectron spectroscopies," *J. Synchrotron Rad.*, 17, pp. 631–643, 2010.

[29] J. P. Hill, D. S. Coburn, Y.-J. Kim, T. Gog, D. M. Casa, C. N. Kodituwakku, and H. Sinn, "A 2m inelastic X-ray scattering spectrometer at CMC-XOR, advanced photon source," *J. Synchrotron Rad.*, 14, pp. 361–365, 2007.

[30] J. Schlappa, T. Schmitt, F. Vernay, V. N. Strocov, V. Ilakovac, B. Thielemann, H. M. Ronnow, S. Vanishri, A. Piazzalunga, X. Wang, L. Braicovich, G. Ghiringhelli, C. Marin, J. Mesot, B. Delley, and L. Patthey, "Collective Magnetic Excitations in the Spin Ladder $Sr_{14}Cu_{24}O_{41}$ Measured Using High-Resolution Resonant Inelastic X-Ray Scattering," *Phys. Rev. Lett.*, 103, p. 047401, 2009.

[31] K. P. Schmidt and G. S. Uhrig, "Spectral properties of magnetic excitations in cuprate two-leg ladder systems," *Mod. Phys. Lett. B*, 19, pp. 1179–1205, 2005.

[32] B. W. Batterman, "Effect of dynamical diffraction in x-ray fluorescence scattering," *Phys. Rev.*, 133, pp. A759–A764, 1964.

[33] B. W. Batterman and H. Cole, "Dynamical diffraction of x-rays by perfect crystals," *Rev. Mod. Phys.*, 36, pp. 681–717, 1964.

[34] S. K. Andersen, J. A. Golovchenko, and G. Mair, "New applications of x-ray standing-wave fields to solid-state physics," *Phys. Rev. Lett.*, 37, pp. 1141–1145, 1976.

[35] J. Zegenhagen, "Surface-structure determination with x-ray standing waves," *Surf. Sci. Rep.*, 18, pp. 199–271, 1993.

[36] J. Zegenhagen, B. Detlefs, T.-L. Lee, S. Thiess, H. Isern, L. Petit, L. Andre, J. Roy, Y. Y. Mi, and I. Joumard, "X-ray standing waves and hard X-ray photoelectron spectroscopy at the insertion device beamline ID32," *J. Electron Spectrosc. Rel. Phenom.*, 178, pp. 258–267, 2010.

[37] M. J. Bedzyk, D. H. Bilderback, G. M. Bommarito, M. Caffrey, and J. S. Schildkraut, "X-ray standing waves – a molecular yardstick for biological membranes," *Science*, 241, pp. 1788–1791, 1988.

[38] O. Dhez, H. Ade, and S. G. Urquhart, "Calibrated NEXAFS spectra of some common polymers," *J. Electron Spectrosc. Rel. Phenom.*, 128, pp. 85–96, 2003.

[39] B. Watts and C. R. McNeill, "Simultaneous surface and bulk imaging of polymer blends with x-ray spectromicroscopy," *Macromol. Rapid Commun.*, 31, pp. 1706–1712, 2010.

[40] E. Anders, "Pre-biotic organic-matter from comets and asteroids," *Nature*, 342, pp. 255–257, 1989.

[41] G. J. Flynn, L. P. Keller, C. Jacobsen, and S. Wirick, "An assessment of the amount and types of organic matter contributed to the Earth by interplanetary dust," in *Space Life Sciences: Steps toward origin(s) of life* (M. P. Bernstein, M. Kress, and R. Navarro-Gonzalez, eds.), 33 of *Adv. Space Res.*, pp. 57–66, 2004.

[42] S. Anders, H. A. Padmore, R. M. Duarte, T. Renner, T. Stammler, A. Scholl, M. R. Scheinfein, J. Stöhr, L. Seve, and B. Sinkovic, "Photoemission electron microscope for the study of magnetic materials," *Rev. Sci. Instrum.*, 70, pp. 3973–3981, 1999.

[43] Y. R. Ma, B. Aichmayer, O. Paris, P. Fratzl, A. Meibom, R. A. Metzler, Y. Politi, L. Addadi, P. U. P. A. Gilbert, and S. Weiner, "The grinding tip of the sea urchin tooth exhibits exquisite control over calcite crystal orientation and Mg distribution," *Proc. Natl. Acad. Sci. U.S.A.*, 106, pp. 6048–6053, 2009.

[44] J. Stöhr, "Exploring the microscopic origin of magnetic anisotropies with X-ray magnetic circular dichroism (XMCD) spectroscopy," *J. Magn. Magn. Mater.*, 200, pp. 470–497, 1999.

[45] J. Raabe, C. Quitmann, C. H. Back, F. Nolting, S. Johnson, and C. Buehler, "Quantitative analysis of magnetic excitations in Landau flux-closure structures using synchrotron-radiation microscopy," *Phys. Rev. Lett.*, 94, p. 217204, 2005.

[46] G. van der Laan, B. T. Thole, G. A. Sawatzky, J. B. Goedkoop, J. C. Fuggle, J. M. Esteva, R. Karnatak, J. P. Remeika, and H. A. Dabkowska, "Experimental proof of magnetic x-ray dichroism," *Phys. Rev. B*, 34, pp. 6529–6531, 1986.

[47] F. Nolting, A. Scholl, J. Stöhr, J. W. Seo, J. Fompeyrine, H. Siegwart, J. P. Locquet, S. Anders, J. Lüning, E. E. Fullerton, M. F. Toney, M. R. Scheinfein, and H. A. Padmore, "Direct observation of the alignment of ferromagnetic spins by antiferromagnetic spins," *Nature*, 405, pp. 767–769, 2000.

[48] C. S. Fadley, "X-ray photoelectron spectroscopy: Progress and perspectives," *J. Electron Spectrosc. Rel. Phenom.*, 178, pp. 2–32, 2010.

[49] C. Kittel, *Introduction to Solid State Physics*. John Wiley, 8th ed., 2005.

[50] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*. Brooks Cole, First ed., 1976.

[51] P. Aebi, J. Osterwalder, R. Fasel, D. Naumovic, and L. Schlapbach, "Fermi-surface mapping with photoelectrons at UV energies," *Surf. Sci.*, 307, pp. 917–921, 1994.

[52] A. Damascelli, Z. Hussain, and Z. X. Shen, "Angle-resolved photoemission studies of the cuprate superconductors," *Rev. Mod. Phys.*, 75, pp. 473–541, 2003.

[53] X. Y. Cui, E. E. Krasovskii, V. N. Strocov, A. Hofmann, J. Schäfer, R. Claessen, and L. Patthey, "Final-state effects in high-resolution angle-resolved photoemission from Ni(110)," *Phys. Rev. B*, 81, p. 245118, 2010.

[54] H. Ding, T. Yokoya, J. C. Campuzano, T. Takahashi, M. Randeria, M. R. Norman, T. Mochiku, K. Kadowaki, and J. Giapintzakis, "Spectroscopic evidence for a pseudogap in the normal state of underdoped high-T_c superconductors," *Nature*, 382, pp. 51–54, 1996.

[55] G. Margaritondo, "Photoelectron spectromicroscopy and spectronanoscopy at synchrotrons: Growing impact on life sciences and materials science," *J. Electron Spectrosc. Rel. Phenom.*, 178, pp. 273–291, 2010.

[56] K. Siegbahn, U. Gelius, H. Siegbahn, and E. Olson, "Angular distribution of electrons in ESCA spectra from a single crystal," *Phys. Lett. A*, A 32, pp. 221–222, 1970.

[57] C. S. Fadley and S. A. L. Bergström, "Angular distribution of photoelectrons from a metal single crystal," *Phys. Lett. A*, A 35, pp. 375–376, 1971.

[58] J. Osterwalder, P. Aebi, R. Fasel, D. Naumovic, P. Schwaller, T. Kreutz, L. Schlapbach, T. Abukawa, and S. Kono, "Angle-scanned photoelectron diffraction," *Surf. Sci.*, 331, pp. 1002–1014, 1995.

[59] D. P. Woodruff, "Surface structural information from photoelectron diffraction," *J. Electron Spectrosc. Rel. Phenom.*, 178, pp. 186–194, 2010.

[60] R. Timm, P. R. Willmott, and J. R. Huber, "Parallel epitaxy of TiN(100) thin films on Si(100) produced by pulsed reactive crossed-beam laser ablation," *Appl. Phys. Lett.*, 71, pp. 1966–1968, 1997.

[61] T. Greber, Z. Sljivancanin, R. Schillinger, J. Wider, and B. Hammer, "Chiral recognition of organic molecules by atomic kinks on surfaces," *Phys. Rev. Lett.*, 96, p. 056103, 2006.

[62] I. Lindau, P. Pianetta, S. Doniach, and W. E. Spicer, "X-ray photoemission spectroscopy," *Nature*, 250, pp. 214–215, 1974.

[63] L. Kover, "X-ray photoelectron spectroscopy using hard X-rays," *J. Electron Spectrosc. Rel. Phenom.*, 178, pp. 241–257, 2010.

[64] M. Sing, G. Berner, K. Goss, A. Mueller, A. Ruff, A. Wetscherek, S. Thiel, J. Mannhart, S. A. Pauli, C. W. Schneider, P. R. Willmott, M. Gorgoi, F. Schaefers, and R. Claessen, "Profiling the Interface Electron Gas of $LaAlO_3/SrTiO_3$ Heterostructures with Hard X-Ray Photoelectron Spectroscopy," *Phys. Rev. Lett.*, 102, p. 176805, 2009.

7

Imaging Techniques

7.1 Introduction

In 2006, Dr. Marco Stampanoni of the TOMCAT beamline of the Swiss Light Source, happened across a letter in *Nature* from 2004 describing the structure of fossil embryos dating from the Cambrian period, some 500 million years ago (MYA) [1]. The study was based on the collection of a large set of fossils covering the different stages of embryonic development of the early metazoa (multicellular animal) *Markuelia secunda*. Harvesting these specimens was a gargantuan effort – after about a dozen specimens had been discovered by several groups from sites in Australia, Siberia, and the United States, over 200 excellently preserved specimens were harvested from more than 8000 kg of material originating from the Hunan province of southern China [2]. This represents a mass yield of approximately 10^{-9}! Needless to say, each specimen is of enormous value.

In this early study, the samples were investigated using scanning electron microscopy. Clues as to the internal structures of the round, egglike specimens (with diameters of a fraction of a millimetre) were best revealed in those samples which had been fractured, although the extent to which the exposed internal morphology had been weathered remained a moot point. Because of the samples' high intrinsic value, there was a great deal of unwillingness, however, to slice open and effectively destroy the best-preserved samples in which the outside of the embryo 'egg' remained intact. Ironically, these samples yielded the least information... This is where Dr. Stampanoni entered the story.

Upon reading the article Dr. Stampanoni contacted Dr. Philip Donoghue of the University of Bristol, who was collaborating with Beijing University – the institute which had originally found the fossils – and offered him beamtime to investigate some of the intact samples. Using x-rays, the three-dimensional internal structure can be obtained in an entirely nondestructive manner – in other words, the samples could be sliced, but *virtually*, using computed x-ray tomography. The results that emerged (see Figure 7.1) were stunning in their detail, revealing features below the micronscale and providing unequalled insights into the body-plan assembly of these creatures which existed at a crucial time in the phylogeny (i.e. the history of evolutionary lineage) of multicellular animals [3].

An Introduction to Synchrotron Radiation: Techniques and Applications, First Edition. Philip Willmott.

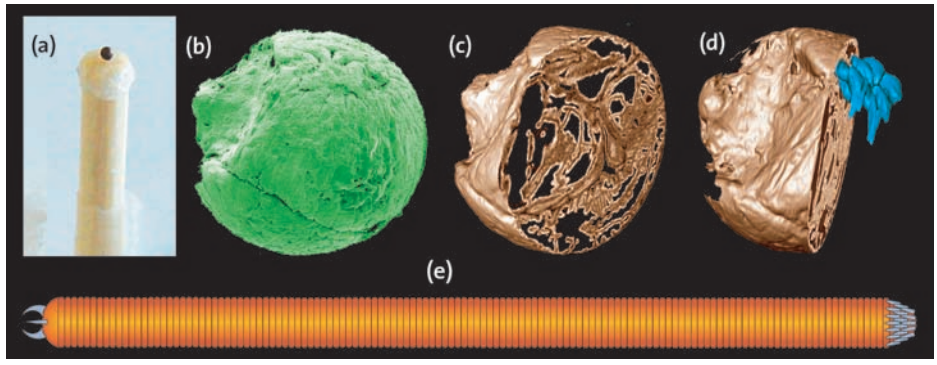

Figure 7.1 *Microtomography study of the 500-million-year-old fossil* Markuelia secunda. *(a) The diameter of the fossil embryo was less than half a millimetre. (b) Its external structure could be determined using conventional scanning electron microscopy. (c) The internal structure was revealed using x-ray tomography – here a 'virtual' cut is shown, while in (d), the tail of the embryo is highlighted in blue. (e) The supposed structure of* Markuelia. *The head is to the left, the tail to the right. Adapted from [3] with permission of Macmillan Publishers Ltd.*

It is the relative transparency of condensed matter to hard x-rays that makes them such a powerful tool for investigating the internal structures of objects in a nondestructive manner. Indeed, this benefit of x-rays was the first to be recognized and was what caused such a sensation immediately after their discovery more than a century ago (Figure 1.5). Radiographic imaging remains the most widespread application of hard x-radiation.

Imaging and microscopy techniques using x-rays extend across many x-ray methods. Many of these have already been discussed within the framework of other phenomena – for example, scanning transmission x-ray microscopy (STXM) and photoemission electron microscopy (PEEM) were introduced as two distinct imaging modes associated with x-ray absorption near-edge structure (XANES) in Sections 6.7 and 6.8. So, although x-ray imaging using synchrotron radiation is ubiquitous in third-generation facilities, we only discuss the areas of tomography and lensless imaging in this chapter.

7.2 Computed Microtomography

7.2.1 Introduction

The word 'tomography' is derived from the Greek words 'tomos' meaning 'to slice' and 'graph' meaning 'image'. Tomography in its most general sense is the generation of a three-dimensional image by analysing several transmission radiographic projections[1] taken of a specimen at different angles. This requires radiation which has an absorption length of the same order of magnitude as the size of the specimen under investigation. Nowadays, computed tomography (CT) with a resolution of approximately 500 µm is

[1] Standard x-ray plate images performed at dentists and emergency rooms at hospitals are the most commonly encountered type of transmission radiographic projections.

