2

Generic methodologies for nanotechnology: characterization

2.1 GENERAL CLASSIFICATION OF CHARACTERIZATION METHODS

In this section we will confine ourselves to a general description of characterization methods for both imaging (*microscopy*) and analysis (*spectroscopy*); in subsequent sections we discuss in more detail some of the common techniques employed for the investigation of nanostructures. Considering the large range of possible physical analysis techniques, very few microscopical analytical techniques are currently in widespread use. Generally there exist around ten possible primary probes (e.g., electrons, X-rays, ions, atoms, light (visible, ultraviolet, infrared), neutrons, sound) which may be used to excite secondary effects (electrons, X-rays, ions, light, neutrons, sound, heat, etc.) from the region of interest in the sample. The chosen secondary effects may then be monitored as a function of one or more of at least seven variables (e.g., energy, temperature, mass, intensity, time, angle, phase). This implies a theoretical maximum of around 700 single-signal characterization techniques (although some permutations of probe and effect are physically impractical) as well as the technically more difficult multi-signal techniques. To date, around 100 different techniques have been attempted, most of which use ions, electrons, neutrons or photons as the primary probes.

When ions are used as probes, generally ions and/or photons are emitted from the sample and analysed; examples include secondary ion mass spectrometry (SIMS), Rutherford backscattering spectrometry (RBS) and proton-induced X-ray emission

Nanoscale Science and Technology Edited by R. W. Kelsall, I. W. Hamley and M. Geoghegan
 ISBN 0-470-85086-8

(PIXE). When electrons are used as probes, generally electrons and/or photons are emitted and analysed; examples include scanning electron microscopy (SEM), electron probe microanalysis (EPMA), analytical transmission electron microscopy (TEM) including energy-dispersive X-ray (EDX) analysis and electron energy loss spectroscopy (EELS), Auger electron spectroscopy (AES), low-energy electron diffraction (LEED) and reflection high-energy electron diffraction (RHEED). When photons are used as probes, generally electrons and/or photons are emitted and analysed; examples include light microscopy, X-ray diffraction (XRD), X-ray fluorescence (XRF), X-ray absorption spectroscopy (XAS), infrared spectroscopy (IR), Raman spectroscopy and X-ray photoelectron spectroscopy (XPS).

In recent years there has been the development of proximal probe techniques such as atomic force microscopy (AFM), scanning tunnelling microscopy (STM) and scanning tunnelling spectroscopy (STS) which monitor the interaction between a localized probe and a sample surface. In other cases, such as position-sensitive atom probe (POSAP) spectroscopy, the sample itself is effectively the probe.

2.1.1 Analytical and imaging techniques

Before using a particular characterization technique it is essential to consider what information is required about a sample and at what resolution. Simplistically we may think of this information as being divided into:

- morphology (the microstructural or nanostructural architecture);
- crystal structure (the detailed atomic arrangement in the chemical phases contained within the microstructure);
- chemistry (the elements and possibly molecular groupings present);
- electronic structure (the nature of the bonding between atoms).

In addition, information on the mechanical, thermal or electronic properties of the sample may also be required. However, many of these properties will arise as a direct consequence of the structural factors outlined above and, ultimately, we need to establish a correlation between nanostructure and properties and determine how both these factors depend on the processing conditions employed to fabricate the material or the device. Eventually this information may be employed as input to computer-based modelling procedures, which may then allow the prediction of nanostructure–property–processing relationships.

Having decided on the type of information required, the next question to address is from what sample volume or area should this information be obtained? The resolution of a particular technique may be conveniently divided into lateral or spatial resolution (i.e., from what area on a sample surface or from what volume within a sample a particular signal originates?) and depth resolution (i.e., from how far below the sample surface a signal originates). Very simplistically a large number of techniques may be classified as either surface or bulk (volume) analytical techniques. Different types of nanostructure (e.g., films, dots, particles) require different combinations of lateral and

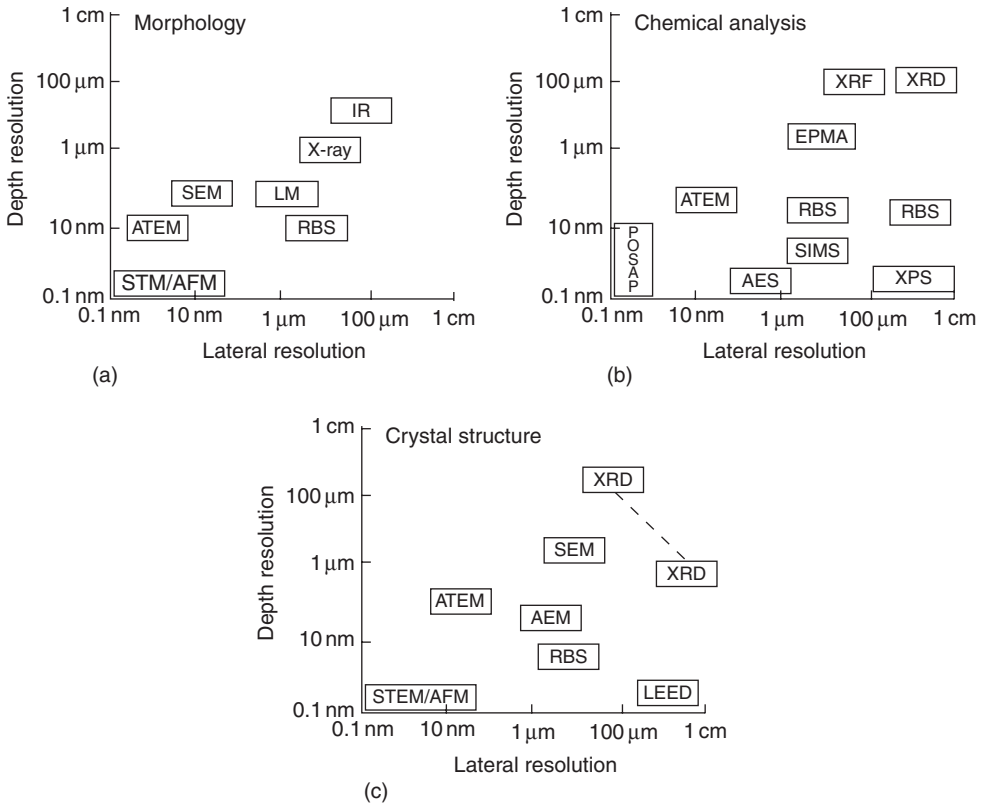

Figure 2.1 Schematic classification of the information content of different imaging and analytical techniques in terms of their lateral and depth resolutions

depth resolutions. Figure 2.1 presents a schematic classification of some of the more common techniques in terms of these two resolution parameters.

2.1.2 Some scattering physics

Electromagnetic radiation (e.g., X-rays, light) can be described in terms of a stream of photons, whilst correspondingly, particles (e.g., neutrons, ions and electrons) can be described in terms of an associated wavelength. This wave–particle duality of incident radiation is summarized by the de Broglie relationship, $\lambda = h/mv$, where λ and v are the wavelength and velocity, h is Planck's constant and m is the mass. When a beam of incident radiation interacts with matter it will be scattered. For each type of radiation there are a variety of possible scattering processes that may occur, each of which will have a particular scattering cross section σ which represents an area the scattering particle (e.g., electron, nucleus or atom in the material) appears to present to the incident particle or radiation. Alternatively the mean free path Λ (which is proportional to the inverse of σ) represents the average distance travelled by the incident particle

between scattering events of a particular type. The frequency of occurrence of different scattering events will depend on the thickness of the specimen and the probabilities will follow Poisson statistics. When many scattering events occur, averaging approaches such as computer-based Monte Carlo simulations are commonly used to model the overall process.

During scattering, the incident wave may undergo changes in amplitude and/or phase, and scattering processes may generally be divided into two types:

- elastic scattering involves no energy transfer during the scattering process (hence no wavelength change of the scattered radiation), although there may be changes in the direction of the incident wave following scattering;
- inelastic scattering involves energy transfer during scattering (hence a gain or loss in the energy of the associated particles or quasiparticles) and changes in the direction of the incident wave.

Another distinction is made between (spatially) coherent scattering, where the phase relationship between scattered waves from neighbouring scattering centres is preserved, and incoherent scattering, where any phase relationships between scattered waves are lost. Generally small-angle elastic scattering is coherent, while inelastic scattering is incoherent; large-angle elastic scattering becomes rapidly more incoherent with increasing scattering angle.

2.1.2.1 X-rays and their interaction with matter

X-rays are produced via the electron bombardment of a metal target or anode. The emission spectrum consists of (a) white radiation or *bremsstrahlung* (from the German for 'braking radiation'), the wavelengths of which are related to the electron energy losses (ΔE) as they are decelerated through Planck's equation, $\Delta E = hc/\lambda$, and (b) characteristic radiations, of discrete wavelengths, which arise from ionization of inner shell electron levels followed by electron relaxation. The emitted X-rays are generally collimated and made monochromatic using filters or crystal monochromators which may be flat or curved to provide a focusing action. Alternatively extremely high intensity X-ray beams are produced by accelerating electrons around the storage ring of a synchrotron and undulating their trajectories ('wiggling') to release electromagnetic radiation of varying wavelength.

When X-rays interact with matter, elastic scattering occurs when X-rays are scattered by the electrons in the material and involves the interaction between the negatively charged electron cloud and the electromagnetic field of the incident X-rays. Electrons respond to the applied field, oscillate and emit an electromagnetic wave (X-ray) identical in wavelength and phase to the incident X-ray. Inelastic scattering occurs when the incident X-ray photons give up all or part of their energy to individual electrons associated with atoms. These electrons are excited to higher energy levels or ionized and escape from the solid as photoelectrons (Section 2.8.1). The ionized atoms in the solid undergo de-excitation to produce a variety of secondary signals. Alternatively, X-rays can lose part of their energy to an electron in a high-energy collision known as

Compton scattering. Overall, X-rays tend to undergo mainly elastic scattering with heavy elements (i.e., those with a large atomic number Z), while for light elements inelastic scattering predominates.

2.1.2.2 Electrons and their interaction with matter

Electrons are produced either by thermionic (heat) or by (electric) field emission from sharp metallic tips. Low-angle (1–10°), coherent elastic scattering of electrons occurs via the interaction of the electrons with the electron cloud associated with atoms in a solid. High-angle, incoherent elastic (back)scattering (10–180°) occurs via interaction of the negatively charged electrons with the nuclei of atoms. The cross section for elastic scattering of electrons varies as the square of the atomic number of the element.

Inelastic scattering of electrons occurs through smaller angles than elastic scattering and the cross section varies linearly with atomic number. Inelastic scattering of electrons occurs predominantly via the four major mechanisms outlined below.