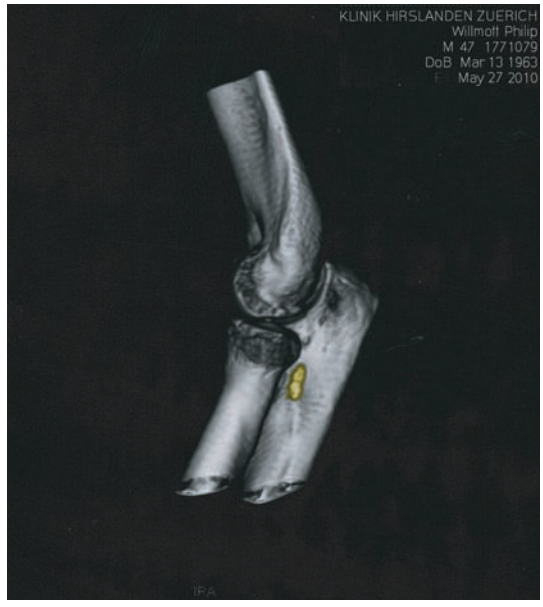

Figure 7.2 *A CT scan of a male elbow. The patient suffered an olecranic fracture, resulting in a small bone fragment, highlighted in gold, breaking off the outside of the elbow. The poor fellow.*

a standard diagnostic medical imaging tool in hospitals around the world (Figure 7.2). It is estimated that approximately 100 million CT scans per year are performed in the United States alone.

High-resolution x-ray tomographic microscopy (XTM) is a synchrotron-based non-destructive microscopy technique for creating quantitative, three-dimensional images of solid objects, including their detailed internal architecture, with a spatial resolution which can extend below the micron range [4–6]. It is useful for a wide range of materials, including inorganic substances such as rock, ceramic or metal, and also for objects made from bone and soft tissue. In recent years, optimization of XTM beamlines using high x-ray fluxes (at the partial cost of the relatively unimportant parameter of monochromacity, see below) has driven the speed of data acquisition so that it is now possible to record entire XTM scans in a few seconds or even, in favourable circumstances, on the subsecond timescale. Time-resolved biomedical studies using XTM of processes such as breathing or blood circulation are now becoming feasible.

The elements of a tomographic reconstruction, the so-called *voxels* (short for 'volumetric pixel'), provide a quantitative three-dimensional structure, which one can then view as one pleases, for example, sliced across any desired cross-section or rendered to highlight boundaries between materials of different density, to name just two possibilities. The contrast between voxels is in general higher than that between neighbouring pixels in a standard radiographic projection, as in the latter one records for any given pixel the integrated contribution of all the voxels that lie in path of the x-rays impinging on that pixel (Figure 7.3).

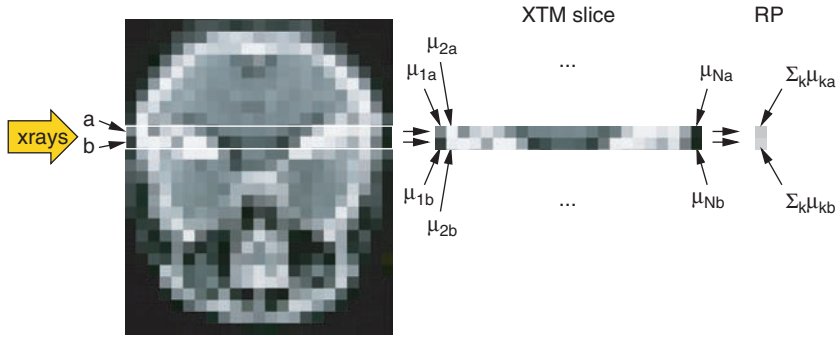

Figure 7.3 *Tomographic versus radiographic imaging. The contrast in attenuation strength between adjacent voxels μ_{ka} and μ_{kb} can be high in a reconstructed tomogram slice. Along the x-ray propagation direction, however, the contrast between the integrated attenuation strength of neighbouring pixels in the radiographic projection (RP) will be small. Note that in this figure the lighter the shade of grey of a pixel, the higher its associated attenuation coefficient μ.*

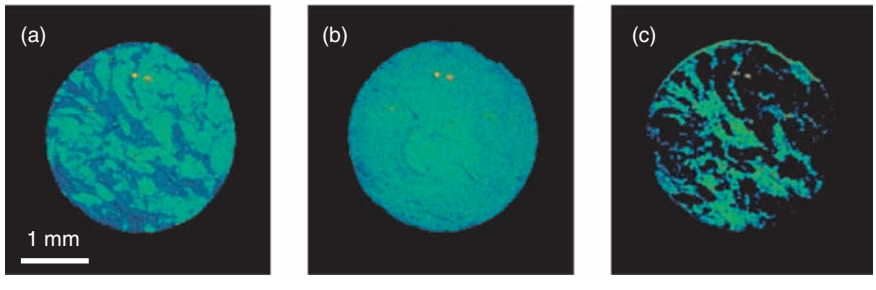

Figure 7.4 *Diffusion of a 'contaminant plume' of caesium in a geological sample. Contrast in the elemental distribution of Cs was enhanced by recording radiographic projections below and above the Cs K-edge at 35.985 keV. (a) Tomogram at 35 keV, (b) the same tomogram recorded at 37 keV and (c) the difference (a)–(b), clearly revealing the Cs-distribution. Courtesy Daniel Grolimund, Paul Scherrer Institut.*

The tunability of synchrotron radiation can be exploited to record tomograms above and below absorption edges of certain elements, thereby enhancing three-dimensional maps of their concentrations. An example of the diffusion of the contaminant caesium in a geological medium is shown in Figure 7.4.

7.2.2 General Concepts

A typical experimental setup for recording XTM data is shown schematically in Figure 7.5. A sample is rotated through 180° around an axis perpendicular to the quasiparallel incident x-ray beam. The transmitted signal is recorded at regular intervals during the rotation using a scintillator, magnification optics and CCD camera. The geometric resolution of the image is therefore determined by five factors, namely the point-spread function of the scintillator, the magnification factor of the microscope

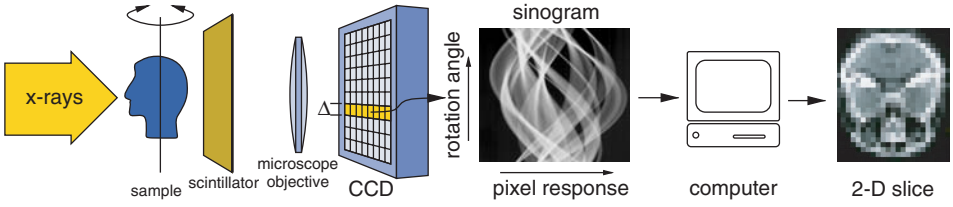

Figure 7.5 *The principle of x-ray computed tomography. A parallel beam of x-rays is allowed to pass through a sample, and the transmitted signal is detected by a scintillator, which is optionally magnified using a microscope objective lens onto a CCD array. The response from each row of pixels (height Δ) as a function of sample rotation angle, the so-called 'sinograms', are fed into a computer, from which two-dimensional tomograms are generated. Finally, the 3-D reconstruction is created from the stack of 2-D tomograms.*

objective optics M, the signal-to-noise ratio of the CCD image, and the linear pixel size Δ and point-spread function of the CCD camera.

The number of projections N of the sample which must be recorded within the angular range is given by the so-called 'sampling theorem' and is equal to

$$N = N_p \times \pi/2, \tag{7.1}$$

whereby N_p is the number of pixels in the row.

From each row in the CCD array a tomogram is generated. To obtain this, the response of the N_p pixels in the row as a function of rotation angle, known as a *sinogram*, is fed into a computer. An example of a sinogram of a simple function is shown in Figure 7.6.

The response of a detector pixel is proportional to the intensity of the transmitted x-rays that impinge upon it after travelling along a path $L = M\,\Delta z$ through an object, that is

$$I = I_0 e^{-\mu_1 \Delta z}\, e^{-\mu_2 \Delta z}\, e^{-\mu_3 \Delta z} \cdots e^{-\mu_M \Delta z} \tag{7.2}$$

$$= I_0 \exp\left(\sum_{k=1}^{M} \mu_k \Delta z\right). \tag{7.3}$$

We define the *projection measurement* p as

$$p = -\ln\left(\frac{I}{I_0}\right) = \sum_{k=1}^{M} \mu_k \Delta z \approx \int_L \mu(z) dz. \tag{7.4}$$

In other words, p for a given pixel at a given angle is the line integral of the object along the line L. Given these sets of line integrals for different angles, the task is to reconstruct the attenuation distribution throughout the entire object.

Each tomogram is reconstructed from its respective sinogram using one of several algorithms based on the 'Radon transform', developed by Johann Radon in 1917 [7], called filtered back projection (FBP). The basic idea behind FBP is simple to understand. Consider the 2-D absorbing disc shown in Figure 7.7. In one direction, the radiographic absorption profile is recorded. This projection is then run back through

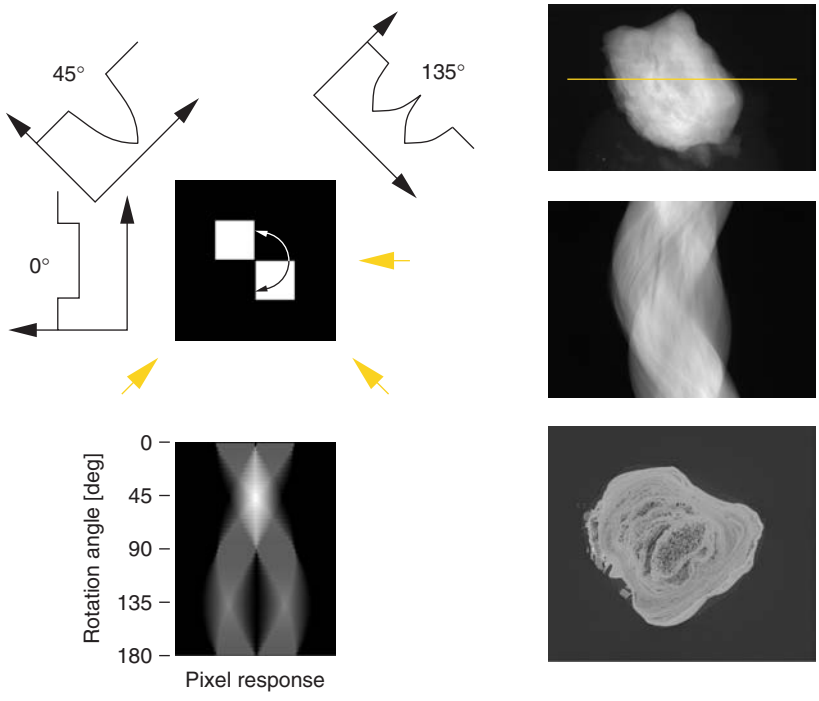

Figure 7.6 *The generation of sinograms. Left: example of a simple sinogram resulting from rotation through 180° of two absorbing blocks touching each other along diagonal corners. Transmission profiles of the three high-symmetry angles of 0, 45 and 135° are also shown. Adapted from http://en.wikipedia.org/wiki/Radon_transform. Right: a slice in the plane marked by the yellow line through a kidney stone results in the shown sinogram. From this, the tomogram at the bottom was reconstructed. The field of view was 3.7 mm. Courtesy Andreas Pasch, Inselspital, Bern. Analysis courtesy Federica Marone, Paul Scherrer Institut.*

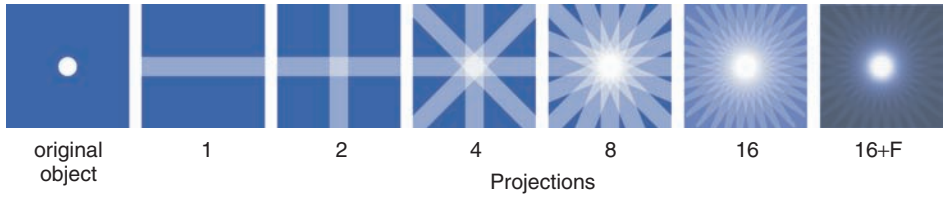

Figure 7.7 *Illustration of filtered back projection in XTM. The transmission profile of an absorbing disc is recorded in several directions. For each direction, the profile is run back through the image space. Those regions in the space where the back projections overlap add up. The more projections are added, the closer the summation of them resembles the original object. However, the reconstructed image suffers from blurring and starlike artefacts. In order to suppress this effect, each back projection is first filtered, whereby certain spatial frequencies like low-frequency blurring or high-frequency noise are suppressed.*

the image, whereby the projection intensity is evenly distributed among all the pixels along each ray path. This is then repeated for a set of projection angles spanning 180°, whereby for each angle the back projection is added to the image. The smaller the angular shift, the more closely the final overlapping set of back projections resembles the original object (Figure 7.7).

Although the general features of the original object can be retrieved in this manner, there is significant distortion due to overlapping projections producing starlike artefacts and a blurring, which becomes increasingly evident near the centre of the reconstruction. These artefacts can be suppressed by careful filtering of the absorption profiles. This is carried out by Fourier analysis – each profile is Fourier-transformed and then treated with a filter which removes certain ranges of the Fourier components. Lower frequency components tend to be responsible for blurring, while sharper, high-frequency, features might be suppressed in order to smooth statistical noise. A common type of filter is shown in Figure 7.8. In real-space, the filter function, or 'kernel', has the appearance of a distorted 'w', which is convoluted with the back projection profile. In Fourier space, the FT of the profile is therefore *multiplied* by the FT of the kernel (according to the convolution theorem, see Section 5.5.1). The detailed shape of the kernel's FT determines which frequencies are filtered out.

Advances in the processing power of modern computers now allow tomogram slices to be rapidly and more directly constructed using fast-Fourier transform algorithms, exploiting the so-called 'Fourier slice theorem'. Shown schematically also in Figure 7.9, this states that the one-dimensional Fourier transform of the projection of a two-dimensional object onto a line is equal to the one-dimensional slice through the centre of the two-dimensional Fourier transform of the same object. So, by Fourier transforming in one dimension of a given projection, we obtain a line in the 2-D-Fourier transform of the object. By collecting all the projections between 0 and 180° (the sinogram), we fill the entire Fourier-space representation of the object. It is then a simple procedure to first interpolate this data from its polar grid to a Cartesian coordinate system and then recover the object (i.e. the tomogram) by inverse Fourier transforming the data.