Phonon scattering

The incident electrons excite phonons (atomic vibrations) in the sample, typically the energy loss is <1 eV, the scattering angle is quite large ($\sim 10°$) and the mean free path Λ is $\sim 1\mu$m for carbon. This is the basis for heating the specimen by an electron beam.

Plasmon scattering

The incident electrons excite collective, resonant oscillations (plasmons) of the valence electrons of a solid. The energy loss from the incident beam is 5–30 eV and Λ is ~ 100 nm, causing this to be the dominant scattering process in electron–solid interactions.

Single-electron excitation

The incident electron transfers energy to single electrons, resulting in ionization of atoms. The mean free path for this event is of the order of micrometres. Lightly bound valence electrons may be ejected from atoms, and if they escape from the specimen surface, they may be used to form secondary electron images in the SEM. Energy losses for such excitations typically range up to 50 eV. If inner shell electrons are removed, the energy loss can be up to keV. For example, the energy loss required to ionize carbon 1 s (K shell) electrons is 284 eV. The energy loss of the incident beam can be used in EELS analysis and the secondary emissions (e.g., X-ray or Auger electron production) produced when the ionized atom relaxes can also be used in analysis (energy-dispersive or wavelength-dispersive X-ray spectroscopy) as discussed in Section 2.7.3.

Direct radiation losses

The principal direct radiation losses are the bremsstrahlung X-ray emission caused by the deceleration of electrons, as described above. The energy losses can approach the total incident beam energy in the limit of full deceleration.

2.1.2.3 Neutrons and their interaction with matter

Low-energy thermal neutrons produced in nuclear reactors are most commonly used in materials analysis. Neutrons are neutral particles but have a relatively large mass and can interact with atomic nuclei through nuclear and magnetic forces. They can undergo elastic scattering and also inelastic scattering where the low-energy (slow) thermal neutrons induce atomic vibrations (phonons). High-energy (fast) neutrons can displace and even ionize atoms.

2.1.2.4 Ions and their interaction with matter

In comparison with electrons, ions are relatively heavy, negatively or positively charged particles and various effects occur upon their interaction with matter. These include ion backscattering (Rutherford backscattering spectrometry), the excitation of electrons and photons, X-ray bremsstrahlung, the displacement of atoms and sputtering, as well as the possible implantation of ions within the surface of the material. The latter is extensively used for doping semiconductors.

2.1.2.5 Elastic scattering and diffraction

The atomic scattering amplitude (factor) $f(\theta)$ describes the amplitude of X-rays, electrons or neutrons that are elastically scattered by individual atoms. For X-rays and electrons, this function decreases with increasing angle. In the case of X-rays, it is primarily the electrons (rather than the nucleons) which give rise to elastic scattering and $f_x(\theta)$ is defined as the ratio of the amplitude scattered by the atom to that of a single electron. It is a dimensionless number, equal to Z at low scattering angles but decreases below Z with increasing angles because of interference arising from path differences between the scattered rays from all Z electrons. In the case of electrons, it is primarily the protons in the atomic nuclei which give rise to elastic scattering. However, because of the shielding effect of the electrons, which is proportional to f_x, the electron scattering amplitude $f_e(\theta)$ is proportional to $(Z-f_x)$. It is normally expressed as a scattering length, hence electron intensities (the square of the amplitude) are expressed as areas or atomic scattering cross sections. On average, f_e is proportional to $Z^{3/2}$. Similarly, neutrons are scattered by the nucleons in atoms, but since they also possess spin, they interact and are scattered by unpaired electrons in magnetic materials. The neutron scattering amplitude (again expressed as a scattering length) varies in an irregular way with Z and also between different isotopes of the same element; in particular, hydrogen and oxygen scatter strongly in comparison with heavy atoms. Neutron scattering therefore complements X-ray and electron scattering in providing information about the positions of light atoms and atoms of similar Z. The important practical difference between X-ray, electron and neutron elastic scattering is that the magnitude of the scattering amplitude for electrons is roughly 10^4 times greater than the magnitude for X-rays, which are in turn greater than the magnitude for neutrons.

Since the elastically scattered X-rays, neutrons and electrons are coherent, they interfere and give rise to diffraction patterns; the diffracted beams arise from strong constructive interference from all the atoms in the material. The nature of the diffraction

pattern depends on the nature of the material. For perfect crystals, with long-range periodic order, the diffraction pattern is sharply defined. As crystals become smaller and also when they become increasingly imperfect (due to the presence of defects and internal stresses), the diffraction pattern is less sharply defined. In amorphous materials (such as glasses and liquids) the diffraction pattern shows diffuse maxima which are related to the average interatomic spacings between the atoms; a radial profile through the diffraction pattern essentially gives a radial distribution function of the atoms surrounding a particular scattering centre in the amorphous solid. In short, in all cases diffraction patterns reveal the nature (i.e., the structure and the imperfections) of the material under investigation.

In crystalline solids, atoms are arranged periodically in regular, three-dimensional repeat motifs known as unit cells and this also leads to a corresponding periodic distribution of electrons. Waves coherently scattered from such periodically arranged centres therefore undergo interference leading to intense diffracted waves in specific directions which will depend on the exact three-dimensional periodicity (the crystallography). For radiation of wavelength λ incident on a crystal whose lattice planes are regularly spaced by a distance d, the allowed angles θ for diffraction are given by the Bragg equation

$$m\lambda = 2d\sin\theta \tag{2.1}$$

for integer m. Summing the atomic scattering amplitudes of atoms in a particular plane of Miller indices (hkl), taking into account the path differences between scattered waves, allows the calculation of a structure factor F_{hkl}, the scattering amplitude for diffraction from a particular set of (hkl) atomic planes. The analysis of the angular distribution and the intensities of these diffracted waves (usually for a given wavelength of incident radiation) therefore provides information on the atomic identities and arrangements in the material. An introduction to the relationship between crystallography and diffraction is provided in Section 2.7. Incoherent high-angle elastic scattering is reasonably unaffected by this periodicity and depends only on the nature of the atoms not their relative positions and generally follows the decreasing $f(\theta)$ function at high angles.

A summary of the analytical applications of a variety of interactions between radiation and matter is presented in Table 2.1. It also provides a basis for the classification of many of the imaging and analytical techniques that are discussed in the following sections.

2.2 MICROSCOPY TECHNIQUES

Radiation–matter interactions provide a wide range of analytical and spectroscopic techniques, and although it is possible to obtain an analytical signal from a single nanostructured system or an averaged signal from a collection of nanostructured systems, images are a very powerful means of presenting information. Imaging the morphology, as well as the structural and chemical arrangements in nanostructures, necessarily requires microscopical techniques capable of high spatial and/or depth resolution.

Table 2.1 Summary of radiation–matter interactions and their analytical applications for characterization of nanomaterials (see Section 2.1 for a list of acronyms)

Primary radiation	Modification of matter	Modification of radiation	Secondary radiation (signal)	Applications
Photons (e.g., X-rays)	Elastic vibration	Elastic and quasi-elastic scattering		IR, Raman, light microscopy, XRD
	Recoil electron	Inelastic scattering		Compton spectrometry
	Bond breaking	Inelastic scattering: absorption		Photolithography
	Ionization			Optical absorption, XAS
			Photoelectrons	XPS
	De-excitation		Photons	Luminescence, XRF
			Auger electrons	Auger, XPS
Electrons	Thermal vibration	Elastic and quasi-elastic scattering		Backscattered electron imaging in SEM/TEM imaging and electron diffraction
	Plasmon	Inelastic scattering: energy loss		EELS
Electrons	Bond breaking	Inelastic scattering: absorption		Electron beam lithography
	Atom displacement Sputtering			
	Ionization		Bremsstrahlung Secondary electrons	EDX or EPMA SEM
		Inelastic scattering: energy loss		EELS or energy-filtered TEM
	De-excitation		X-rays	EDX or EPMA
			Auger electrons	AES
Neutrons	Thermal vibration	Elastic and quasi-elastic scattering		Neutron diffraction and reflectometry
	Bond breaking	Inelastic scattering		
	Atom displacement Sputtering			
Ions	Thermal vibration	Elastic and quasi-elastic scattering		RBS
	Bond breaking	Inelastic scattering: absorption	Emission of ions and radiation (e.g., gamma rays)	NRA
	Atom displacement Sputtering		Secondary ions	SIMS FReS Ion beam lithography
	Ion implantation			Implantation or doping
	Ionization		Bremsstrahlung	
	De-excitation		X-rays	PIXE

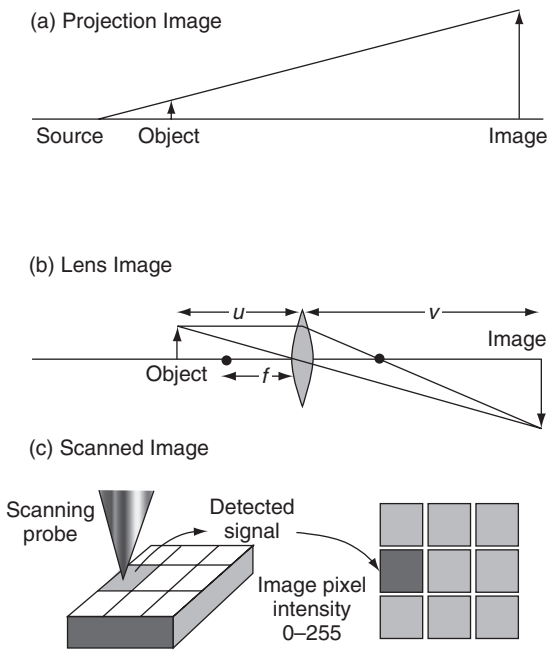

Figure 2.2 Formation of (a) a projection image, (b) a magnified image by a lens and (c) a scanned image

2.2.1 General considerations for imaging

The general purpose of microscopy is to resolve finer details than are discernible with the naked eye, which is 0.1–0.2 mm at the near point of vision, 250 mm. For nanostructured materials we need to resolve details down to a level of 100 nm and below. In transforming an object to an image, there are essentially three methods available, as shown in Figure 2.2:

- A projection image formed in parallel; e.g., field ion microscopy (FIM).
- A lens image formed in parallel; e.g., TEM and light microscopy. Here a lens of focal length f will form an image of an object at a magnification v/u, where u and v are the object and image distances, respectively. The relationship between these parameters is given approximately by the thin lens formula

$$\frac{1}{f} = \frac{1}{v} + \frac{1}{u} \tag{2.2}$$

  from which the magnification v/u can be obtained.