The final three-dimensional reconstruction consists of a stack of all the reconstructed tomograms. The resolution of the reconstruction is determined by the dimensions of

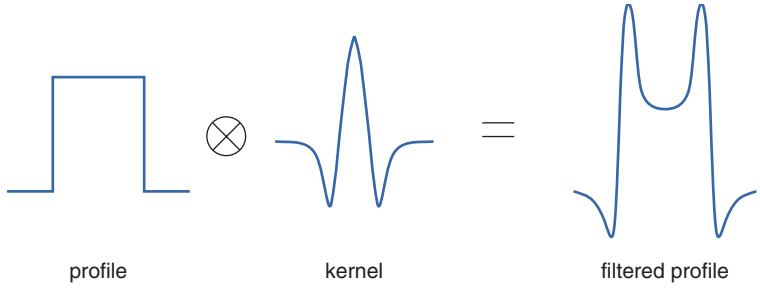

Figure 7.8 *Filters for back projections. A 'kernel' is convoluted with the original profile to produce a filtered profile which compensates the blurring due to overlapping of adjacent back projections. The detailed shape of the kernel determines the spatial frequencies most affected.*

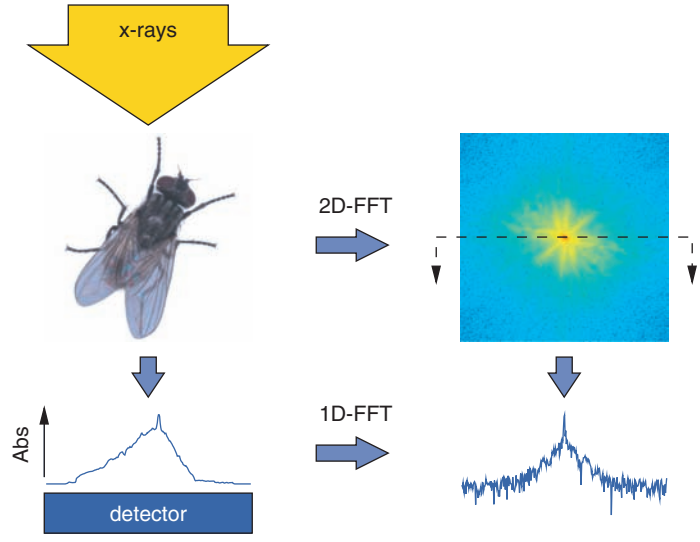

Figure 7.9 *The Fourier slice theorem. The one-dimensional (fast) Fourier transform (1-D-FFT) of a projection of an object yields the same result as the central slice of the 2-D-FFT of the same object.*

the voxels. Tomographic images comprise an array of these voxels on a regular three-dimensional grid, whereby each voxel is associated with a certain physical value, such as the average linear absorption coefficient within that voxel.

7.2.3 Practical Considerations

Each radiographic projection of a data set must be corrected for two factors: 'dark' noise generated by the detector electronics of the CCD camera in the absence of an x-ray beam, and inhomogeneities in the incident x-ray beam profile. The first factor is corrected by recording a 'dark-field' image in the absence of x-rays, while the second 'flat-field' image records the direct beam with the sample removed.

The sample sits on a stage, which rotates through 180°. While rotating, the rotation axis can drift laterally (the *eccentricity*), and can also *wobble*. Good modern goniometers offer eccentricities and wobbles of the order of a few hundred nanometers and a microradian, respectively.

XTM images can sometimes contain features which are artefacts of the technique. Regions in the centre of the object appear to have lower attenuation than identical regions near the periphery of the object if significantly polychromatic radiation ('pink beam') is used. The radial distribution of the photon energy of a polychromatic x-ray beam emanating from a bending magnet or wiggler at a synchrotron facility is such that the high energy photons are more concentrated in the centre of the beam than the soft x-rays, which become more dominant at the peripheries. Because the lower-energy x-rays in the beam are more strongly absorbed than are the higher-energy x-rays, the energy-distribution spectrum of the beam changes as it passes through the object. The emergent beam therefore contains a higher proportion of hard x-rays, and this

phenomenon is therefore known as 'beam-hardening'. This shift in the distribution of photon energies in the x-ray beam manifests itself during image reconstruction as an apparent decrease in attenuation in the centre of the object.[2]

In many XTM experiments, a high degree of monochromacity is not essential, while the need for high flux is paramount. Multilayer monochromators, described in Section 4.3.3, provide bandwidths of the order of $\Delta E/E = 0.01$, and the Bragg-reflected flux is some hundred times greater than when using single-crystal monochromators.

The speed of data acquisition is nowadays so high that XTM can be used to study dynamical processes which exhibit changes on the timescale of a second to a few seconds. Three interlinked factors determine the feasibility of using ultrafast XTM, namely the brilliance of the x-ray beam, the detector performance (sensitivity and readout time) and accurate synchronization of sample position and exposure time [8]. Instead of rotating the sample in a steplike manner and recording the radiographic projections between steps, the sample is continuously rotated and the rotational position must be recorded accurately 'on the fly' by an encoder triggered by the detector.

Because the readout times for CCD chips containing millions of pixels are relatively long, and certainly not in the desired ms-range, CMOS (complementary metal-oxide semiconductor) technology is an obvious alternative for fast-XTM. The advantage of CMOS sensors is their rapid and parallel readout times. This, however, requires more electronic architecture associated with the light-sensitive pixels, which therefore tend to be smaller for a given pixel repetition length and hence their light-gathering power is lower than that of CCDs. Lastly, the effects of centrifugal distortion of the sample must be considered. The outwards-acting acceleration of a sample rotating at 1 Hz around a radius of 1 mm is $0.04\,\mathrm{m\,s^{-2}}$. Although this is only about 4×10^{-3} of Earth's gravitational acceleration, it can produce unacceptable 'equatorial bulging' in soft-matter samples.

7.2.4 Phase-Contrast Tomography

7.2.4.1 Introduction

Using x-ray absorption as the sole source of contrast draws exclusively on ray optics to describe and interpret image formation. This approach ignores another, potentially more useful source of contrast – phase information. Phase-sensitive techniques, which can be understood using wave optics rather than ray optics, offer ways to augment or complement standard absorption contrast by incorporating phase information. In phase-contrast imaging, small phase shifts in different parts of an x-ray beam passing through a sample are converted into contrast in the projected image.

Clinical and biological studies benefit particularly from the development of phase-sensitive techniques. Absorption contrast works well in distinguishing between hard and soft tissue. Heavier elements – like calcium in bones and teeth – have a much higher absorption cross-section than the lighter elements that constitute soft tissues. However, in many clinical situations, such as mammography, there is a need

[2] Note that CT experiments that use broadband laboratory x-ray sources also suffer from beam hardening, although in this case, border regions tend to seem anomalously bright (in other words, they seem to have a *low* attenuation). The x-ray beam from a lab source is spatially homogeneous with regards to its spectral content, but because borders tend to present a shorter absorption path for the x-rays, a disproportionately high fraction of soft x-rays are transmitted.

to distinguish between different kinds of soft tissue – between tumors and normal tissue, for instance. Because the absorption is small, and differences in density and composition are slight, standard absorption x-ray imaging is not ideal.

The behaviour of x-rays as they travel through a sample is described using the complex index of refraction $n = 1 - \delta + i\beta$, discussed in Chapter 2. Here, β describes the absorption of x-rays, while the phase-shift term δ incorporates refraction. Phase contrast is useful in biological and medical studies because it falls off less quickly at higher energies than absorption contrast: $\delta \propto E^{-2}$, whereas $\beta \propto E^{-4}$. Because phase contrast relies only on refraction of x-rays and not on absorption, imaging can be performed at higher energies where the absorbed radiation dose is less, thereby reducing potential damage to tissues. At typical tomography hard x-ray energies, the magnitude of the phase-shift term can be up to 1000 times greater than the absorption term – of the order of 10^{-6}, compared to 10^{-9} (see Figure 7.10).

7.2.4.2 *Edge Diffraction and Refraction*

The effect of δ alone (ignoring absorption) is to cause the wavefronts within the medium to stretch apart by an amount $(1 - \delta)^{-1}$ as the phase velocity of the x-rays increases from c in vacuum to c/n in the medium. Hence two parallel x-ray beams which are initially in phase become out of phase by an amount

$$\phi = \frac{2\pi L\delta}{\lambda} \tag{7.5}$$

if one of these passes through a refracting sample of length L and refractive index n (Figure 7.11). Hence, the part of the beam propagating through the medium a distance $L = \lambda/\delta \sim 100\,\mu\text{m}$ for hard x-rays, will have advanced by a phase shift of 2π relative to the unperturbed part of the beam. This is approximately two orders of magnitude larger than the longitudinal coherence length of synchrotron radiation in the hard x-ray regime. One might conclude from this that any phase shifts will be scrambled and hence no useful information can be extracted. However, in coming to this conclusion, we have

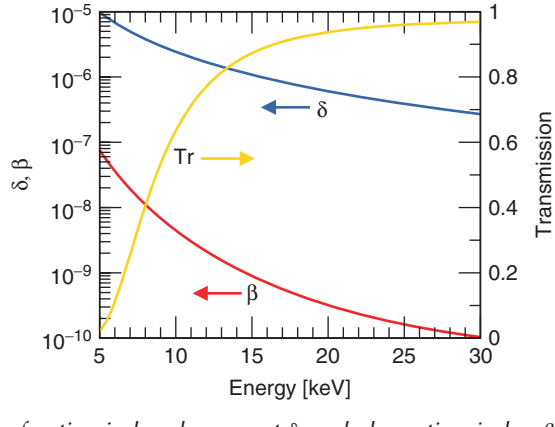

Figure 7.10 *The refractive index decrement δ and absorption index β for flesh from 5 to 30 keV, plus the transmission through a 1-mm sample. To calculate these curves, the chemical formula for flesh was taken as $C_{12}H_{60}O_{25}N_5$ and the density as 1.071 g cm^{-3}.*

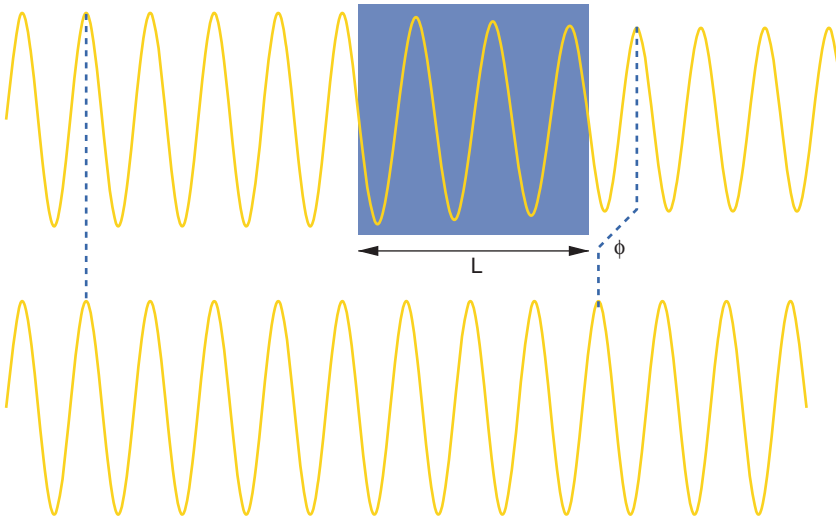

Figure 7.11 *The part of a beam which travels through a medium of refractive index n has a phase velocity $v_p = c/n > c$ – the peaks and troughs speed up. The result is that the emerging beam is phase shifted by an amount ϕ relative to the part of the beam travelling through vacuum.*

used ray optics. As we will now discover, refraction and diffraction effects couched in terms of wave optics reveal that phase-induced edge enhancement does indeed occur.

Here, we discuss edge-enhanced imaging due to the two related phenomena of Fresnel edge diffraction and refraction. Under the correct conditions, these effects cause the borders of objects to be highlighted, somewhat like a line drawing, even for object with very low absorption contrast. First, we should identify under which geometrical arrangement one observes and exploits edge enhancement and diffraction phenomena.

When a coherent wave is scattered by an object, interference between the differently scattered components produces diffraction features. This requires a certain propagation distance, however, for the different parts of the scattered wave to overlap. Observed close to the object, therefore, the scattered pattern still resembles the object, albeit with increasing evidence of diffraction phenomena with distance Z_0 from the object. This is the so-called Fresnel-diffraction regime, which extends until the diffraction features (fringes) are separated by a distance similar to the linear dimensions of the scattering object itself (a in Figure 7.12), that is, when

$$Z_0 \sim \frac{a^2}{\lambda}.$$

At very large distances $Z_0 \gg a^2/\lambda$, in the so-called far-field – or Fraunhofer-diffraction – regime the object profile is completely obscured by the diffraction effects, and the pattern is in fact the square of the Fourier transform of the object (assuming the object to be smaller than the spatial coherence of the incident beam). This is further discussed in Section 7.3.

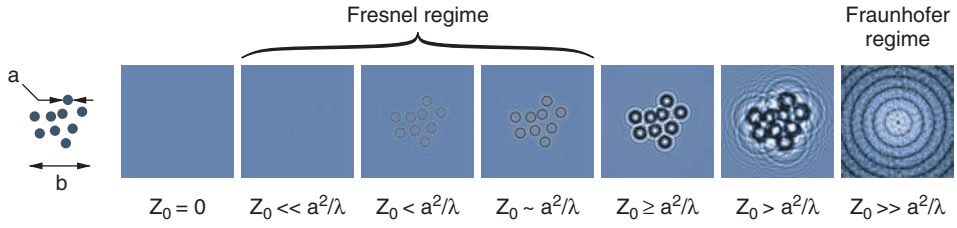

Figure 7.12 *The regimes of Fresnel and Fraunhofer diffraction. As Z_0, the distance from a transparent sample to the detector, is increased from zero the projected image of the object (here, nine spheres of diameter a spread over an area of characteristic length b) develops fringes at the boundaries between components with different refractive index. In the Fresnel regime, from $Z_0 > 0$ to $Z_0 \sim a^2/\lambda$ the object remains recognizable but becomes increasingly fuzzy at the boundaries due to the formation of fringes. At very large distances, $Z_0 \gg a^2/\lambda$, the detector records the far-field diffraction pattern of the object in the so-called Fraunhofer-diffraction regime. This is the condition for 'normal' diffraction and SAXS experiments. Courtesy Franz Pfeiffer, Technische Universität München.*

Fresnel edge diffraction is observed in both fully opaque and transparent objects. In order to understand its origin, we begin by introducing the concept of the Huygens–Fresnel construction. This imagines that every point along a wavefront (let us say, for convenience's sake, the wave maximum) is itself a point emitter of radiation of the same wavelength. In vacuum, interference between the spherically propagating wavelets emanating from the row of point sources causes all directions other than that perpendicular to the row to cancel out, thereby generating the next wavefront parallel to the previous (Figure 7.13(a)). If some sort of object lies in the path of part of the beam, the wavelets passing through that object assume a wavelength $\lambda/n > \lambda$ and this fraction of the beam is refracted and will subsequently interfere with other parts of the beam that travelled along a different path.