- A scanned image formed in serial; e.g., scanning electron microscopy (SEM), scanning transmission electron microscopy (STEM), confocal laser scanning microscopy (CLSM), scanning tunnelling microscopy (STM) and scanning force microscopy (SFM).

Signals used to form an image can be reflected (e.g., reflected light microscopy or backscattered electrons in SEM), transmitted (e.g., transmitted or polarized light microscopy, or electrons in TEM and scanning TEM) or stimulated (e.g., secondary electrons, cathodoluminesence, X-ray emission or electron beam induced current (EBIC) in SEM, tunnelling current in STM or force in SFM). The concept of picture elements (pixels) is applicable to all images collected in either serial or parallel. Each pixel in the image is defined by x and y coordinates and may be assigned an intensity value; e.g., in terms of grey levels a number between 0 and 255, allowing images to be stored digitally on a computer. Imaging techniques capable of a 3D representation of the object (e.g., tomographic techniques) involve 3D volume elements (voxels). Digital images may be processed to reduce noise (e.g., by use of a framestore), alter contrast (e.g., by use of a non-linear intensity scaling) or analysed and quantified mathematically.

2.2.2 Image magnification and resolution

The magnification is defined as the ratio of the image size to the object size. This can be altered by changing the focal length of the lenses (Equation 2.2). However, making an image bigger does not necessarily increase the resolution. The (spatial) resolution is the smallest separation of two object points that are discernible as separate entities in the image.

When illumination from a point object passes through an aperture – either a physical aperture or a lens – diffraction (scattering and interference) occurs at the aperture giving rise to a series of cones (Airy's rings) with a characteristic intensity distribution. As shown in Figure 2.3, the diameter of the central Airy disc D_d (which contains most of the intensity) is inversely proportional to aperture diameter and is given by $1.22\lambda/n_r \sin \alpha$, where λ is the wavelength of the radiation, n_r is the refractive index of the medium between object and lens, and α is the semi-angle subtended by the aperture. Note that $n_r \sin \alpha$ is often called the numerical aperture (NA). Hence, due to diffraction, each point in the object becomes a small Airy disc in the image, known as a disc of confusion. The *Rayleigh criterion* states that two points in the object may be just distinguished when the maximum intensity of one Airy disc in the image

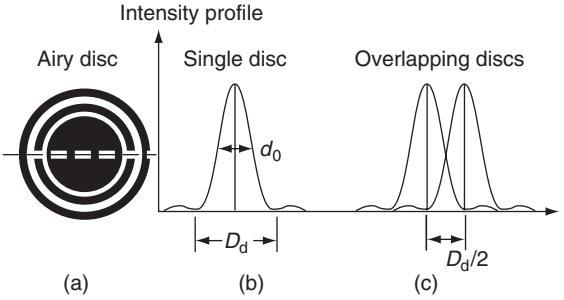

Figure 2.3 Schematic diagram of the Airy diffraction pattern in the image of a point, the intensity variation along the diameter of the Airy pattern and the overlap of the images of two points at the resolution limit

corresponds to the minimum of the second. Hence the limit of resolution r set by diffraction is given by

$$r = \frac{D_d}{2} = \frac{0.61\lambda}{n_r \sin\alpha}. \tag{2.3}$$

In light microscopy the limit of resolution may be decreased by decreasing λ and using green, blue ($\lambda = 0.4\,\mu m$) or UV light ($\lambda = 0.2\,\mu m$), or increasing n_r and using an oil immersion objective lens. Generally the achievable resolution of light microscopy is $0.2\,\mu m$, which is is not generally sufficient for the characterization of nanostructures. However, some specialized light microscopy techniques are applicable that are not diffraction-limited and that allow much lower limits of resolution, such as scanning near-field optical microscopy (SNOM) as discussed in Section 2.2.4.1. In electron microscopy the accelerated electron wavelength is $\sim 10^5$ times smaller than that of light, which leads to increased resolution. For electrons accelerated through a potential V, the corresponding de Broglie wavelength is given by

$$\lambda = \frac{h}{(2m_e eV)^{1/2}}. \tag{2.4}$$

ignoring relativistic effects, where m_e and e are the electron mass and charge, respectively. For an accelerating potential of 100 kV, this gives an electron wavelength of 0.0037 nm. For electrons in a vacuum $n_r = 1$, and electron lenses employ very small values of α ($\sim 10^{-3}$ radians), thus the diameter of the disc of confusion due to Airy diffraction simplifies to $D_d = 1.22\lambda/\alpha$, giving a limit of resolution of ~ 0.2 nm. Electron microscopy techniques are discussed in Section 2.3. Finally, for scanned imaging techniques, the resolution is determined primarily by the size of the scanned probe: the smaller the probe, the higher the resolution. However, many scanning techniques such as SEM and STEM necessarily employ condenser lenses to demagnify the probe, which may limit the attainable resolution.

The depth of field z is the range of positions of the object, on either side of the object plane, which give no loss in resolution. Since the resolution is limited by diffraction, it is relatively simple to show that

$$z = \frac{0.61\lambda}{n_r \sin\alpha \tan\alpha}. \tag{2.5}$$

Electron microscopes operate with small values of α, thus the depth of field is relatively large compared to light microscopy, where α is much bigger. Typical values for a magnification of $\times 1000$ are $z = 40\,\mu m$ (SEM) and $1\,\mu m$ (light microscopy). A large depth of field is particularly useful for imaging three-dimensional (3D) structures and devices.

The depth resolution of a signal and hence an image depends on either the penetration depth of incident radiation or the escape depth of the monitored signal. This can vary from a few atom layers (e.g., XPS, Auger, SIMS) to micrometres (e.g., X-ray emission in SEM).

A large number of imaging techniques (for example, TEM) essentially provide a two-dimensional (2D) projection of what is in reality a 3D object; essentially there is no depth resolution and information is averaged through thickness. Tomographic techniques essentially reconstruct the 3D image, and hence the object, from a set of 2D projections measured along different directions of the incident radiation. An example of tomographic reconstruction using scanning force microscopy is given in Section 2.5.2. Alternatively, confocal techniques are based on images measured at different depths within a structure with a very small depth of field; this focal series may also be reconstructed to form a 3D representation of the object.

2.2.3 Other considerations for imaging

Image points in a particular microstructural feature will only be discernible if they have an intensity I_f greater than the intensity I_b of neighbouring points in the background, regardless of the resolution. The contrast C is given by $C = (I_f - I_b)/I_b$. Signals always contain some degree of noise N which will be given by Poisson statistics as $N = \sqrt{n}$, where n is the number of signal counts at a particular point. One criterion for visibility is the Rose criterion, which states that $(I_f - I_b)$ must be greater than 5 times the noise level.

Lenses for light or electrons are not perfect but possess physical and geometrical aberrations that distort the image and lead to a loss in resolution and quality above and beyond the diffraction-limited resolution. These may be summarized as follows:

- *Chromatic aberration* arises because lenses have different focal lengths for different wavelengths of radiation. If the radiation is not monochromatic (of a single λ or energy) then the lens will cause the radiation to be focused over a range of positions; this leads to a point in the object becoming a disc of confusion in the image.
- *Spherical aberration* arises due to the difference in path lengths of rays travelling through different portions of the lens. An axial ray travelling along the optic axis through the centre of a lens will be focused further away from the lens than a non-axial ray travelling through the extremities of the lens. This also leads to a point in the object becoming a disc of confusion in the image.
- *Astigmatism* arises when the focal length of the lens differs in two axial planes perpendicular to each other. For electromagnetic lenses this may be corrected by the use of externally imposed fields known as stigmators.

In light microscopy it is possible to reduce both chromatic and spherical aberrations by using combinations of concave and convex lens elements. However, this is not usually possible for electromagnetic lenses, where the only solution is to reduce the aperture angle, α, which makes the diffraction limited resolution worse. Also it is possible in both light microscopy, SEM and TEM to make the radiation more monochromatic, thereby minimizing chromatic aberrations. Recent developments in instrumentation allow the computer-controlled correction of spherical aberration in electromagnetic lenses in SEM, TEM and STEM; aberration-corrected STEM can produce sub-angstrom probes for imaging, as evinced by the SuperSTEM facility at Daresbury Laboratories, UK.

2.2.4 Light microscopy

As stated the resolution limits for visible or UV light microscopies are insufficient for the direct visualization of nanostructures. However, such techniques may be extremely useful for the analysis of larger-scale devices and architectures constructed from nanoscale elements. One important point concerning light microscopy is that the *detection limit* for dark lines on a white background (white ground/bright field) is considerably less than the resolution (diffraction) limit, while for dark ground/dark field epi-illumination conditions (white lines on a dark background) there is essentially no detection limit other than that of the photon detector. Indeed, epi-illumination fluorescence light microscopy is based on the detection of probes (e.g., DNA markers) at the Angström level (although not their resolution). However, we do not intend to go into great detail on these topics here and the interested reader is referred to the bibliography for a comprehensive description. Key aspects to note are that, owing to the nature of the radiation, the sample can be viewed under ambient conditions and can also be subjected to dynamic experiments whilst being simultaneously observed.

Light microscopy techniques divide into reflection and transmission methods and both may require some sample preparation such as polishing or thin sectioning. Both transmission and reflection light microscopy can exploit the optical anisotropy (birefringence) of specimens by use of incident plane or circularly polarized light and can give colour contrast that depends on specimen anisotropy, orientation absorption and (in the case of transmitted light) thickness.

Two powerful techniques that have arisen in recent years are confocal methods and near-field optical techniques. As stated the diffraction-limited resolution of light microscopy is generally too large for application to the majority of nanostructures, however, scanning optical microscopy (SOM) has had applications. Here a beam of light is scanned across the surface of a specimen. Scanning is performed either by using a spinning disc with a pattern of holes or by using a focused probe such as a laser. The signal is either reflected, fluorescent or transmitted light. To achieve a good resolution, small-diameter spots are required with a large convergence angle which thus limits the depth of field. Confocal scanning light microscopy involves focusing the imaging signal onto a detector aperture in a plane conjugate with the objective focal plane. As shown in Figure 2.4, this

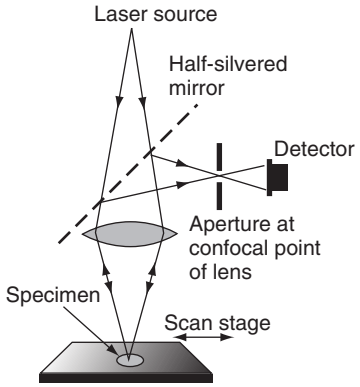

Figure 2.4 Schematic diagram of a confocal scanning light microscope

effectively selects signals from just the focal plane of the objective (i.e., with a very small depth of field). This allows optical sectioning of the image of the specimen, which is useful for quantitative measurements of topography; the sections can be stored and recombined to give an extended focus image.