So, for example, the pattern on a detector at a distance Z_0 from a parallel beam of x-rays partially blocked by an opaque screen is not a step function, but instead exhibits interference fringes at the projection of the edge of the screen (Figure 7.13(b)). Although it goes beyond the scope of this book to derive an expression for this pattern (which involves the so-called 'Fresnel integrals', which must be evaluated numerically), we can gain a physical insight using the Huygens–Fresnel principle. If the propagation of a planar wavefront can be considered to be formed by a row of point emitters as argued above, then removing those emitters below the opaque screen means that they cannot interfere with those above the screen and the constructed wavefront must therefore deviate from a perfect plane wave, particularly in the neighbourhood of the screen edge. In contrast, the missing contribution to the planar wave for the constructed wave far above the screen is negligible, as here the amplitudes of the missing wavelets, which drop off inversely with the propagation radius, are very small.

Where can we expect the first interference minimum to occur? Consider Figure 7.13(c). We estimate x_1, the lateral displacement of the first dark fringe relative to the geometric projection of the screen edge, by assuming that here a ray A associated with the forward-propagating planar wavefront interferes destructively with ray B originating

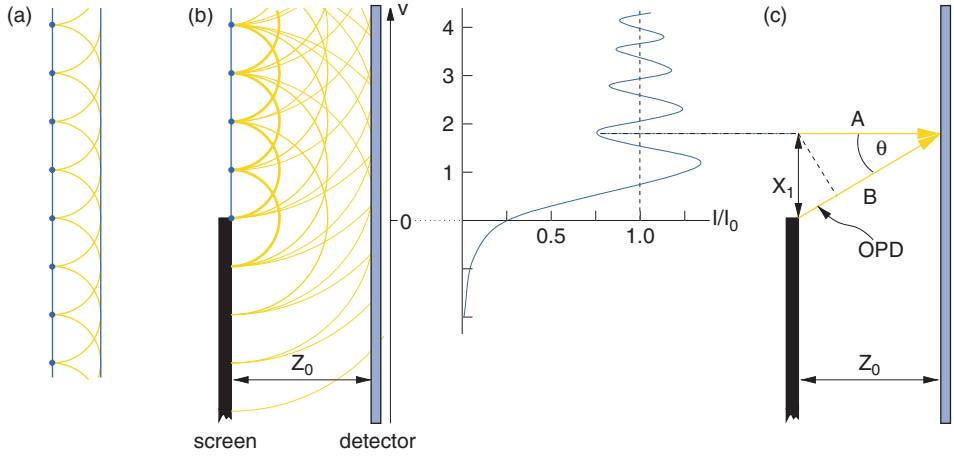

Figure 7.13 *Edge diffraction and enhancement. (a) A propagating wavefront can be constructed by assuming each point on that wavefront acts as a point source emitting spherical waves. Only in the direction of propagation of the plane wave is interference between the wavelets constructive, resulting in a new wavefront. (b) If the wavefront is partially blocked by an opaque screen, the wavelets from this part are missing, and the new wavefront can no longer form a perfect plane wave – interference fringes occur. (c) Simplified geometric explanation for the occurrence of the first dark interference fringe, whereby the optical path difference (OPD) between the forward propagating plane wave and a wavelet originating from the screen edge is* $\lambda/2$.

from the last point source before the plane wave is occluded by the screen edge. We next make the assumption that x_1 is small compared to the screen–detector distance Z_0, in other words θ is small and $\sin\theta \approx \tan\theta$, from which

$$\frac{x_1}{Z_0} \approx \frac{\text{OPD}}{x_1}, \tag{7.6}$$

where $\text{OPD} = \lambda/2$ is the optical path difference. From this, we obtain

$$x_1 = \sqrt{\frac{Z_0\lambda}{2}}. \tag{7.7}$$

Hence, the further away the detector is from the screen, the larger the fringe separations. We therefore introduce the dimensionless parameter

$$v = x\sqrt{\frac{2}{Z_0\lambda}}. \tag{7.8}$$

Fringe patterns from an opaque edge are therefore invariant if plotted as a function of v, as in Figure 7.13(b). Here, we see that in fact $v \approx 1.8$ for the first fringe minimum, that is, we have underestimated x by a factor of 1.8. This is explained by realizing that we have ignored all the other wavelets between rays A and B which still contribute partially constructively. Because the positions of the fringes are proportional to $\sqrt{\lambda}$, one

still observes edge diffraction (albeit more smeared out to a greater or lesser extent) for relatively broadband x-ray sources.

Note that directly opposite the screen edge the fringe intensity equals $I_0/4$, where I_0 is the beam intensity far away from the screen. This is simply explained by the fact that half the wavefront is occluded here and hence the amplitude in the forward direction is halved. But the intensity is proportional to the square of the amplitude, thus here it drops to one quarter. Note also that the constructive-interference fringes have *higher* intensities than I_0 – the wavelets blocked by the screen that would otherwise have interfered destructively with the rest of the wavefront are missing and so cannot contribute – the amplitude (and hence intensity) is therefore enhanced. This effect is termed *edge enhancement*.

Lastly, it should be noted that not only do the fringes become weaker away from the screen, they also become narrower. From Equation (7.6), we see that for an increase in the OPD of 2π radians between adjacent fringes Δx, the separation between fringes on the detector is approximately $2\pi / \sin\theta$.

What are typical dimensions for observing edge enhancement using hard x-rays? Edge enhancement is used in tomography to highlight internal and external boundaries of heterogeneous samples, hence the first fringe maximum should be resolved, but none of the subsequent, more narrowly spaced, fringes. The separation between the first two fringe maxima is $\Delta v \approx 1.2$. The sample–detector distance Z_0 should therefore be adjusted so that the lateral resolution (pixel size) matches this fringe separation. An example of the benefits of edge enhancement is shown in Figure 7.14.

Although in the above discussion we assumed that the screen was totally opaque to x-rays, it can be shown that similar effects can be observed for semi- or indeed completely transparent screens, as long as the screen has a different refractive index than the medium through which the x-rays are travelling.

Refraction not only changes the phase velocity of x-rays travelling through a medium, but also bends them if the beam impinges on the medium at a non-normal angle. Consider an x-ray beam passing through a transparent prism, as shown in Figure 7.15. Using Snell's law (Equation (2.16))

$$\begin{aligned} \cos\alpha' &= \cos\alpha / n \\ \alpha'' &= \alpha' + \chi \\ \cos\alpha''' &= n\cos(\alpha''). \end{aligned} \tag{7.9}$$

With $n = 1 - \delta$ we therefore obtain

$$\Delta = \frac{\pi}{2} - \arccos\left\{(1-\delta)\cos\left[\arccos\left(\frac{\cos(\pi/2-\chi)}{1-\delta}\right)+\chi\right]\right\}. \tag{7.10}$$

At glancing incident angles α, the deviation Δ is significantly larger than the minimum residual divergence of collimated synchrotron radiation (of the order of a microradian). Hence, interference can occur between the refracted and unrefracted parts of the beam, and is most prominent at those points in the object under observation where the incident x-rays impinge at a glancing angle, such as at its borders.

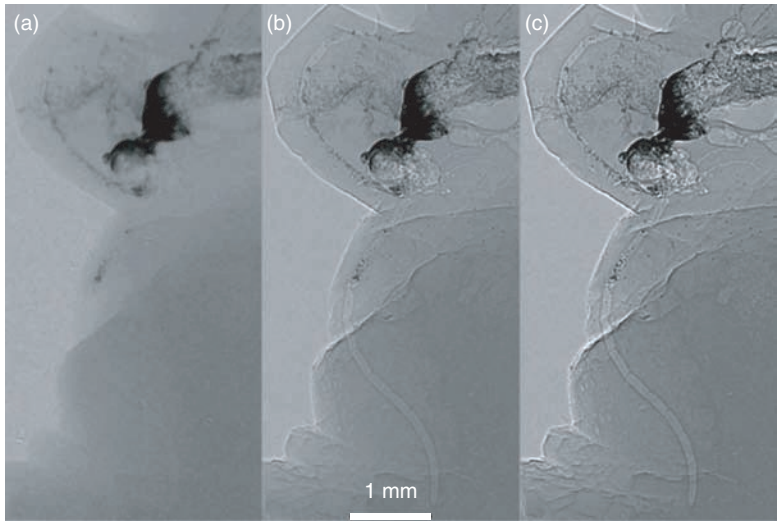

Figure 7.14 *Radiographic projections of an insect from the Cretaceous period trapped in opaque amber recorded using 20 keV radiation and a CCD camera with 7.5 μm pixel size and at different sample–detector distances (a) 10 mm, (b) 500 mm, and (c) 990 mm. The separation between the first two fringe maxima is 0.6, 4.7 and 6.6 μm, respectively. Hence, in the near-contact image (a) the fringe separation is much smaller than the pixel size and no edge enhancement is observed. In (b), the sharpness and clarity of the insect's features are much improved and are optimal in (c), in which the separation of the first fringes is only marginally smaller than the pixel size. Adapted from [9] with permission of Springer Science and Business Media.*

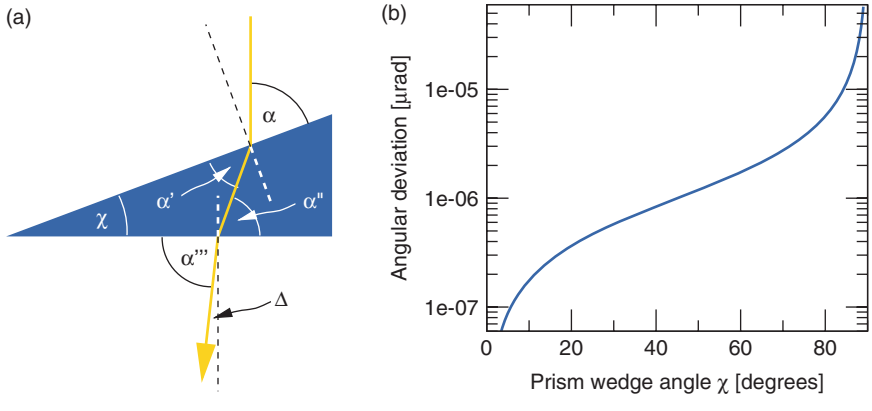

Figure 7.15 *Refraction of an x-ray beam through a prism. (a) The incident angle of the x-rays on the prism is $\alpha = \pi/2 - \chi$, where χ is the prism wedge angle. The emerging beam is deviated from the incident beam by an amount $\Delta = \pi/2 - \alpha'''$. (b) A plot of Δ as a function of χ.*

In summary, if phase-contrast-optimized radiographs are used in tomographic reconstructions, the three-dimensional data contains edge detection superimposed on the absorption contrast, which can be crucial for resolving features in biological specimens.

7.2.4.3 *Differential Phase-Contrast Imaging*

In the previous subsection, the information contained in the tomograms is qualitative insofar that the 'bulk' of the information is associated with absorption contrast, while boundaries are highlighted by edge detection. This is exceedingly useful in recognizing the internal architecture of soft-matter samples. Quantitative analysis in which a three-dimensional map of the refractive index (or more accurately, the refractive index *decrement* δ) is often desired. There are several different approaches to obtain this information by converting phase differences into intensity contrast [10–15] which go beyond the scope of this book. Here, we show one recent technique, differential phase-contrast (DPC) imaging, developed to obtain quantitative phase-contrast data using grating interferometers [16].

The principle behind DPC imaging is shown in Figure 7.16. The interference pattern produced by a grating G_1 with grating periodicity g_1 changes with distance downstream. At special distances (d_m) known as 'Talbot lengths', it has a regular periodicity equal to $g_1/2$. At this position, a second grating G_2 is placed with periodicity $g_2 = g_1/2$. Immediately behind G_2 is an area detector. If now a refracting sample is placed upstream of G_1, there will be slight distortions of the interference pattern at G_2 and therefore changes of the transmitted signal, which provide quantitative information about the phase

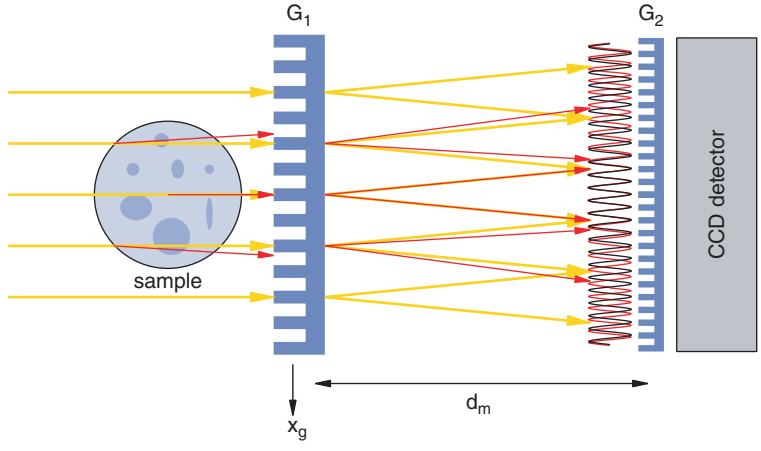

Figure 7.16 *Principle of differential phase-contrast imaging based on a grating interferometer. A phase grating G_1 with grating pitch g_1 produces at certain distances d_m interference fringes with a periodicity $g_2 = g_1/2$, where a second analyser grating G_2 with a pitch g_2 is positioned. In the absence of a refracting sample ('phase object') the interference pattern (shown as the black sinusoidal curve) and G_2 are in perfect registry. Introducing the sample causes the x-ray beam and fringe pattern (shown in red) to become distorted. The transmitted signal through G_2 changes in a manner that reflects the phase gradient. Quantitative phase information can be extracted from other contributions by shifting G_1 laterally (in the direction x_g) and recording the fringe pattern. Adapted from [16] with permission of IUCr.*

gradient of the illuminated object. In order to separate this phase information from other contributions, such as absorption and scattering, the first grating G_1 is shifted laterally (x_g) over one grating period g_1, while recording at least three images. The intensity oscillation at each pixel in the area detector can therefore be recorded. The phase ϕ of the oscillation can be related to the phase profile Φ of the wavefront and to the refractive-index decrement δ by

$$\phi = \frac{\lambda d_m}{g_2} \frac{\partial \Phi}{\partial x_g} \tag{7.11}$$

$$= \frac{2\pi d}{g_2} \int_{-\infty}^{+\infty} \frac{\partial \delta}{\partial x_g} \partial z. \tag{7.12}$$

The integral on the right is the line integral of the differential phase through the object and can be treated in the reconstruction of the tomogram in exactly the same way as is absorption in conventional tomography.

An example of DPC imaging of a rat's brain using 25 keV photons is shown in Figure 7.17. The characteristic size of the sample was 10 mm, but it was necessary to resolve the internal structure with a resolution of considerably less than 100 µm. Note that the x-ray transmission at 25 keV through a voxel of brain tissue of this length is more than 99.5%. This high transparency means that absorption-based contrast techniques are impractical.

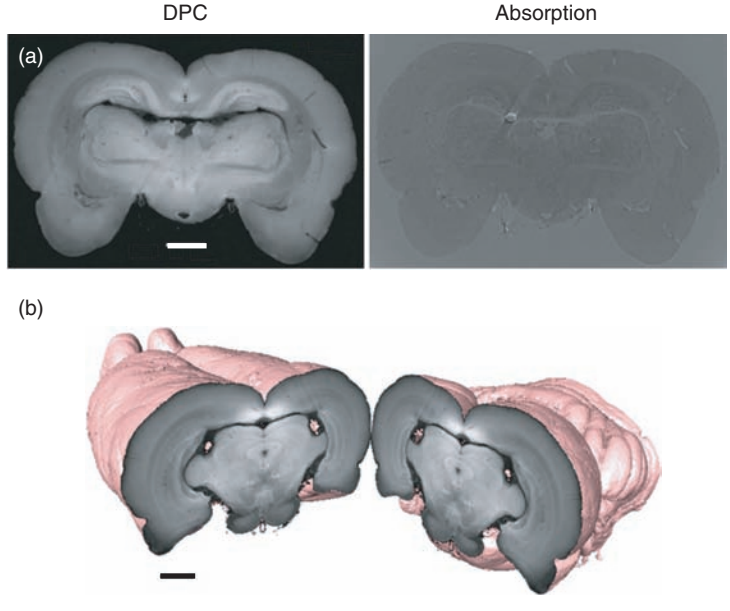

Figure 7.17 *DPC tomography of a rat's brain. (a) Comparison of the contrast of the same tomographic slice recorded using DPC (left) and conventional absorption tomography (right). (b) A three-dimensional isosurface of the rat brain sectioned across the middle to reveal the inner structure. In both figures the scale bar is 1 mm. Adapted from [16] with permission of IUCr.*

7.2.5 Soft X-ray Tomography

An alternative approach to phase-contrast tomography to obtain high contrast from low-Z samples, is to work in an energy regime where their absorption is strong. For biological samples containing primarily carbon, oxygen and nitrogen, this is the water window, described in Section 6.7.2. The transmission curves of different biological materials, shown in Figure 6.28, demonstrate that different architectural structures within cells, such as mitochondria, lipid walls and large protein structures, can be distinguished in this photon-energy regime and that typical cell dimensions of a few microns are ideally suited to this x-ray range of energies.

Soft x-ray tomography (SXT) has therefore been developed to investigate the structure of cells with a resolution of approximately 50 nm or better [17]. Its advantages over more established biological imaging techniques are the spatial resolution (compared to light microscopy) and the ability to ascertain the internal architecture without having to section the cell, normally a necessary step in electron microscopy. Both these improvements have revealed new insights in cell biology [18, 19].

A schematic of the setup used for SXT is shown in Figure 7.18. As the samples one is interested in are of the order of a micron in size, and features as small as a few tens of nanometer should be distinguished, direct radiographic projections and the use of scintillators are impractical. Instead, the x-ray optics setup is in principle the same as that used in an optical full-field microscope, in which the sample is illuminated by focusing the x-rays on to it with a condenser lens, and the image is magnified using an objective lens. Essentially, this is the setup used for transmission x-ray microscopy (TXM) used with both soft and hard x-rays, with the addition of a rotation stage for the sample.

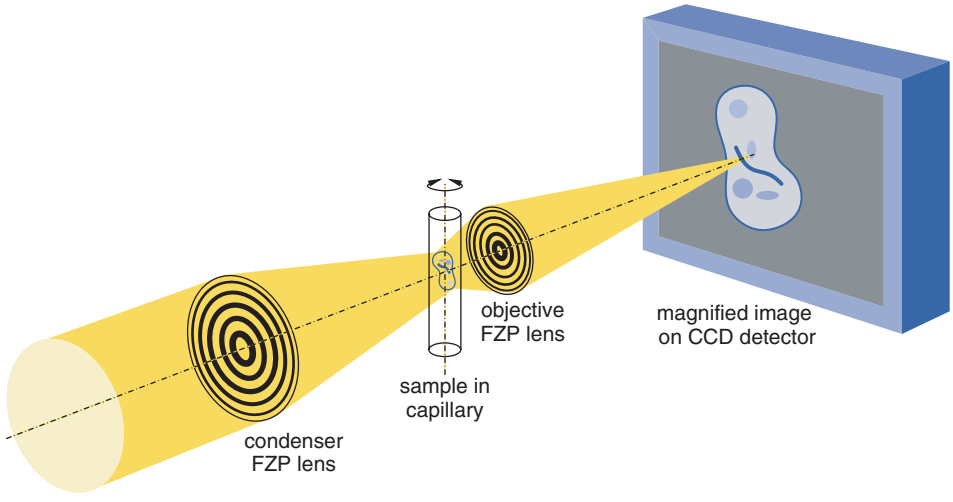

Figure 7.18 *Schematic of the SXT setup. Soft x-radiation with photon energies lying within the water window are focused onto the sample using a Fresnel zone plate (FZP) condenser lens. The objective FZP lens magnifies the specimen image on an area detector. Three-dimensional tomographic images are generated by rotating the specimen, which resides in a very thin-walled silica capillary tube, in steps of approximately 1°.*

The main advantage of SXT, namely the high absorption strength of the specimens and consequent high imaging contrast, is, paradoxically, also one of its major drawbacks. Biological samples absorb a large fraction of the incident soft x-ray flux. In order to avoid that they are consequently very quickly destroyed, they are rapidly frozen in liquid nitrogen in a process called 'flash freezing' (also used in protein crystallography), which inhibits the diffusion of photoinduced radicals through the specimen. The rapidity of the cooling is essential in order to suppress the formation of large ice crystals, which would otherwise destroy the cell structure through expansion. This obviously means that no *in-vivo* samples can be investigated, and that living processes cannot be followed in any one specimen. Therefore, in order to follow, for example, cell mitosis or infection, individual cells must be monitored using light microscopy and flash-frozen at different stages of the process. An example of a red blood cell infected by malaria is shown in Figure 7.19 [20]. This analysis, using SXT of the different stages of malarial infection within red blood cells and the mechanisms behind antimalarial drugs, has contributed significantly to the understanding of malarial infection and multiplication.

In addition, the sample environment and optics must reside in vacuum. First, air at one atmosphere has an absorption length at 540 eV (just below the oxygen K-edge) of less than 1 mm. Secondly, formation of ice on the cryogenically-cooled sample surface from air-born water cannot be tolerated – even a 10 μm-thick ice layer absorbs 60% of the beam at 540 eV.

Another limitation of SXT concerns the depth of field (DOF) of high-resolution FZPs. The DOF of an imaging optic is defined as the range of distances between the object and

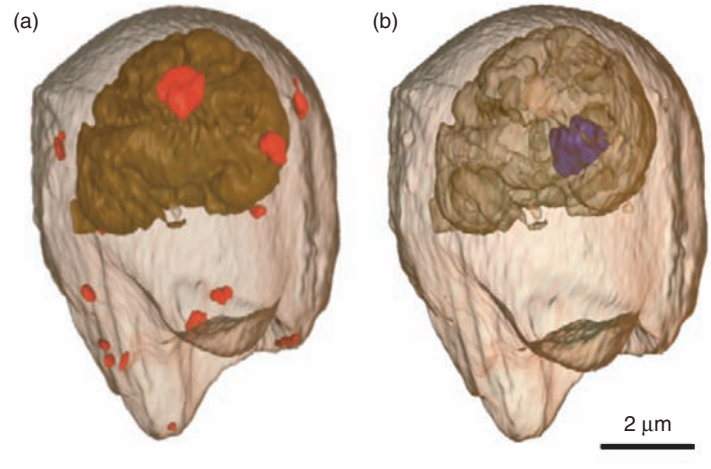

Figure 7.19 *Soft x-ray tomography of red blood cells infected with the malarial parasite* Plasmodium falciparum *in the early to middle stages of infection. The red blood cell, rendered as translucent pink, contains the malaria parasite, shown in brown. (a) The red features are on the surface of the red blood cell and are involved in transporting proteins across the surface. The purple feature within the parasitic body in (b) is an aggregate of so-called hemozoin crystals formed by the breakdown of the red blood cell's haemoglobin, which the malaria parasite uses as a source to build the amino acids needed for multiplication. Adapted from [20] with permission of Elsevier.*

the imaging (objective) lens for which the image resolution does not degrade by more than a factor of two. For FZPs, this quantity is inversely proportional to the number of rings N in the zone plate. As this necessarily has to be large in order to obtain high lateral resolution, the DOF can often be smaller than the size of the sample being investigated. Although this can in principle be corrected for by deconvolution techniques, this requires computational power and an accurate knowledge of the focusing properties. As the resolution in phase-contrast hard x-ray tomography continues to improve, and lensless imaging techniques such as ptychographic x-ray diffraction microscopy mature (see Section 7.3.4), their advantages regarding their low dosage capability to investigate *in-vivo* samples and large depth of field are making them increasingly competitive with soft x-ray imaging techniques [21, 22].

7.2.6 Worked Example – The Spread of Hominoids

The evolution of the hominoids (the great apes) counts as one of the most controversial and emotive debates in the natural sciences. Although it is widely accepted that the first hominoids evolved approximately 20 million years ago (MYA), at the start of the Miocene period, their subsequent spread and diversification is still disputed. The earliest fossils specimens found outside Africa date to some 16.5 MYA. Between 12 and 6 MYA (the end of the Miocene), many different species are known to have existed in Africa, Asia and Europe. Diversification seems to have declined after that, and now, only five genera, the gibbons, orangutans, gorillas, chimpanzees and man, remain.

Conventional theory has it that the principal area of diversification was Africa during the whole of the Miocene period, with staggered migrations to Europe and Asia. In an alternative hypothesis, however, it is suggested that early hominoids migrated from Africa and that most of the diversification happened in south-east Asia, while the original species became extinct in Africa. Only later, it is proposed, was Africa repopulated and migration to Europe occurred (see Figure 7.20). Fossil records from both Africa and Asia should, therefore, provide valuable evidence to test these competing theories [23].

In 2003, fossil teeth from a previously unknown species of hominoid were discovered in Thailand, dating from approximately 12 MYA (see Figure 7.21(a)). The large differences in their sizes indicated that they originated from both males and females of a species which showed strong sexual dimorphism, that is, the male was much larger than the female. This new species was named cf. *Lufengpithecus chiangmuanensis*.[3]

Although traditional palaeontological studies of the external dental morphology provide valuable information, the internal structure of the teeth is equally important. Although sectioning a fossil tooth will yield this information, this involves destroying an exceptionally rare fossil artefact! XTM allows one, however, to *virtually* slice and dice as one pleases while the original remains intact. The individual teeth were investigated using absorption-contrast XTM at 50 keV at the ID19 beamline at the ESRF [9, 24]. At this energy, fossilized material has an attenuation length of approximately 10 mm. The voxel resolution was 30 μm. The use of synchrotron radiation was especially important here – although laboratory-based x-ray machines can provide enough flux and the

[3] The term 'cf.' means 'confere', that is, the species is similar to another fossil species, but incompleteness in its fossil record forbids a more formal attribution.

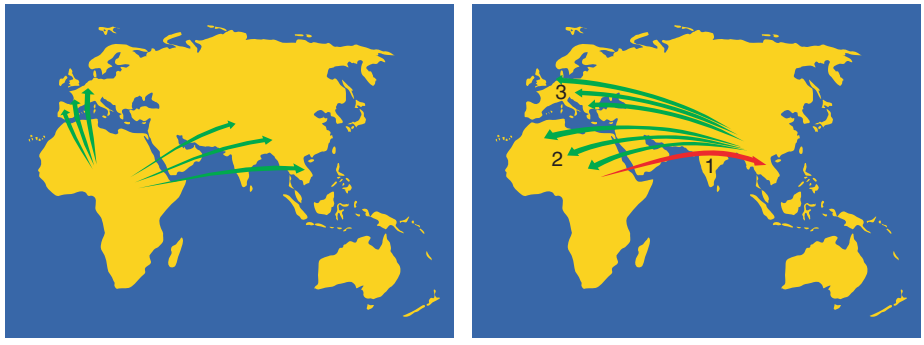

Figure 7.20 *The spread of hominoids. Left panel: the classical hypothesis holds that diversification happened mainly in Africa, starting about 20 MYA, followed by successive migratory waves to Asia and Europe. Right panel: an alternative thesis suggests that after a few million years' evolution in Africa, some species migrated to Asia (1), while those remaining in Africa became extinct. The main centre of evolutionary change is claimed to be south-east Asia, while repopulation of Africa (2) and migration to Europe (3) only happened later. Adapted from [23] with permission of the European Molecular Biology Laboratory.*

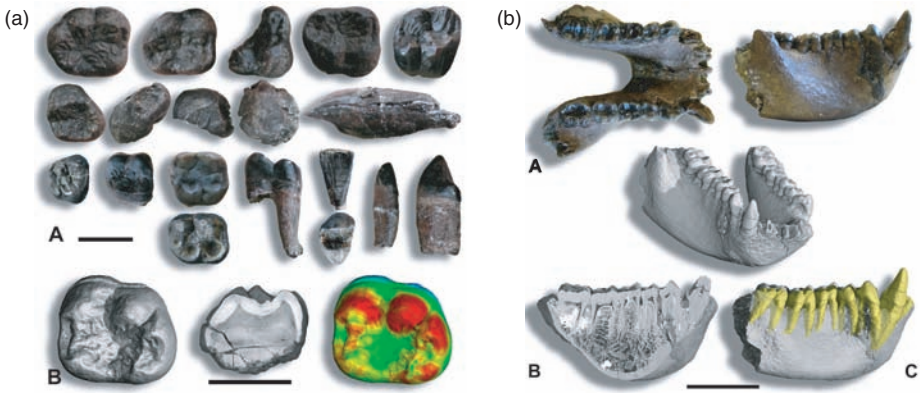

Figure 7.21 *(a) Photographs (A) of fossil teeth discovered of* Lufengpithecus chiangmuanensis *(later renamed* Khoratpithecus chiangmuanensis*). B. XTM analysis of a lower molar. Left to right: 3-D-reconstruction, virtual vertical cut, and a distribution map of enamel. The scale bars represent 1 cm. (b) Photographs (A) of the lower mandible of* Khoratpithecus piriyai. *B: XTM analysis. The central image shows a 3-D-reconstruction. A virtual vertical cut through the right side of the mandible displaying the bone structures and the dental roots is shown lower left. On the lower right (C), the intact teeth are shown. The scale bar is 4 cm. Adapted from [23] with permission of the European Molecular Biology Laboratory.*

high photon energies needed to penetrate several millimetres of mineralized fossil material, they are polychromatic in nature. This introduces the phenomenon of 'beam hardening' associated with polychromatic x-rays. Because softer x-rays are absorbed more strongly, border regions of a sample, which tend to be less thick, will let through a disproportionately larger fraction of the spectrum. As the tooth-enamel resides on the borders, this region will appear to be anomalously bright, leading to a misinterpretation

of the thickness [9]. The problem is further exacerbated because detectors tend to be more sensitive to softer than harder x-rays for the same reason, falsely brightening the border regions still more.