2.2.4.1 Scanning near-field optical microscopy

Scanning near-field optical microscopy (SNOM) is a scanning optical microscope surface (reflection) technique that provides a limit of resolution beyond the diffraction limit. The sample is illuminated in the near field with light of a chosen wavelength passing through an aperture a few tens of nanometres in diameter formed at the end of a single-mode drawn optical fibre coated with aluminium to prevent light loss, thus ensuring a focused beam from the tip. The probe is usually held fixed while the sample is scanned employing the same piezoelectric drivers used in most commercial scanning probe microscopes (Section 2.5). The tip-to-sample distance may be accurately controlled using a feedback circuit and used to generate both optical images (at a resolution below 100 nm) and topographic images (tip displacement images). The most commonly chosen imaging mode is fluorescence, but other variants include UV-visible, IR and Raman techniques, which can provide chemical information using near-field spectroscopy. Despite the promise that it offers, SNOM remains a less widely used technique than AFM, and is generally perceived as being experimentally challenging.

2.3 ELECTRON MICROSCOPY

2.3.1 General aspects of electron optics

Electron microscopy is an extremely important technique for the examination and analysis of both the surface and subsurface of nanostructured systems (SEM) as well as the bulk structure of thin samples, usually averaged through thickness (TEM and STEM).

In many respects, electron optics may be regarded as analogous to light optics; the main differences may be summarized as follows. Firstly, the wavelength of accelerated electrons is very much smaller than the wavelength of visible or UV photons, which implies a much greater resolution for electron imaging. Secondly, electrons interact much more strongly with matter than photons, thus all optical paths in an electron microscope must usually be under a vacuum of at least 10^{-4} Pa. However a major benefit for nanostructures is that even though the signal is derived from a small amount of material, it will be relatively strong compared to that derived using other incident radiations. Thirdly, electrons are charged particles and may thus be focused by magnetic or electric fields. Furthermore, this focused electron beam may be scanned relatively easily using electrostatic fields. Finally, there is therefore no change in n_r during passage through such a lens, so n_r may be taken as unity, which simplifies the optical formulae. However, in order to reduce lens aberrations when using electromagnetic lenses, the value of α is kept very small and only rays close to the optic axis are employed.

2.3.2 Electron beam generation

The electron beam may be produced either by thermionic emission, where thermal energy is used to overcome the surface potential barrier (work function) of a solid source and so allow extraction of electrons from the conduction band of an emitter, or by field emission, in which an extremely high electric field is employed to reduce the surface potential barrier of an emitter. The finite width of the barrier allows quantum tunnelling of electrons out of the source at room temperature (cold field emission), although higher temperatures have been used more recently (Schottky or thermally assisted field emission), which requires a lower field strength.

A typical thermionic electron gun employs a tungsten cathode filament which is resistively heated to 2800 K in a vacuum of 10^{-4} Pa to give the electrons sufficient energy to overcome the work function of the metal. The emitted electrons are collimated and focused using a Wehnelt cylinder to a beam of typically 50 μm in diameter. The anode potential is typically 1–20 kV in the SEM and 100–200 kV in the TEM.

An important electron gun parameter is the brightness of the source β, the current density per unit solid angle. Increased brightness allows either the use of higher currents, which improve sensitivity and image contrast, or smaller probe areas, which yield higher spatial resolution. The brightness may be increased by using a lanthanum hexaboride (LaB_6) filament that possesses a lower work function than tungsten and a smaller source size (~1 μm), giving an order of magnitude increase in β. Alternatively, a field emission gun (FEG), which employs a high field strength surrounding a pointed tungsten cathode of radius 100 Å, gives a source size as low as 5 nm and a brightness increase of 10^4 over a thermionic tungsten filament. Both sources possess a longer lifetime than tungsten filaments but require a higher vacuum in the gun region. These sources also give a much reduced energy spread of the emitted electrons, leading to improved resolution as a result of reduced chromatic aberration in spectroscopic analysis.

2.3.3 Electron–specimen interactions

Figure 2.5(a) summarises the various signals produced as a result of electron–specimen interactions in SEM and shows the calculated volumes and penetration depths which give rise to the different signals from thick specimens. The *interaction volume* is defined as the volume within which 95% of the electrons are brought to rest by scattering, and has a characteristic teardrop shape (as shown in the figure). The depth and lateral width of electron penetration in the specimen are roughly proportional to V^2 and $V^{3/2}$, respectively, where V is the accelerating voltage. Materials of high (average) atomic number will exhibit reduced electron penetration depths and increased lateral spread relative to those of lower atomic number. For comparison, Figure 2.5(b) shows the signals produced in thin specimens appropriate for TEM imaging. Besides the primary electrons, there are various types of emitted electrons which leave the surface of a bulk sample and may be used for imaging or analysis, as summarized in the following sections.

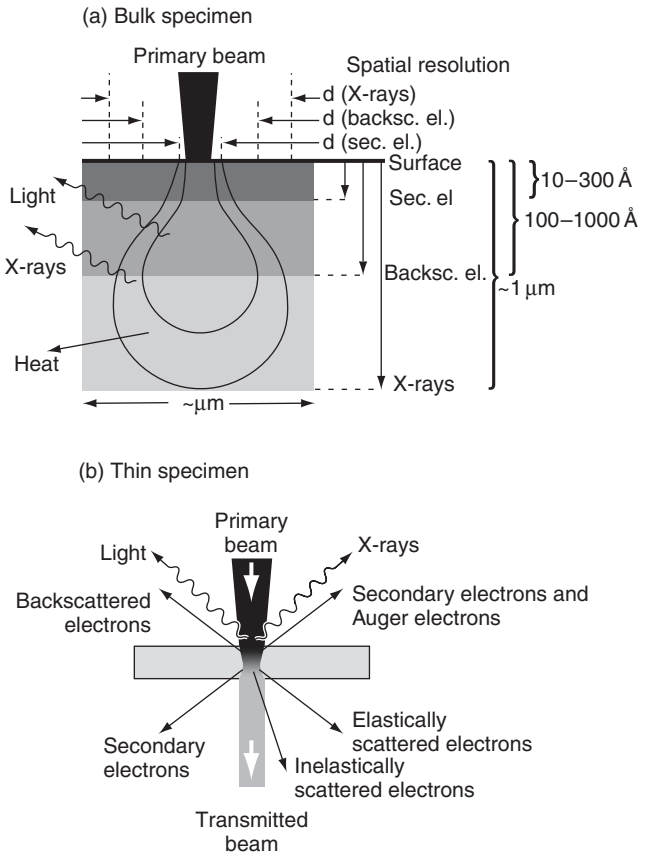

Figure 2.5 Schematic diagram of the beam–specimen interaction in (a) a thick specimen and (b) a thin specimen

2.3.3.1 Secondary electrons

The term 'secondary electrons' is used in general to describe any electrons that escape from the specimen with kinetic energies below about 50 eV. They are most likely to arise from ionized electrons previously associated with atoms close to the surface of the solid which have gained a small amount of kinetic energy and escaped. Alternatively they may be primary electrons which have lost nearly all their energy through scattering and have reached the surface. Secondary electrons are extremely abundant and the secondary electron yield, δ (the number emitted per primary electron), is dependent on the accelerating voltage and can even exceed 1, hence they are extensively used for imaging in SEM.

2.3.3.2 Backscattered electrons

Some primary electrons undergo large deflections and leave the surface with little change in their kinetic energy; these are termed backscattered electrons (BSE). The backscattered

electron yield, η, is almost independent of accelerating voltage and is small compared with δ, but backscattered electrons are also used for imaging in SEM as η is strongly dependent on Z (unlike δ) owing to the fact that they arise from Rutherford backscattering from the nucleus. Owing to this nuclear interaction, backscattered electron imaging can be used to differentiate phases of differing average atomic numbers.

2.3.3.3 Auger electrons and emitted X-rays

Atoms which have undergone inner shell ionization and have been promoted into an excited state by the primary electron beam relax when electrons from higher energy levels drop into the vacant inner shells. This process results in the release of the excess energy between the electron energy levels involved in the transitions (ΔE), producing low-energy (100–1000 eV) *Auger* electrons, or X-rays, or visible photons of wavelength $\lambda = hc/\Delta E$. The Auger yield is small, except in the case of light elements, and is extremely useful for surface analysis (Section 2.8). The energies and wavelengths of the emitted X-rays, which are characteristic of the atom involved, are used in elemental analysis (Section 2.8.1).

2.3.4 Scanning electron microscopy

The scanning electron microscope (SEM) is extremely useful for imaging surface and subsurface microstructure. The basic layout of SEM instrumentation is shown in Figure 2.6.

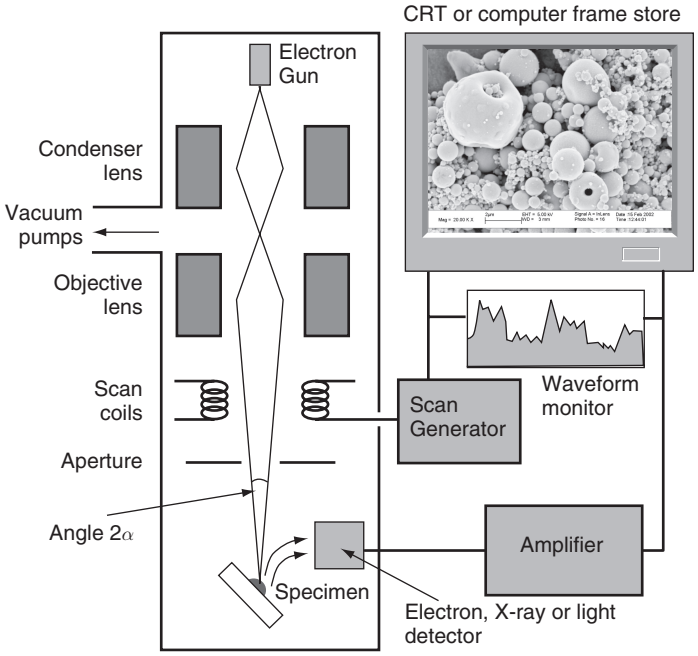

Figure 2.6 Schematic diagram of the layout of a scanning electron microscope

The electron gun usually consists of a tungsten or LaB_6 filament, however FEGs are becoming increasingly necessary for imaging nanostructures at high resolution. The accelerating voltage is usually between 1 and 30 kV. The lower voltages can be used with high-brightness LaB_6 and FEG sources and they are often employed to increase surface detail and obviate sample charging in non-conducting or poorly conducting samples.

In the SEM, two or more condenser lenses are used to demagnify the crossover produced by the gun, while the objective lens focuses the electron probe onto the specimen so that the final probe diameter lies between 2 and 10 nm. The objective aperture limits the angular spread α of the electrons. The focused beam is scanned across the specimen surface in a two-dimensional raster, with the beam passing through the optic axis at the objective lens; meanwhile, an appropriate detector monitors the secondary electrons or other signal, such as backscattered electrons or X-rays, as they are emitted from each point on the surface. Simultaneously, using the same scan generator, a beam is scanned across the recording monitor. The intensity of each pixel on the monitor is controlled by the amplified output of the selected detector and is therefore directly related to the emission intensity of the selected interaction at the corresponding point on the specimen surface. The magnification is simply the ratio of the monitor raster size to the specimen raster size and, as in field ion microscopy, it involves no lenses.

2.3.4.1 Secondary electron imaging

During electron irradiation, a wide spectrum of low-energy secondary electrons are emitted from the surface of the specimen with a high emission cross section, giving images with a high signal-to-noise ratio. They are detected by a scintillator–photomultiplier system known as an Everhart–Thornley (ET) detector positioned to the side of the specimen chamber. Surrounding the scintillator is a grid (a Faraday cage) which has a small positive bias to attract the low-energy secondary electrons that are travelling in all directions away from the specimen surface. High-energy backscattered electrons travelling directly towards the detector will also contribute to the image, however due to the small detector collection angle, this signal is relatively low. The ET detector will collect nearly all secondary electrons and therefore images appear to look down holes and over hills. However, the signal from sample areas in direct line-of-sight of the detector is slightly higher than the signal from 'unseen' parts of the specimen and consequently they will appear brighter, an effect known as shadowing contrast. The detector and a typical secondary electron image are shown in Figure 2.7.

The shape of the interaction volume relative to the specimen surface, and hence the secondary electron yield, depends on the angle of inclination θ_{inc} between the beam and the specimen surface, varying as $1/\cos\theta_{inc}$. This *inclination effect* means that the edges of a spherical particle will appear brighter than the centre and, for studies of specimen topography, specimens are normally tilted 20–40° towards the detector to maximize the signal. Due to increased field strength at sharp edges and spikes on the specimen surface, the emission of secondary and backscattered electrons increases, and such features will appear bright, known as edge contrast. Consequently, owing to the surface specificity and the contrast effects mentioned above, the major use of secondary electron imaging is for topographic contrast. In general, topographic images obtained with secondary electrons have the same appearance as an object illuminated with diffuse

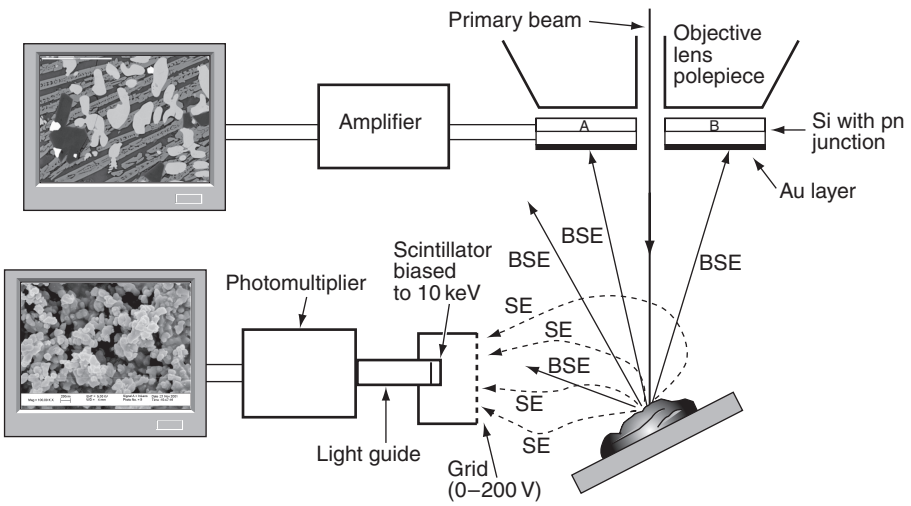

Figure 2.7 Schematic diagram showing a secondary electron detector and a solid-state backscattered electron detector together with typical images

light, in that there are no harsh shadows. Quantitative information on surface topography may be obtained using stereomicroscopy techniques in which two images are taken under identical magnification conditions at slightly different tilt angles (e.g., $-5°$ and $+5°$). The images may be viewed using a special stereo viewer (which allows the brain to interpret them as a 3D picture). For quantitative work, the relative displacement or parallax between two features on the pair of tilted images allows their height separation to be determined.

Since the detected secondary electrons only originate from the top few nanometres of the specimen surface, the emission diameter will only be slightly bigger than the probe diameter (Figure 2.5). The lateral resolution in secondary electron images is therefore of the order of 1–5 nm for typical specimens but will obviously vary with probe size and signal-to-noise considerations. The depth resolution is a function of the accelerating voltage. Secondary electron imaging is therefore the most useful general signal for imaging nanostructures in the SEM due to the high signal level combined with high lateral and depth resolution.

2.3.4.2 Backscattered electron imaging

BSEs have very high energies and low yields compared to secondary electrons. The ET detector may be used to collect only BSEs by making the grid bias slightly negative so as to exclude secondary electrons. However, only those electrons travelling in line-of-sight towards the detector will contribute to the relatively low signal and therefore images will show very strong shadowing contrast and a low signal-to-noise ratio. More normally a large-area solid-state BSE detector is employed; this consists of a thin, annular-shaped semiconductor device (a reverse-biased pn junction) attached to the bottom of the

objective lens, as shown in Figure 2.7. High-energy BSEs incident on this detector cause the formation of electron–hole pairs which, once separated, form a current and may be amplified. The response time is relatively slow, which necessitates the use of slow scan rates.

The resolution in BSE images is typically in the range 25–100 nm, which is significantly worse than in secondary electron images due to the larger penetration depth and thus sampling volume (Figure 2.5). Although the BSE yield η depends sensitively on atomic number Z, it is essential to minimize topographic effects via the use of flat, well-polished specimens if elemental contrast imaging is required.

In a crystalline material, η also depends weakly on the orientation of the crystal with respect to the incident beam. If the beam is parallel to a set of atomic planes, electrons are *channelled* into the crystal structure, giving abnormally large penetration depths, which means that BSEs are less likely to escape to the surface and therefore η is considerably reduced. In a well-polished, undeformed polycrystalline material, grains at different orientations will therefore show different channelling contrasts. Electron backscattered diffraction (EBSD) patterns from a given probe position on the specimen are recorded using a tilted sample and a specially inserted phosphor screen or scintillator to record the angular variation of the BSE yield. When the sample is translated relative to the beam, EBSD may be used to map the relative crystallographic orientations of grains within a microstructure at a resolution $\geq$100 nm.

2.3.4.3 Voltage contrast imaging

Voltage contrast arises since the secondary electron yield is sensitive to the variation in electric fields which may occur at the surface of semiconducting specimens; positive potentials discourage secondary electron emission whereas negative potentials enhance it. Biased semiconductor devices can therefore be imaged in SEM with negatively biased regions appearing bright.

2.3.4.4 Electron beam induced current imaging

Electron beam induced current (EBIC) contrast arises from ionization of valence electrons in the specimen by the incident beam. Normally the electron–hole pairs recombine to produce photons, called cathodoluminescence (Section 2.7.3.2), but in a biased semiconductor specimen the two charge carriers may be separated, giving rise to a current. Measurement of this current allows images to be formed and can give information on conductivity, carrier lifetimes and mobilities. This is useful for the dynamic imaging of integrated devices during operation.

2.3.4.5 Magnetic contrast imaging

Magnetic contrast arises due to the interaction of the external leakage magnetic fields at a specimen surface which can deflect the emitted secondary electrons and BSEs, or via

the interaction of the primary beam with the internal magnetic field of the specimen, which will deflect the beam and change emission yields. In principle both mechanisms allow imaging of magnetic domain structures and can also be employed for Lorentz microscopy in the TEM.

2.3.4.6 Environmental scanning electron microscopy

Environmental scanning electron microscopy (ESEM) or variable pressure SEM relies on imaging in a much degraded vacuum in the SEM chamber, but not the electron gun region. Imaging mechanisms in ESEM (particularly secondary electron imaging) rely on the amplification of the signal by the ionization of gas molecules in the chamber. The introduction of gas into the chamber has a number of benefits as well as drawbacks: (a) It is possible to image uncoated non-conducting samples since specimen charging is compensated for by attraction of ionized gas molecules or electrons to the specimen surface. (b) It is possible to image specimens which are in the hydrated state or even reacting in-situ within the SEM. In particular, dehydration of an SEM sample within the vacuum can lead to large changes in specimen morphology. (c) Image resolution is degraded due to interaction of the incident and scattered electrons beams with the gaseous atmosphere.

2.3.5 Transmission electron microscopy

The conventional transmission electron microscope is a key tool for imaging the internal microstructure of ultrathin specimens. The basic TEM instrumentation is summarized in Figure 2.8. The electron gun is usually thermionic tungsten or LaB_6, however FEGs are becoming increasingly common. The accelerating voltage is considerably higher than in an SEM and is typically 100–400 kV, although a number of specialist high-voltage instruments are designed to operate at 1 MV and above. The benefits of high voltage include increased imaging resolution, due to the decreased electron wavelength, and also increased penetration and thus the ability to study thicker samples. Two or more condenser lenses demagnify the probe to typically 1 μm in diameter, although this can be reduced via use of a condenser–objective nanoprobe system. The condenser excitation controls the beam diameter and convergence. The first condenser (C1) controls the demagnification of the crossover (the spot size), whilst the second condenser (C2) controls the size and convergence of the probe at the specimen, and hence the area of the sample that is illuminated. The specimen must be no more than a few hundred nanometres in thickness, and is usually in the form of a 3 mm diameter disc. The specimen is located between the pole pieces of the objective lens. The combination of the objective lens and the projector lens system provides an overall magnification of around 10^6. The *selected area diffraction* (SAD) *aperture* allows selection of a minimum sample area of ~0.1 μm diameter for electron diffraction. Smaller areas (down to a few nanometres in diameter) can be selected using a focused probe rather than an aperture; this is convergent beam electron diffraction (CBED).