In this manner, it was possible to quantify precisely the thickness and distribution of the dental enamel (see lower right image of Figure 7.21(a)). Surprisingly, it was found that this fossil species is more similar to modern orangutans than to any known fossil hominoids found in Asia. It was therefore tentatively proposed that cf. *L. chiangmuanensis* might be a direct ancestor of modern-day orangutans.

The plot thickened, however, the following year in 2004, when a complete mandible (lower jawbone) from approximately 7 MYA was discovered, also in Thailand, and was assigned to yet another new species, named *Khoratpithecus piriyai* (*Khorat* is the region in Thailand where the fossil was discovered). From traditional palaeontological methods, it seemed clear that these two fossils were more similar to one another than to any other fossil or living species, despite the fact that one fossil was older than the other by some 5 million years. The older fossil was consequently renamed *Khoratpithecus chiangmuanensis*.

K. piriyai's mandible was also imaged using XTM. Because of the much larger size of this fossil mandible, whereby in certain directions the x-rays would have to penetrate through approximately 10 cm of mineralized material, the experiment was performed at the ID17 medical beamline of the ESRF [9, 25] at a still higher photon energy of 70 keV, for which the attenuation length within the fossil was approximately 25 mm. The results are summarized in Figure 7.21(b). The voxel resolution was approximately 50 μm, perfectly adequate for such a large specimen. Indeed, the voxels were further binned to provide an effective resolution of about 100 μm. The low absorption contrast between the different structural elements inside the mandible (the enamel, the underlying mineralized dentine, the roots and the jawbone) meant that phase-contrast imaging was necessary to differentiate between these, something which can only be provided at a synchrotron source.

The XTM data of *K. chiangmuanensis*'s teeth were then used in a virtual reconstruction of its jaws (Figure 7.22). This allowed a closer comparison of the reconstructed jaws of the older species and the mandible of the younger fossil. It emerged that their similarities were even stronger than first believed, thereby strengthening the justification of their classification in a single genus, made earlier on the basis of traditional palaeontological techniques.

XTM imaging of these fossils therefore revealed anatomical characteristics that would otherwise have been impossible to study without destroying the fossils. *Khorapithecus* is the known genus most closely related to modern orangutans, although specialized features in the more modern *K. piriyai* are very unlike modern orangutans. From what is known about the evolutionary development of other lineages, it seems unlikely that so many specializations would have evolved and subsequently disappeared. Hence *K. piriyai* is probably a distant cousin to modern orangutans rather than being a direct ancestor.

The older species *K. chiangmuanensis*, however, shows fewer specialized dental characteristics, and is therefore a more likely candidate for being a direct ancestor of the orangutans. More fossil discoveries are necessary to test the hypothesis, but it remains very possible that the orangutan's lineage was derived from primitive forms of *Khoratpithecus* and that subsequently two branches evolved in different ways, one

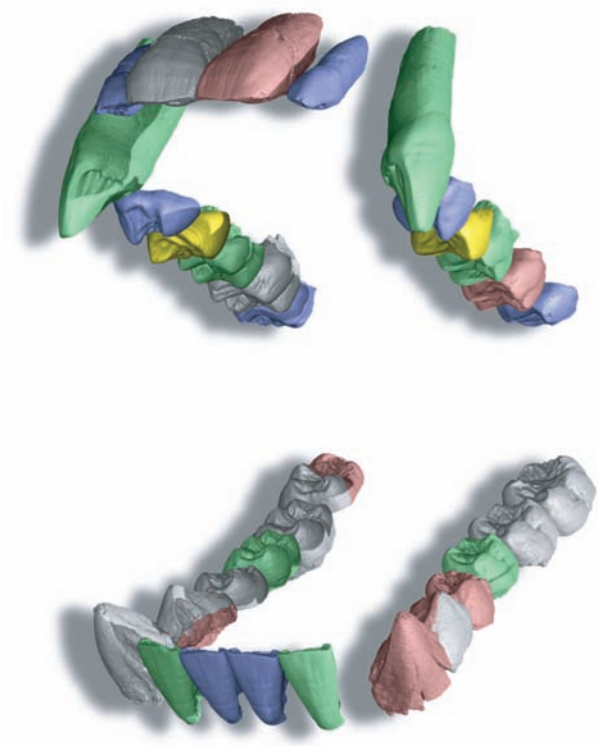

Figure 7.22 *Jaw reconstitution of male* K. chiangmuanensis. *Grey: original male teeth. Pink: symmetrical view of these teeth. Blue: female teeth after adjustment of their size to take into account the sexual dimorphism. Green: missing teeth with equivalents extrapolated from existing teeth after calculation of their probable sizes. Yellow: completely unknown teeth replaced by teeth of orangutans. Adapted from [23] with permission of the European Molecular Biology Laboratory.*

culminating in modern orangutans and the other to later species of *Khoratpithecus*, such as *K. piriyai*.

How do these findings affect the two hypotheses of hominoid evolution? First, the two discovered *Khoratpithecus* species, as well as numerous other fossil hominoids from Asia, show more variation than that of African hominoid fossils during the same period. Importantly, however, fossil remains of primitive species have also been found in Asia, indicating that diversification in Asia started early in the Miocene. Lastly, Asian hominoid fossils are found across large areas of Asia and in geological strata ranging from 16.5 MYA to the present. All these facts support the hypothesis that Asia was an important, perhaps even the principal, centre of hominoid diversification during the Miocene. Lastly, it should be mentioned that this hypothesis does not undermine the 'out-of-Africa' theory of the evolution and spread of *Homo sapiens*. This happened far more recently – the genus *Homo*, of which *Homo sapiens* is the only extant (i.e. surviving) species, first appeared approximately 2.5 MYA, some three million years after the end of the Miocene.

7.3 Lensless Imaging

7.3.1 Introduction

Coherent diffraction is 'normal' diffraction writ big. Whereas in standard crystallography one is interested in determining the atomic components and their positions within a unit cell, coherent diffraction attempts to determine structures on the mesoscopic scale, that is, in the range of nanometres to a few tens of nanometres. 'Normal' imaging using lenses consists of a beam of light or x-rays being scattered by a sample, which in the far field is the Fourier transform of that sample. The inversion back to an image of the sample is performed by the lens. Coherent diffraction is a form of lensless imaging, whereby the inverse Fourier transform is not carried out by a lens, but by computation.

Coherent diffraction imaging (CDI, often alternatively referred to as lensless imaging, or diffraction microscopy) was first conceived by David Sayre more than half a century ago, but had its first successful realization using soft x-rays ($\lambda = 17$ Å, 730 eV) in 1999, on an array of gold dots of 100 nm diameter. The resolution was specified as being approximately 75 nm [26] (see Figure 7.23).

What are the advantages of coherent diffraction imaging? The resolving power of a lens (i.e. the smallest sized feature Δx that a lens can distinguish) is inversely proportional to the lens diameter (Figure 7.24). X-ray lenses, especially 'big' ones, are difficult (essentially impossible) to manufacture such that they produce no aberrations, hence dispensing with them solves this problem. The resolution of lensless imaging is given by

$$\Delta x = 4\lambda \frac{Z_0}{D}, \tag{7.13}$$

where Z_0 is the sample–detector distance, and D is the detector size. It is therefore in principle easier to obtain high resolution in lensless imaging simply by using area detectors which subtend a large angle with the sample. As we saw in Section 5.16.2, the scattering power drops off with the fourth power of the scattering vector, hence the signal intensity drops off sharply away from the direct beam. Lensless imaging is therefore limited by a tradeoff between resolution and signal statistics.

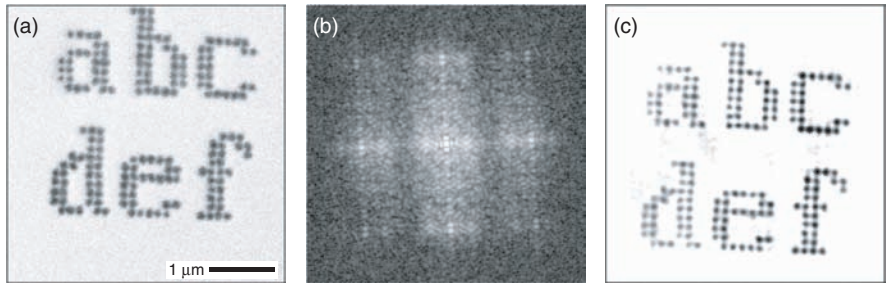

Figure 7.23 *Diffraction microscopy of a gold array. (a) A scanning electron microscope image of the array, comprised of gold dots of 100 nm diameter. (b) The diffraction pattern produced by coherent illumination of the array using 1.7 nm radiation. (c) Reconstruction of the array using a Fienup iterative phasing algorithm. Adapted from [26] with permission of Macmillan Publishers Ltd.*

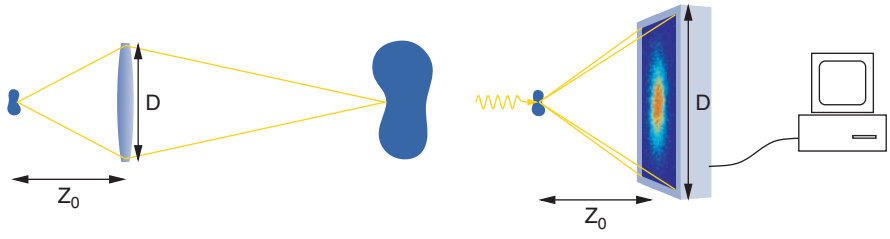

Figure 7.24 *Focusing optics versus lensless imaging. Left: The resolution of imaging using lenses is limited by the size of the lens, D and the sample–lens distance Z_0. Because x-ray lenses are difficult to make aberration-free, especially for large D, lensless imaging (right) becomes attractive. Here, the resolution improves linearly with detector size D. The drawbacks are, of course, the need to inverse-Fourier transform the far-field scattering image of the pattern recorded by the detector, which is hampered by the phase problem; and the low scattered intensity at large angles.*

Second, CDI neatly bridges the resolution gap between optical microscopies (hundreds of nanometres) and electron microscopies (nanometre or subnanometre). Third, and perhaps most importantly, because of the transparency of matter to x-rays, the *internal* structure of objects can be probed. Lastly, because of this weak interaction of x-rays with matter, one can apply the kinematical approximation, thereby minimizing the computational effort in reconstructing the image. Which brings us to the main drawback of the technique, namely the ubiquitous phase problem. As we will see later in this section, sophisticated methods have been developed to overcome this.

Because standard x-ray crystallography is concerned with the arrangement of atoms within the unit cell, the spatial coherence of the x-ray beam in principle needs only to be larger than the linear dimensions of the unit cell. However, the intensity of diffraction maxima increases with the square of N, the number of coherently illuminated unit cells in any given direction, while the width of the signal is inversely proportional to N. From Equation (3.23), we can calculate that for beamlines not immediately concerned with high coherence, the upper limit to the transverse coherence lengths is of the order of 50 µm for hard x-rays of approximately 1 Å wavelength. However, this estimate ignores any degradation of the wavefront of the x-rays produced by optical elements on the path from the x-ray source to the illuminated sample. Refraction of the x-rays as they pass through microscopically inhomogeneous material (such as beryllium windows, polycrystalline filters, etc.), and slope errors associated with monochromator elements and mirrors, result in a 'scrambling' of the phase relationship across the x-ray beam, and a consequent reduction of the coherence. Typical coherence lengths can therefore be an order of magnitude smaller than the theoretical limits. The distortion of the wavefront can in principle be determined and taken into account when retrieving the real-space image, although drifts in the distortion field over the time required to record the data, caused by thermal effects or vibrations, present more serious problems.

In order to obtain useful results in coherent diffraction imaging, two conditions must be met: first, the sample must be entirely bathed in the coherent part of the beam; and second, the maximum path-length difference must be less than the longitudinal coherence length.

Third-generation synchrotron sources exhibit low coherence – the longitudinal (temporal) coherence is limited by the monochromacity of the beam (typically of the order of $\Delta\lambda/\lambda = 10^{-4}$), while the transverse (spatial) coherence is determined by the source size [of the order of 20 (h) $\times$200 (v) μm^2] and the distance of the source from the sample (see Equation (3.23)). The degree of coherence can be improved by inserting micron-sized pinholes in the beam path and by making the source-to-sample distance as long as possible – for this reason, one of the three experimental hutches at the BL29XU RIKEN beamline at SPring8 in Japan is placed one kilometre downstream from the undulator source!

For a given beam brilliance, B, the spatially coherent flux F_{coh} scales with the square of the wavelength, while for a given absolute bandwidth, $\Delta\lambda$, the coherence is linearly proportional to the wavelength. Specifically,

$$F_{coh} = B\left(\frac{\lambda}{2}\right)^2\left(\frac{\Delta\lambda}{\lambda}\right). \tag{7.14}$$

It is therefore easier to produce coherent soft x-ray sources, although the ultimate resolution is less than that when using hard x-rays.

Beamlines that utilize the coherent part of the synchrotron radiation are therefore characterized by a small source size, minimization of optical components, minimization of sources of vibration and a large source-to-sample distance. A beamline optimized for coherence can have transverse coherence lengths of the order of 100 μm or more. Normally samples have dimensions considerably smaller than this, in which case, the need for coherence is somewhat relaxed and one can partially focus the incident beam to increase the signal intensity.