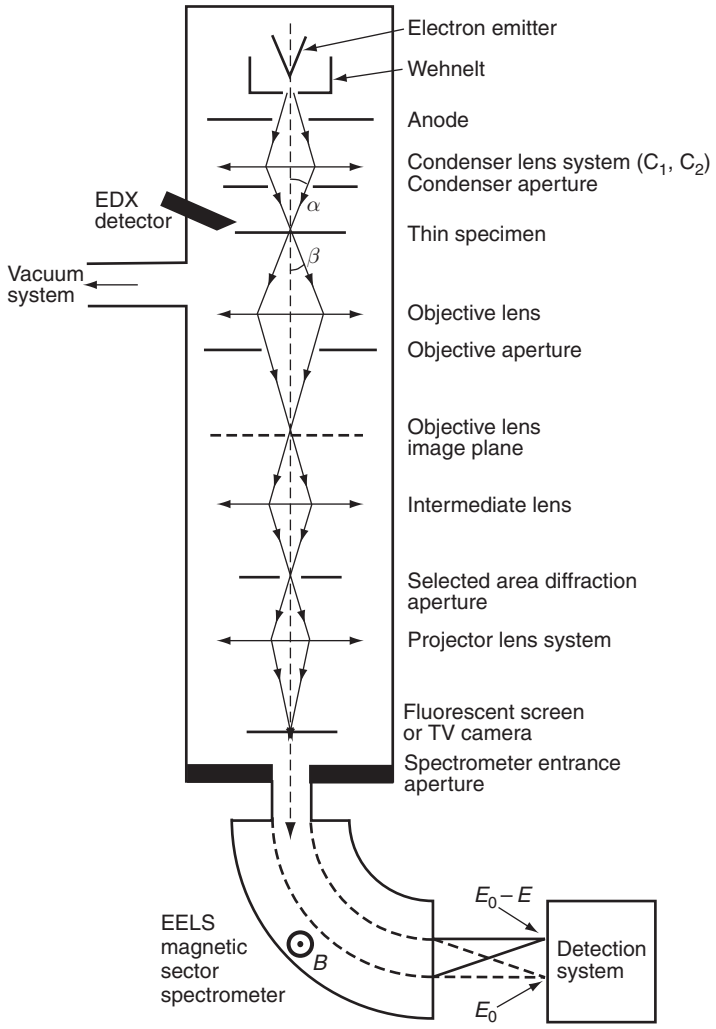

Figure 2.8 Schematic diagram of the layout of an analytical transmission electron microscope

TEM specimens are normally in the form of an ultrathin disc, prepared by cutting, mechanical polishing or chemical dissolution, followed by electropolishing (for conducting materials), chemical polishing (for semiconductors and ceramics) or ion beam milling, including focused ion beam (FIB) machining. Powders may be ground and dispersed onto an amorphous support film on a circular metal grid.

2.3.5.1 Electron diffraction

As described in Section 2.1.2.5, amorphous materials exhibit diffuse diffraction rings related to the average interatomic separations and hence the radial distribution function.

For crystalline materials, the periodic atomic arrangement of atoms scatters the electrons through well-defined angles given by Bragg's law (2.1). Since λ is small (e.g., $\lambda = 0.0037$ nm for 100 keV electrons), the Bragg angles are generally small (of the order of 10^{-3} radians), $\sin\theta \approx \theta$ and therefore $\lambda = 2d\theta$. Hence diffraction occurs only when the planes of atoms are closely parallel to the incident beam, as shown in Figure 2.9.

Since the TEM specimens are so thin, the diffraction conditions along an axis normal to the specimen are significantly relaxed, such that reflections will be excited even when Bragg's law is not exactly satisfied. Thus when the beam is parallel to a major zone axis there will be a large number of (*hkl*) planes in the zone close to their Bragg angles, giving rise to a large number of diffracted beams. The intensity of a particular diffracted beam, $\mathbf{g} = (hkl)$, is a function of the particular diffracting plane in question, the deviation away from the exact Bragg condition, and the sample thickness.

If the specimen area is polycrystalline, then electron diffraction produces a pattern of concentric rings of radii r which exhibit virtually all possible d spacings, owing to the fact the crystallites are randomly oriented (Figure 2.9). The diffraction pattern may be analysed using the relationship $r/L = 2\theta = \lambda/d$, where L is known as the camera length. Thus r is directly proportional to $1/d$; the constant of proportionality $L\lambda$ is known as the camera constant. For a single-crystal specimen, the diffraction pattern will consist of points, spaced at a distance proportional to $1/d$, aligned in a direction perpendicular to the orientation of the (*hkl*) planes as shown in Figure 2.9. The electron diffraction pattern essentially represents a scaled section through the reciprocal (Fourier space) lattice normal to the electron beam direction; the beam direction is simply the vector cross-product of two reciprocal lattice vectors in the plane. Analysis of the diffraction patterns obtained along different incident beam directions, which represent differing projections of the reciprocal lattice, allows the extraction of the real space lattice and hence the unit cell of the material.

Any restriction in the dimensions or morphology of the sample, such as plate-like precipitates oriented parallel to incident beam direction, will lead to changes in the form

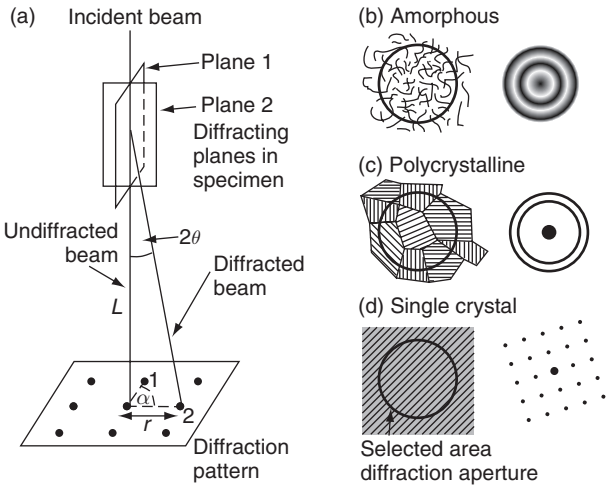

Figure 2.9 Schematic diagram of (a) the geometry of electron diffraction in the TEM and the form of the diffraction pattern for (b) amorphous, (c) polycrystalline and (d) single-crystalline sample regions

of the diffraction pattern such as streaking of the diffraction spots. Furthermore, materials with ultrafine grain sizes will exhibit a degree of broadening in their polycrystalline ring patterns, as described in Section 2.6.1. Both these effects are highly relevant for the analysis of nanostructured materials.

If the incident electron beam is essentially parallel then, for a crystalline material, sharp spots are obtained in the electron diffraction pattern. As shown in Figure 2.9, the area of the specimen contributing to the diffraction pattern can be defined by the position of the SAD aperture in the TEM image. However, if the SAD aperture is removed and the incident beam is focused onto the specimen (i.e., it is convergent), then there are a range of possible incident beam directions and thus Bragg angles. This constitutes the CBED technique, and leads to a broadening of the normal diffraction spots into discs, the diameters of which are proportional to the beam convergence semi-angle α. For small diffraction angles a plane in reciprocal space is sampled, whereas at larger diffraction angles other reciprocal lattice planes are intersected and appear in the pattern. CBED patterns can give information on the 3D crystal structure, the symmetry of the unit cell, the sample thickness and the exact lattice parameters to an accuracy of 1 part in 10^4. This ability to identify the structure and orientation from individual microstructural regions, either using a SAD aperture or a focused probe, allows the identification of crystallographic orientation relationships between two adjacent nanostructural components (e.g., a layer and a substrate when viewed in cross section or, alternatively, a matrix and a precipitate).

2.3.5.2 TEM imaging

A TEM image is a two-dimensional projection of the internal structure of a material. There are three basic contrast mechanisms which contribute to all TEM images.

Mass thickness contrast

During transmission through a thin TEM sample, electrons undergo a range of scattering processes which change their energy and angular distribution. The intrinsic bore of the microscope column prevents electrons that have been scattered through an angle greater than $\sim 10^{-2}$ mrad from contributing to the image. In the presence of a sample, regions that are thicker or of higher density will scatter the electrons more strongly and hence more electrons will be scattered through an angle greater than 10^{-2} mrad making these areas appear darker in the image. This simple mass–thickness contrast is exhibited by all specimens, whether amorphous or crystalline. Biological samples are often deliberately stained with a heavy metal such as osmium or uranium in order to increase the image contrast from the structural features into which the stain becomes incorporated; see Figure 8.10(b).

Contrast may be increased by placing an objective aperture in the back focal plane of the objective lens, which will further limit the angular range of electrons allowed to contribute to the image. If this objective aperture is centred on the optic axis, then in the absence of a specimen the image is bright, hence this is termed *bright-field* imaging (Figure 2.10).

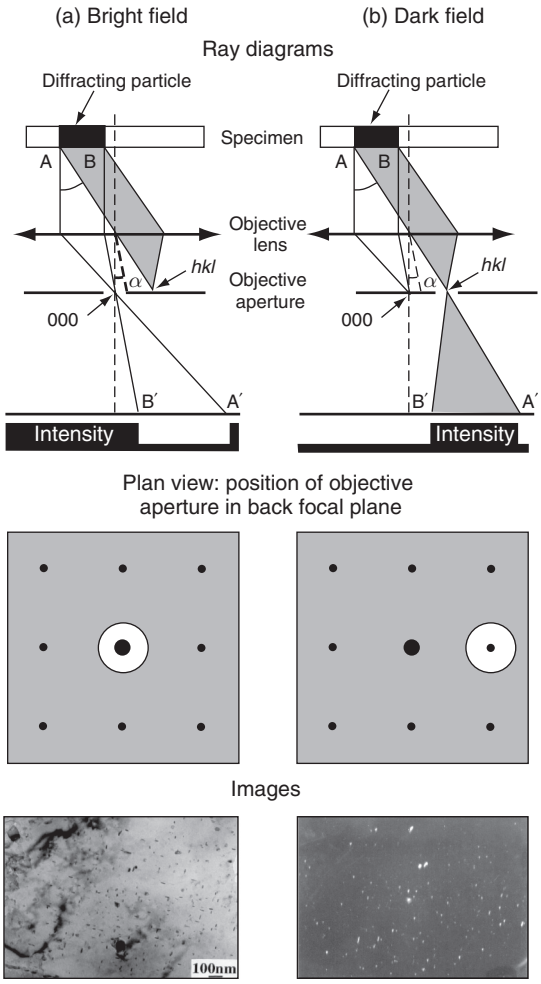

Figure 2.10 TEM ray diagrams for the formation of (a) bright-field and (b) dark-field images together with some typical images of carbide precipitates in a vanadium steel. Image courtesy of Dr Keijan He, Institute for Materials Research, University of Leeds

Diffraction contrast

Diffraction contrast is the major contrast mechanism in crystalline specimens, especially at medium magnifications. It arises from the differing amplitudes of the undiffracted beam and diffracted beams, resulting in intensity variations in the image formed by the different beams. If the objective aperture is centred on the undiffracted beam, then we have a bright-field image as discussed: however, in addition to mass–thickness contrast, any region of the specimen which is in the right orientation (relative to the electron beam direction) for strong diffraction to occur will also appear dark. Correspondingly, if the objective aperture is centred on a diffracted beam, either by displacing the aperture off the optic axis or by tilting the beam so that the diffracted beam travels along the optic axis (giving fewer aberrations in the image) and is used to form the

image, then this is known as *dark-field imaging*, also shown in Figure 2.10. Selection of a particular diffraction spot associated with a specific crystal structure can be used, for example, to identify diffracting crystallites. A large range of common microstructural features show strong diffraction contrast.

Bend (extinction) contours

The closer a crystal plane is to a Bragg orientation, the more strongly it diffracts and the darker it appears in the bright-field image (and correspondingly the lighter it appears in the dark-field image). Most thin specimens are buckled, in which case some areas are closer to a major zone axis than others, leading to bright and dark regions known as bend contours. Similar effects allow direct observation of strain fields in TEM images.

Contrast from variations in specimen thickness

As the direct and undiffracted beams propagate through a crystal, they undergo complementary fluctuations in intensity, the periodicity of which (the extinction distance) depends on the specimen, the operative Bragg planes and their deviation from the exact Bragg angle. Hence wedge-shaped crystals or, for example, inclined grain boundaries show corresponding changes in intensity called thickness fringes, from which the specimen thickness can be calculated.

Contrast from crystal defects

When atoms are displaced by a vector **R** from their lattice positions due to the strain field of a crystal defect, the diffraction conditions are changed and the defect may appear as a different contrast in bright-field and dark-field images depending on which diffracted beams are used for imaging. Generally the appearance of a defect in a dark-field image using a particular Bragg reflection **g** will depend on the magnitude of the product **g.R**. If this scalar product is zero, then this implies that this particular set of **g** planes remain unchanged by the presence of the defect, therefore the diffracted intensity is unaffected and the image is the same as that from a perfect crystal (i.e., the defect is invisible). Other values of this scalar product will lead to some degree of contrast, except for certain special cases. Determination of **R** for a particular defect may be achieved by determining two or more **g** vectors for which the defect is invisible.

Examples of planar defects include stacking faults and grain boundaries. The defect plane in a stacking fault separates the crystal into two parts, shifted through a vector **R** (in addition there might also be a misorientation). Interference between the beams diffracted from the upper and lower parts of the crystal which have undergone different phase changes gives rise to a set of fringes running parallel to the intersection of the fault with the specimen surface. Furthermore, on crossing a grain or phase boundary, the crystal orientation or structure may change and thus the diffraction conditions are unlikely to be similar on both sides of the boundary. If the boundary is inclined, one grain may be strongly diffracting and show thickness fringes whereas the other may show little contrast. If both grains are strongly diffracting, then we may obtain interference between the two sets of fringes producing Moire fringes of a different, usually larger, spacing to the individual interplanar spacings.

An example of a line defect is a dislocation. If a crystal is not in a Bragg condition, then at the core of an edge or screw dislocation the lattice planes are severely bent and must at some point reach the Bragg angle and so diffract. Thus in a bright-field image the dislocation core will normally appear as a dark line.

Examples of three-dimensional defects include precipitates and voids or pores. Incoherent precipitates show diffraction and mass–thickness contrast due to differences in orientation, structure factor or density with respect to the matrix. Coherent precipitates may produce Moire fringes due to interference between two Bragg reflections. Coherent or semicoherent precipitates may also strain the matrix, giving rise to contrast in a similar way to that observed at dislocations. Voids will show some mass–thickness contrast in that they will appear bright. If the void is faceted then effectively the void will produce thickness variations and thus thickness fringes; the void will appear dark on bright thickness fringes, and vice versa.

Phase contrast: lattice imaging

Phase contrast arises when electrons of different phases interfere to form an image. It is present in virtually all TEM images but is generally only visible as a speckled background at high magnifications. Phase contrast is used to form an atomic resolution lattice image by orienting the specimen perpendicular to a major crystallographic zone axis and letting at least two beams pass through the objective aperture; i.e., the undiffracted (000) beam and also the $+\mathbf{g}$ beam and maybe also the $-\mathbf{g}$ beam. The beams interfere and reproduce the periodicity of the object; i.e., a one-to-one correspondence between the observed fringes and the $\mathbf{g}$ lattice planes. If many beams are allowed to pass through the objective aperture then a number of intersecting lattice fringes will be observed, giving rise to spots at the positions of the atomic columns and therefore a structure image (a two-dimensional projection of the atomic structure), as shown in Figure 2.11(d).

2.3.6 Scanning transmission electron microscopy

The general principle of STEM is similar to that of SEM, except that the electron detector is now located below a thin TEM specimen as shown in Figure 2.12. The small probe (of the order of angstroms), produced by a FEG and a condenser–objective lens system, is scanned across the specimen and the signal detected and imaged on a monitor or frame store. The resolution is determined mainly by the probe diameter, which is usually of the order of a nanometre, although recent developments in the correction of spherical aberration in the probe-forming lens have demonstrated sub-angstrom resolution. Many conventional TEMs have a separate scanning attachment; however, dedicated STEMs operate entirely in scanning mode under UHV conditions. Bright-field STEM imaging uses an axial detector to detect electrons scattered through relatively small angles, and images contain diffraction contrast. Dark-field STEM imaging is essentially incoherent and employs an annular detector to detect electrons scattered through higher angles. High-angle annular dark field (HAADF) image intensity is roughly proportional to the square of atomic number (so-called Z contrast) and is therefore extremely useful for mapping phases at ultra-high resolution.

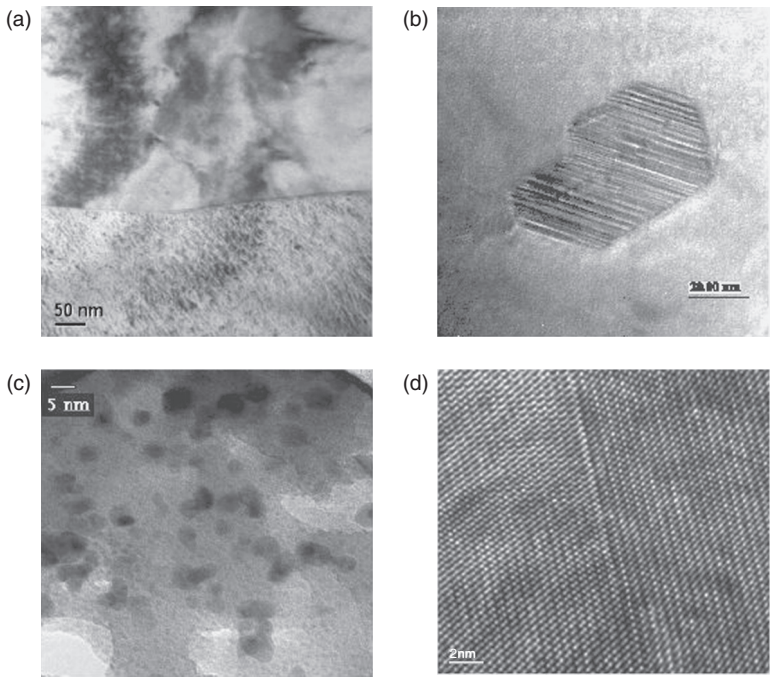

Figure 2.11 Examples of TEM images that exhibit diffraction contrast and phase contrast: (a) ferrite–cementite interface in a steel, (b) twinned titanium boride precipitate in a titanium alloy matrix, (c) nickel catalyst particles on a silica support and (d) grain boundary in zirconia polycrystal. Images courtesy of Drs Andy Brown and Andrew Scott, Institute for Materials Research, University of Leeds

2.4 FIELD ION MICROSCOPY

In the field ion microscope (FIM), the specimen consists of a conducting or semiconducting fine needle-tip of radius r (usually between 5 and 100 nm) prepared by etching or electropolishing techniques, such that individual atoms stand out in a terrace of atomic-scale ledges. The tip is held at a positive potential (between 5 and 20 keV), with respect to a negatively charged conductive fluorescent screen of radius R, in a high vacuum and is maintained at cryogenic temperatures. A low-pressure inert gas, such as hydrogen or helium, is introduced into the chamber, the atoms of which become polarized in the electrostatic field between the screen and the tip and are therefore attracted to the tip by the increased field intensity. The polarized atoms are then ionized and lose electrons (which channel into the tip) either at one point of contact or by 'hopping' over the surface. The resulting positive ions are then accelerated towards the fluorescent screen, giving an instantaneous fluorescent signal. In principle, each fluorescent point on the screen corresponds to an atom in the tip and the high-contrast image consists of a pattern of bright dots on a dark background, which reproduces the atomic arrangement of the tip.

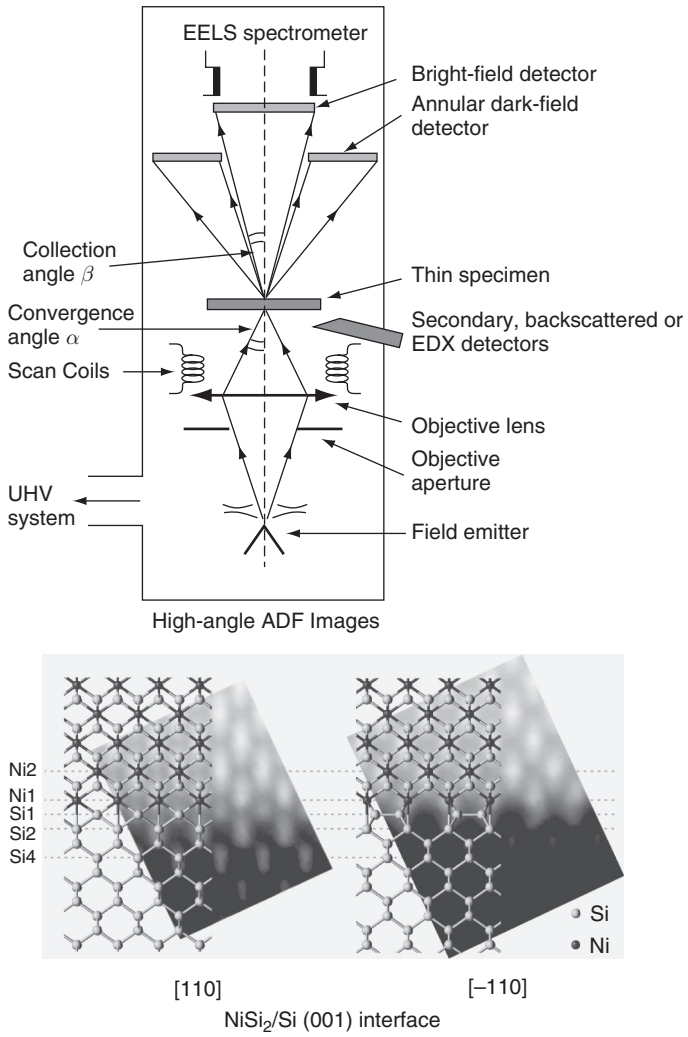

Figure 2.12 Schematic diagram of the layout of a STEM together with HAADF images of a silicon–nickel disilicide interface in two projections with the directly interpretable atomic structure model superimposed. Images courtesy of Dr Andrew Bleloch, Daresbury Laboratories

The image magnification is equal to the ratio of the radius of the screen to that of the tip (R/r) and is of the order of 10^6. More importantly, the limit of resolution is that of an individual atom and the FIM was the first instrument to provide images of atoms and their arrangement at the tip, as well as the location of vacancies and line imperfections. Since no lenses are involved, there are no depth-of-field or diffraction limitations to be considered.