One can divide up samples into two types – noncrystalline and single crystal. In the first case, one is interested in the signal close to the forward scattering direction around the direct beam. In the case of crystalline samples, one can investigate the details of the scattering around the Bragg peaks associated with the crystal, which can lend some important advantages, described below.

To understand the physics of coherent diffraction, we next turn to the concept of speckle.

7.3.2 Speckle

Speckle is the result of interference between wavelets produced by scattering of x-rays caused by spatial variations in the electron density within an object (see Figure 7.25). It therefore has its origins in the same phenomena responsible for diffraction patterns, and is the square of the Fourier transform of the electron density distribution of the object. To obtain a stable speckle pattern the object must be coherently illuminated – if this is not the case the speckles become averaged out due to the fluctuation of the form of the wavefront incident on the sample. In the forward direction around the direct beam, this incoherent signal is referred to as small-angle x-ray scattering (SAXS), covered in Section 5.16.

There are three characteristic length-scales related to speckle: the wavelength λ of the x-rays, measured in Angstroms; the typical linear dimensions of the sample a, which should be smaller than the transverse coherent illumination cross-section (of the order

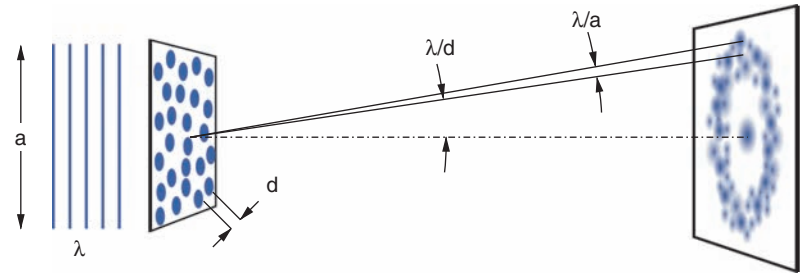

Figure 7.25 *Scattering of coherent illumination of an object produces a speckle pattern with an angular extent of the order of λ/d, and 'speckle' features within the pattern separated from each other by λ/a.*

of 100 μm); and d, the typical length describing the variation in electron density (or distribution of 'particles') within the illuminated object, which can range between 10 and 1000 nm (Figure 7.25). The x-rays are on average scattered through an angle of the order of λ/d, which can be as small as 100 μrad (less than 0.01°). The speckles arise through interference between all the scattered wavelets across the coherently illuminated area and therefore have angular separations that can be as small as $\lambda/a \sim 1$ μrad. Note that in order to resolve such features, a detector having 50 μm-sized pixels would need to be at least 50 m downstream of the sample, for hard x-radiation in the Angstrom range. Normally, however, the illuminated area is limited to only a few microns, and sample–detector distances of approximately 10 m are more common. Alternatively, softer x-radiation can be used, although the increased absorption at lower photon energies sets an upper limit to the size of objects that can be studied.

Note that because the speckles arise from interference between the scattered wavefronts from all the particles in the sample, the entire speckle pattern will change even if only one particle moves. The speckle pattern therefore contains information on the positions of all the particles, which can be obtained by inverting the pattern, involving solving the phase problem. This is nowadays commonly achieved by applying phase-retrieval techniques such as the so-called 'difference-map' algorithm [27–29] (see Figure 7.26).

7.3.3 Noncrystalline and Crystalline Samples

In order to obtain the maximum amount of information about the three-dimensional internal structure of a noncrystalline object, one must rotate it through at least 180°. The reason for this is best understood by considering the Ewald construction shown in Figure 7.27. In reciprocal space, only that part of the pattern that lies on the surface of the Ewald sphere will be recorded, hence in order to obtain the whole speckle pattern centred around the forward-scattered beam (the (000) point), it must be rotated by at least 180°, and with sufficiently small steps so that information does not go missing in between.

In contrast to noncrystalline objects, crystalline samples also have strong Bragg peak signals far away from the direct beam. As we have already discussed in Section 5.13.2, the diffraction pattern of finite-sized crystals is the convolution of the diffraction pattern of an infinitely large crystal with the Fourier transform of the function describing the

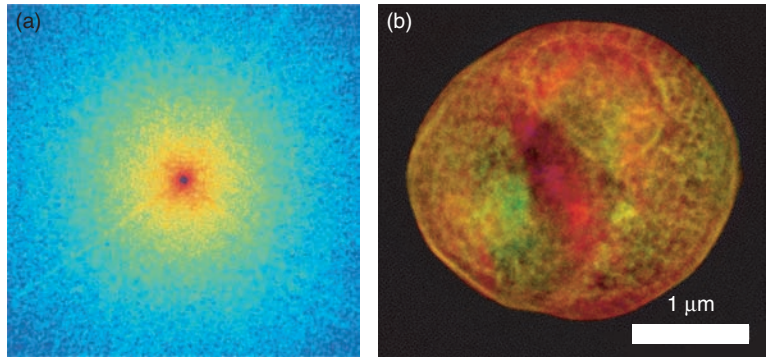

Figure 7.26 *Lensless imaging of a yeast cell. (a) The speckle pattern produced by illuminating the cell with coherent radiation of wavelength 1.65 nm (750 eV) and a transverse coherence length of 12.8 µm was recorded on a CCD array containing 1300 × 1340, 20 µm-sized pixels 150 mm downstream of the sample. This gave a resolution of approximately 20 nm. (b) The yeast cell was reconstructed from speckle patterns taken at nine different angular positions using the difference-map algorithm [27]. Reprinted from [28] with permission of IUCr.*

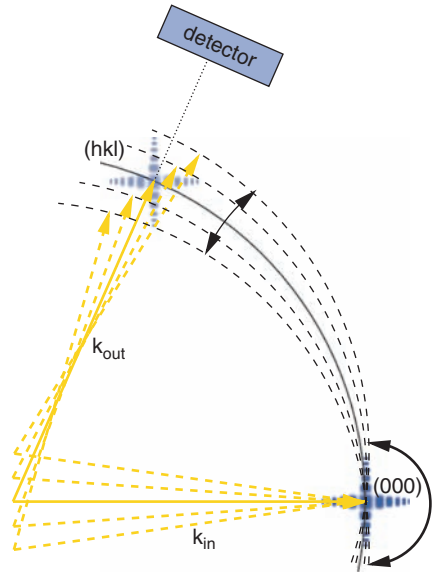

Figure 7.27 *The speckle pattern for crystalline objects can be recorded around Bragg peaks far away from the direct beam. The angular range one must rotate the sample for the Ewald sphere to pass across the scattering pattern and thereby record the entire pattern is much reduced compared to the minimum of 180° required for noncrystalline samples, which only scatter significantly in the forward direction.*

Figure 7.28 *The Bragg peaks of the diffraction pattern of a nanocrystalline sample reflect the 'shape function' of the object.*

boundary of the actual crystal (the 'shape function'), hence each Bragg peak has the form of the shape function (see Figure 7.28).

The exact shape of the Bragg peaks can be further modified by other phenomena such as internal strain [30]. Importantly, however, it reflects the entire internal structure of the crystal, and the shape function is just the speckle for ordered systems.

This has three important consequences. First, coherent diffractive imaging can be carried out far away from the direct beam. This means one can dispense with the beamstop (which excludes features in the reconstruction above a certain size), while diffuse scattering in the forward direction produced by other sources than the sample is avoided. Secondly, because the Bragg peak shape is only affected by the electron density of the crystalline part of the object, it is highly sensitive to crystalline defects and strain fields [30, 31]. Lastly, one must rotate the sample through a much smaller angular range than in the case of noncrystalline samples in order for all the features to pass through the Ewald sphere (Figure 7.27), significantly relaxing the technical specifications of the rotation stage.

7.3.4 Ptychography

Ptychography is an experimental method originally developed in the 1970s for electron microscopy. Recently, it has been applied to x-rays, with dramatic results [32, 33]. Ptychography comes from the Greek 'to fold' which refers to the overlap of and interference between coherently illuminated adjacent Bragg reflections.

Modern x-ray ptychography can be thought of as a marriage of coherent diffractive imaging (see previous section) and scanning transmission x-ray microscopy (STXM, see Section 6.7). The only (but crucial) differences between STXM and ptychography are that in the latter, coherent x-radiation is used and the transmitted signal is spatially resolved using an area detector, rather than recording just the integrated signal (as in STXM). Ptychography has consequently been recently rechristened 'scanning x-ray diffraction microscopy' (SXDM). The configuration of SXDM and the principle behind it is shown in Figure 7.29.

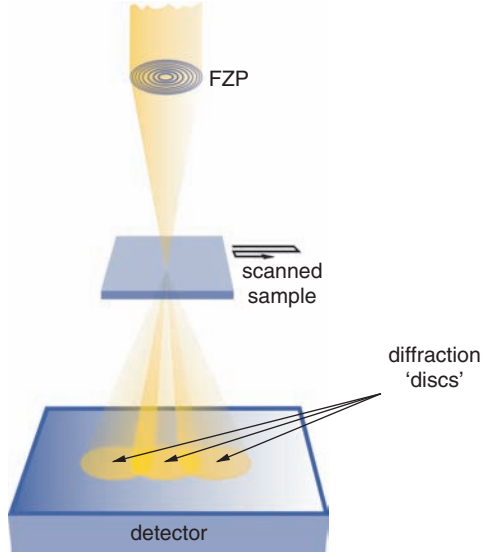

Figure 7.29 *The principle of SXDM. An x-ray beam is either tightly focused on to a crystalline sample using a Fresnel zone plate, or is a secondary source defined by a pinhole. Because of the beam divergence, the diffraction peaks become spread out and overlap with each other on the area detector. After recording each diffraction image, the sample is translated by a distance slightly less than the focal spot size.*

The incident beam on the sample is either so tightly focused that the beam divergence is larger than the angular separation of adjacent diffraction maxima, or is defined by a micronsized pinhole.

In the case of crystalline samples, the diffraction pattern on an area detector therefore contains diffraction 'discs' as a result of the beam divergence, which partially overlap. In these regions of overlap, the diffraction maxima interfere with each other and the intensity depends on their relative phases. In the case of noncrystalline samples, the same overlapping information is obtained by keeping the scanning shift of the sample between recording images less than the extent of the localized illumination.

It is this 'folding' of the diffraction or speckle features that contains the extra crucial information that allows one to extract the phases required for reconstructing the object. There remains an ambiguity regarding the sign of the phase for any given diffraction image, but this is resolved by observing how the overlap intensity changes as one translates the sample. The importance of this aspect has been succinctly highlighted by Bunk *et al*., in which they recorded SXDM data in the visible on an insect wing using different degrees of overlap (defined as the ratio of the distance travelled by the sample between diffraction images to the illumination diameter). The results are shown for the insect wing illuminated with laser light in Figure 7.30. At an overlap of 20%, one begins to be able to distinguish the wing veins, which become increasingly clear until at 60% overlap, the optimal condition has been achieved [34]. Above this, there is very little gain in resolution for an increase in data volume which is inversely proportional to the square of the sample shift distance.

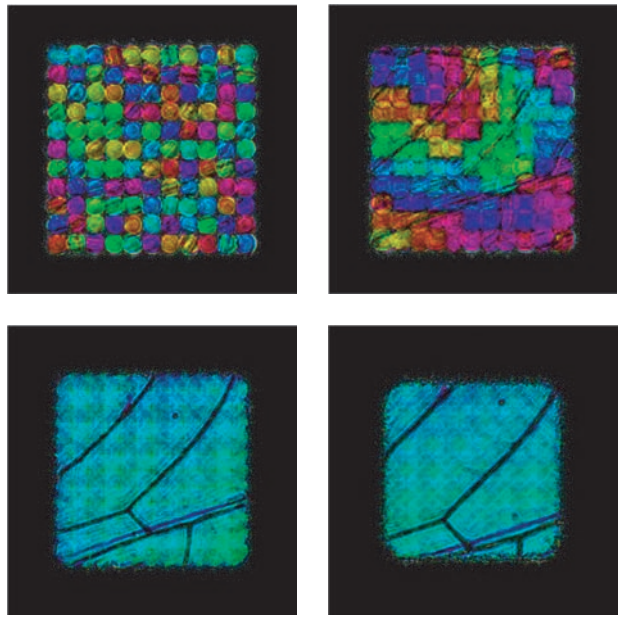

Figure 7.30 *Amplitude and phase images retrieved from SXDM data recorded on a fly wing using different overlaps: (a) 0%; (b) 20%; (c) 40%; (d) 60%. Brightness codes for amplitude, while the hue represents the phase. Note that for (c) and (d), the colour (i.e. the phases) is constant, indicative of a real-valued object. Adapted from [34] with permission of Elsevier.*

In conventional STXM, the resolution is limited by the size of the focal spot. The extra information obtained by using coherent radiation and an area detector in SXDM means that the resolution is significantly improved. In the first reported example of SXDM, the beam was focused to a few hundred nanometres on the sample, while the final SXDM image had a resolution of a few tens of nm. Importantly, because the sample is scanned, large fields of view can be obtained, of the order of several hundred square microns or more.

Scanning SAXS is a novel technique that is essentially identical to SXDM, but uses a partially coherent beam whereby the transverse coherence length is smaller than the illuminated area. Scanning SAXS is able to span distances between neighbouring scattering centres, but not the entire sample. This technique bridges the resolution gap between local probes, such as scanning tunnelling microscopy and electron microscopy, and macroscopic techniques such as tomography. Scanning SAXS has a resolution capable of resolving features down to 1 nm (although, because of the reduced coherence, only an average of their orientation is obtained) and can map over arbitrarily large extended areas [35, 36].

7.3.5 X-ray Photon Correlation Spectroscopy

X-ray photon correlation spectroscopy (XPCS) sits somewhat uncomfortably in a chapter on imaging, as it does not produce any images. However, because it uses the phenomenon

of speckle to extract information about dynamical processes, I decided to include this subsection as an adjunct to the section on lensless imaging.

We have already argued in Section 7.3.2 that because speckle arises from interference between scattered wavelets from all objects illuminated by the coherent x-ray beam, any movement of even one of these objects will alter the entire speckle pattern. Photon correlation spectroscopy (PCS) in the visible spectrum, using lasers, has been an established technique for studying dynamics in optically transparent systems for several decades. However, PCS is unable to probe length scales less than about 200 nm, and is excluded from systems that are opaque, or are embedded in an opaque medium. By extending PCS to the hard x-ray regime, it was recognized in the landmark paper by Sutton *et al*. in 1991 [37] that XPCS can be used as a powerful method to investigate dynamics on length scales given by the speckle pattern, which can range from over 200 nm to as little as 1 nm; and to timescales that are limited either by the repetition rate of the x-ray detector (which nowadays can have frame rates above 10 kHz for area detectors, and orders of magnitude more still using scintillator point detectors), or more commonly by the limited photon flux and scattering power of the sample. With the advent of hard x-ray free-electron lasers, XPCS is likely to experience a renaissance in the field of atomic-resolution dynamics.

In the first paper by Sutton *et al*., speckle was produced by scattering so-called antiphase domains in the binary alloy Cu_3Al. Although no dynamics were reported, a subsequent paper demonstrated fluctuations in the speckle pattern around the (1/2, 1/2, 1/2) superlattice reflection in the domain structure of the related alloy Fe_3Al above an order-disorder transition temperature [38].

In these papers, Sutton *et al*. used a point detector with an analysing pinhole, and could therefore only follow the temporal evolution of a single speckle feature. By using an area detector, the statistics for dynamics on a scale of length d can be enormously improved by averaging over all speckles which lie on a ring of radius $Q = 2\pi/d$ (see Figure 7.31).

7.4 Concluding Remarks

Imaging using x-rays is an enormously profitable enterprise, in which insights on the submicron scale can often be rapidly obtained – indeed, the first demonstration of the power of x-rays was made by Röntgen with his radiograms of human hands. The impact these must have had on society can hardly be appreciated in a modern world in which medical radiographic imaging is so commonplace. However, the detailed architectural information provided by tomography on a microscopic scale can still both surprise and delight. The immediacy of the feedback and the relative ease of at least qualitative interpretation supplied by x-ray imaging makes it one of the most attractive synchrotron techniques.

As still more powerful and coherent x-rays sources such as XFELs and ERLs begin to come online, their associated enormous power densities means that any optical components such as focusing lenses will be circumvented if they are not absolutely necessary. Lensless imaging is emerging not just as an attractive alternative to conventional x-ray microscopy, but may in many cases be the only realistic approach in investigating

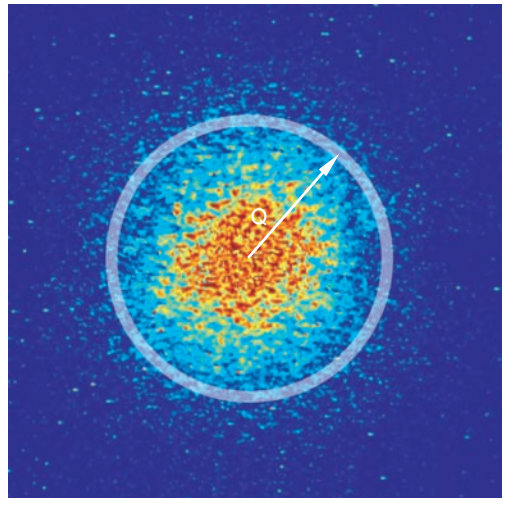

Figure 7.31 *By using an area detector, the statistics of dynamical processes on a scale of $d = 2\pi/Q$ can be significantly enhanced by averaging the speckle signal along the circumference of a ring of radius Q (shown here as a semitransparent white circle).*

complex and large-scale atomic structures. This drive has already resulted in several advances in scattering and diffraction analysis [12, 26, 29, 33] and still more exciting developments can surely be expected in the near future.

References

[1] X. P. Dong, P. C. J. Donoghue, H. Cheng, and J. B. Liu, "Fossil embryos from the middle and late Cambrian period of Hunan, south China," *Nature*, 427, pp. 237–240, 2004.

[2] X. P. Dong, "Cambrian fossil embryos from western Hunan, south China," *Acta. Geol. Sinica Engl. Ed*., 83, pp. 429–439, 2009.

[3] P. C. J. Donoghue, S. Bengtson, X. P. Dong, N. J. Gostling, T. Huldtgren, J. A. Cunningham, C. Yin, Z. Yue, F. Peng, and M. Stampanoni, "Synchrotron x-ray tomographic microscopy of fossil embryos," *Nature*, 442, pp. 680–683, 2006.

[4] J. H. Kinney and M. C. Nichols, "X-ray tomographic microscopy (XTM) using synchrotron radiation," *Ann. Rev. Mater. Sci*., 22, pp. 121–152, 1992.

[5] U. Bonse and F. Busch, "X-ray computed microtomography (μ-CT) using synchrotron radiation (SR)," *Prog. Biophys. Mol. Biol*., 65, pp. 133–169, 1996.

[6] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. Soc. Indust. Appl. Math., 2001.

[7] J. Radon, "On the determination of functions from their integral values along certain manifolds," *IEEE Trans. Med. Imaging*, 5, pp. 170–176, 1986.

[8] R. Mokso, F. Marone, and M. Stampanoni, "Real time tomography at the Swiss Light Source," *AIP Conf. Proc*., 1234, pp. 87–90, 2010.

[9] P. Tafforeau, R. Boistel, E. Boller, A. Bravin, M. Brunet, Y. Chaimanee, P. Cloetens, M. Feist, J. Hoszowska, J. J. Jaeger, R. F. Kay, V. Lazzari, L. Marivaux, A. Nel, C. Nemoz, X. Thibault, P. Vignaud, and S. Zabler, "Applications of x-ray synchrotron microtomography for non-destructive studies of paleontological specimens," *Appl. Phys. A*, 83, pp. 195–202, 2006.

[10] P. Cloetens, R. Barrett, J. Baruchel, J. P. Guigay, and M. Schlenker, "Phase objects in synchrotron radiation hard x-ray imaging," *J. Phys. D*, 29, pp. 133–146, 1996.

[11] T. J. Davis, D. Gao, T. E. Guruyev, A. W. Stevenson, and S. W. Wilkins, "Phase-contrast imaging of weakly absorbing materials using hard x-rays," *Nature*, 373, pp. 595–598, 1995.

[12] A. Snigirev, I. Snigireva, V. Kohn, S. Kuznetsov, and I. Schelokov, "On the possibilities of x-ray phase contrast microimaging by coherent high-energy synchrotron radiation," *Rev. Sci. Instrum*., 66, pp. 5486–5492, 1995.

[13] A. Momose, T. Takeda, Y. Itai, and K. Hirano, "Phase-contrast x-ray computed tomography for observing biological soft tissues," *Nature Med*., 2, pp. 473–475, 1996.

[14] D. Chapman, W. Thomlinson, R. E. Johnston, D. Washburn, E. Pisano, N. Gmür, Z. Zhong, R. Menk, F. Arfelli, and D. Sayers, "Diffraction enhanced x-ray imaging," *Phys. Med. Biol*., 42, pp. 2015–2025, 1997.

[15] A. Momose, "Phase-sensitive imaging and phase tomography using x-ray interferometers," *Opt. Express*, 11, pp. 2303–2314, 2003.

[16] S. A. McDonald, F. Marone, C. Hintermüller, G. Mikuljan, C. David, F. Pfeiffer, and M. Stampanoni, "Advanced phase-contrast imaging using a grating interferometer," *J. Synchrotron Rad*., 16, pp. 562–572, 2009.

[17] G. McDermott, M. A. Le Gros, C. G. Knoechel, M. Uchida, and C. A. Larabell, "Soft x-ray tomography and cryogenic light microscopy: the cool combination in cellular imaging," *Trends Cell Biol*., 19, pp. 587–595, 2009.

[18] G. Schneider, E. Anderson, S. Vogt, C. Knoechel, D. Weiss, M. Le Gros, and C. Larabell, "Computed tomography of cryogenic cells," *Surf. Rev. Lett*., 9, pp. 177–183, 2002.

[19] D. Y. Parkinson, G. McDermott, L. D. Etkin, M. A. Le Gros, and C. A. Larabell, "Quantitative 3-D imaging of eukaryotic cells using soft x-ray tomography," *J. Struct. Biol*., 162, pp. 380–386, 2008.

[20] E. Hanssen, C. Knoechel, N. Klonis, N. Abu-Bakar, S. Deed, M. Le Gros, C. A. Larabell, L. Tilley, and P. Falciparum, "Cryo transmission x-ray imaging of the malaria parasite,," *J. Struct. Biol*., 173, pp. 161–168, 2011.

[21] K. Giewekemeyer, P. Thibault, S. Kalbfleisch, A. Beerlink, C. M. Kewish, M. Dierolf, F. Pfeiffer, and T. Salditt, "Quantitative biological imaging by ptychographic x-ray diffraction microscopy," *Proc. Nat. Acad. Sci. USA*, 107, pp. 529–534, 2010.

[22] M. Stampanoni, R. Mokso, F. Marone, J. Vila-Comamala, S. Gorelick, P. Trtik, K. Jefimovs, and C. David, "Phase-contrast tomography at the nanoscale," *Phys. Rev. B*, 81, p. 140105, 2010.

[23] P. Tafforeau, "Synchrotron light illuminates the orang-utan's obscure origins," *Science in School*, 5, pp. 24–27, 2007.

[24] Y. Chaimanee, D. Jolly, M. Benammi, P. Tafforeau, D. Duzer, I. Moussa, and J.-J. Jaeger, "A middle Miocene hominoid from Thailand and orangutan origins," *Nature*, 422, pp. 61–65, 2003.

[25] Y. Chaimanee, C. Yamee, P. Tian, K. Khaowiset, B. Marandat, P. Tafforeau, C. Nemoz, and J.-J. Jaeger, "Khoratpithecus piriyai, a late Miocene hominoid of Thailand," *Am. J. Phys. Anthropol*., 131, pp. 311–323, 2006.

[26] J. W. Miao, P. Charalambous, J. Kirz, and D. Sayre, "Extending the methodology of x-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens," *Nature*, 400, pp. 342–344, 1999.

[27] V. Elser, "Phase retrieval by iterated projections," *J. Opt. Soc. Am. A*, 20, pp. 40–55, 2003.

[28] P. Thibault, V. Elser, C. Jacobsen, D. Shapiro, and D. Sayre, "Reconstruction of a yeast cell from x-ray diffraction data," *Acta Crystallogr. A*, 62, pp. 248–261, 2006.

[29] D. Shapiro, P. Thibault, T. Beetz, V. Elser, M. Howells, C. Jacobsen, J. Kirz, E. Lima, H. Miao, A. M. Neiman, and D. Sayre, "Biological imaging by soft x-ray diffraction microscopy," *Proc. Nat. Acad. Sci*., 102, pp. 15343–15346, 2005.

[30] M. A. Pfeifer, G. J. Williams, I. A. Vartanyants, R. Harder, and I. K. Robinson, "Three-dimensional mapping of a deformation field inside a nanocrystal," *Nature*, 442, pp. 63–66, 2006.

[31] I. Robinson and R. Harder, "Coherent x-ray diffraction imaging of strain at the nanoscale," *Nature Mater*., 8, pp. 291–298, 2009.

[32] J. M. Rodenburg, A. C. Hurst, A. G. Cullis, B. R. Dobson, F. Pfeiffer, O. Bunk, C. David, K. Jefimovs, and I. Johnson, "Hard x-ray lensless imaging of extended objects," *Phys. Rev. Lett*., 98, p. 034801, 2007.

[33] P. Thibault, M. Dierolf, A. Menzel, O. Bunk, C. David, and F. Pfeiffer, "High-resolution scanning x-ray diffraction microscopy," *Science*, 321, pp. 379–382, 2008.

[34] O. Bunk, M. Dierolf, S. Kynde, I. Johnson, O. Marti, and F. Pfeiffer, "Influence of the overlap parameter on the convergence of the ptychographical iterative engine," *Ultramicroscopy*, 108, pp. 481–487, 2008.

[35] P. Fratzl, H. F. Jakob, S. Rinnerthaler, P. Roschger, and K. Klaushofer, "Position-resolved small-angle x-ray scattering of complex biological materials," *J. Appl. Crystallogr*., 30, pp. 765–769, 1997.

[36] O. Bunk, M. Bech, T. H. Jensen, R. Feidenhans'l, T. Binderup, A. Menzel, and F. Pfeiffer, "Multimodal x-ray scatter imaging," *New J. Phys*., 11, p. 123016, 2009.

[37] M. Sutton, S. G. J. Mochrie, T. Greytak, S. E. Nagler, L. E. Berman, G. A. Held, and G. B. Stephenson, "Observation of speckle by diffraction with coherent x-rays," *Nature*, 352, pp. 608–610, 1991.

[38] S. Brauer, G. B. Stephenson, M. Sutton, R. Brüning, E. Defresne, S. G. J. Mochrie, G. Grübel, J. Als-Nielsen, and D. L. Abernathy, "X-ray intensity fluctuation spectroscopy observations of critical dynamics in Fe_3Al," *Phys. Rev. Lett*., 74, pp. 2010–2013, 1995.

Appendix

Physical Constants Relevant to Synchrotron Radiation

Table 1

Constant	Symbol	Value
Speed of light in vacuum	c	$2.99792458 \times 10^{8}\,\mathrm{m\,s^{-1}}$
Planck's constant	h	$6.62606876 \times 10^{-34}\,\mathrm{J\,s}$
Reduced Planck's constant	$\hbar = h/2\pi$	$1.054571596 \times 10^{-34}\,\mathrm{J\,s}$
Rest mass of an electron	m	$9.10938188 \times 10^{-31}\,\mathrm{kg}$
Rest mass energy of an electron	mc^2	$81.871041\,\mathrm{fJ} = 510.99888\,\mathrm{keV}$
Elementary charge	e	$1.60217653 \times 10^{-19}\,\mathrm{C}$
Permeability of free space	$\mu_0 = 4\pi \times 10^{-7}$	$1.2566371 \times 10^{-6}\,\mathrm{V\,s/(A\,m)}$
Permittivity of free space	$\epsilon_0 = 1/\mu_0 c^2$	$8.85418782 \times 10^{-12}\,\mathrm{A\,s/(V\,m)}$
Fine structure constant	$\alpha = \mu_0 c e^2/2h$	1/137.03599976
Boltzmann's constant	k_B	$1.3806503 \times 10^{-23}\,\mathrm{J\,K^{-1}}$
Avogadro's number	N_A	$6.02214199 \times 10^{23}\ \mathrm{mol^{-1}}$
Absolute zero	θ_0	$-273.15\ ^{\circ}\mathrm{C}$
Gas constant	$R = kN_A$	$8.314472\,\mathrm{J\,K^{-1}\,mol^{-1}}$
Normal pressure	p_n	101 325 Pa

An Introduction to Synchrotron Radiation: Techniques and Applications, First Edition. Philip Willmott.

Index

Note: Page numbers in *italics* refer to figures, page numbers in **bold** refer to tables.

An Introduction to Synchrotron Radiation: Techniques and Applications, First Edition. Philip Willmott.