So much for the (potential) advantages of FIM. Apart from the difficulties associated in the preparation of such a fine tip (and the geometrical constraints on a specimen which this imposes) as well as the exclusion of contamination or adsorbed surface

species, there is the problem that the field strength at the tip must be sufficiently high to ionize the gas atoms introduced into the chamber, but not so high as to result in field evaporation or emission of atoms from the tip. In practice only those high melting point elements with high cohesive energies (e.g., Fe, Mo, W, Rh and Pt) and their compounds, which can resist field evaporation at the field strengths needed for inert gas ionization, fulfil this requirement. No breakthrough has been found in the case of gases which ionize more easily. On the other hand, this problem is turned to advantage in the atom probe FIM (APFIM), which involves the controlled progressive removal via field evaporation of (ionized) atoms from the tip. Moreover, the desorbed atoms can be individually chemically identified using a time-of-flight mass spectrometer giving atomic-scale resolution maps of elemental distributions. A recent variant, the position-sensitive atom probe (POSAP), allows a three-dimensional reconstruction of the atomic structure in the tip across an area of about 20 nm^2.

In a sense, APFIM has come full circle from the field emission microscope (FEM), invented prior to the FIM, in which the polarities of the tip and the screen are reversed and which, of course, forms the basis of the field emission gun in electron microscopy. In principle the emitted electrons recorded on the conductive fluorescent screen arise from individual atoms projecting from the tip surface as in FIM, but in practice the images suffer from noise and perturbations in the electron trajectories.

2.5 SCANNING PROBE TECHNIQUES

2.5.1 Scanning tunnelling microscopy

Scanning tunnelling microscopy (STM) is the ancestor of all scanning probe microscopies, which allow the real space imaging of surfaces with atomic resolution and employ no illumination and no lenses! The experimental arrangement is shown in Figure 2.13. A sharp conducting tip (often tungsten), acting as the anode, is brought close to the surface of the specimen (the cathode). A bias voltage ranging from 1 mV to 1 V is applied between the tip and the sample. Above the specimen surface there is an electron cloud due to surface atoms, and when the tip is brought within about 1 nm of the sample surface, electrons can quantum mechanically tunnel across the gap, causing a current to flow. The direction of electron tunnelling across the gap depends on the sign of the bias voltage. It is this current that is used to generate an STM image. The tunnelling current falls off exponentially with distance between tip and surface, and if the tip or sample is scanned laterally using piezoelectric drivers, the STM image reflects the variation in the sample surface topography. If the system is carefully damped from mechanical vibrations, the STM image will possess a sub-angstrom resolution vertically and atomic resolution laterally. The tunnelling current also depends on the atomic species present on the surface and their local chemical environment. In principle, no vacuum is required except when studying adsorption of species on surfaces. However, many STMs are operated under UHV as any oxide or contaminant can interfere with the tunnelling current. Furthermore, it is important to note that an STM cannot image insulating materials, except under conditions where they have appreciable conductivity, such as at high temperature.

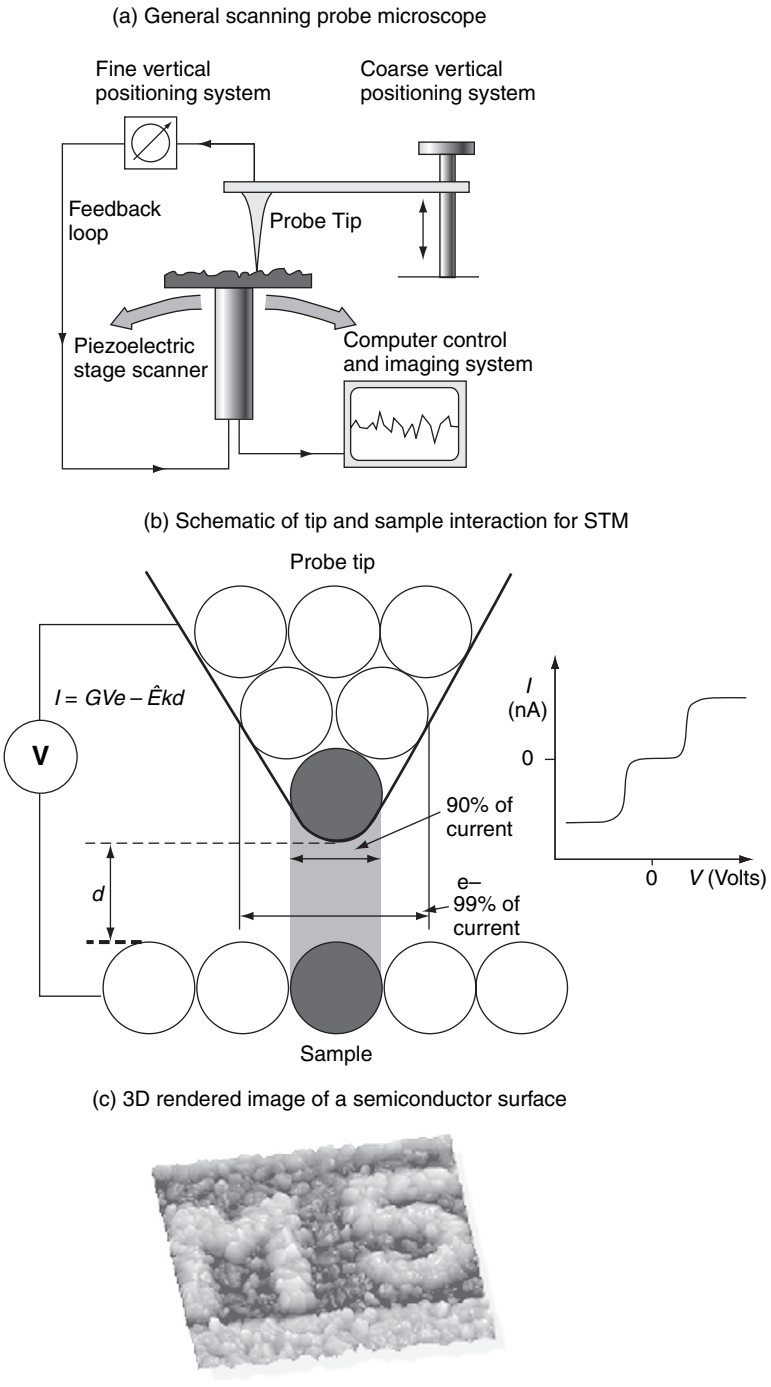

Figure 2.13 Schematic diagram of (a) a general scanning probe microscope, (b) the tip–sample interaction and *I–V* curve for STM, and (c) a typical STM image. Image (c) courtesy of Thermomicroscopes

An STM can be designed to scan a sample in one of two different modes: constant height or constant current. In constant-height mode, the tip travels in a horizontal plane above the sample and the tunnelling current varies as a function of the surface topography and the local surface electronic states of the sample. The tunnelling current, measured at each location on the sample surface, constitutes the data set, the topographic image. In constant-current mode, the STM uses a feedback system to keep the tunnelling current constant by adjusting the height of the scanner at each measurement point. For example, when the system detects an increase in tunnelling current, it adjusts the voltage applied to the piezoelectric scanner to increase the distance between the tip and the sample. In constant-current mode, the motion of the scanner therefore constitutes the data set. If the system keeps the tunnelling current constant to within a few percent, the tip-to-sample distance will be constant to within < 0.01 nm. Comparing the two modes, constant-height mode is faster because the system doesn't have to move the scanner up and down, but it provides useful information only for relatively smooth surfaces. Constant-current mode can measure irregular surfaces with high precision, although the measurement takes considerably more time.

To a first approximation, an image of the tunnelling current provides a map of the topography of the sample. However, since the tunnelling current corresponds to the electronic density of states at the surface, it actually measures a surface of constant tunnelling probability. In fact, STMs can directly measure the number of filled or unfilled electron states near the Fermi level at each particular (x, y) point on the surface, within an energy range determined by the bias voltage. This can cause problems if the sample surface is oxidized (which will dramatically lower the tunnelling current) or contaminated, and for atomic resolution studies the sample is often cooled and examined under UHV conditions.

The ability of the STM to probe the local electronic structure of a surface, in principle with atomic resolution, leads to the various techniques of scanning tunnelling spectroscopy (STS). STS techniques may involve recording 'topographic' (i.e., constant-current) images of the surface using different bias voltages and then directly comparing them, or taking current (i.e., constant-height) images at different heights z. Finally, true STS involves ramping the bias voltage with the tip positioned over a particular surface feature of interest, while recording the variation in tunnelling current; this produces a current versus voltage (I–V) curve characteristic of the surface electronic structure at a specific (x, y) location on the sample surface. STMs can be set up to collect I–V curves at every point in a data set, providing a three-dimensional map of electronic structure. With a lock-in amplifier, dI/dV (conductivity) or dI/dz (work function) versus V curves can be collected directly. STS methods are extremely powerful as they allow the surface electronic properties of a material to be investigated with near atomic resolution. In comparison, many other surface analytical spectroscopies average data from relatively large areas ranging from micrometres to millimetres.

2.5.2 Atomic force microscopy

Atomic force microscopy (AFM) was developed about five years after STM. While STM is limited to the surfaces of conducting specimens, AFM can be used to study non-conducting materials, such as insulators and semiconductors, as well as electrical conductors. As shown in Figure 2.14, AFM again uses a sharp tip, about 2 μm long and down to a minimum of 20 nm in diameter, which is scanned closely over the specimen surface,

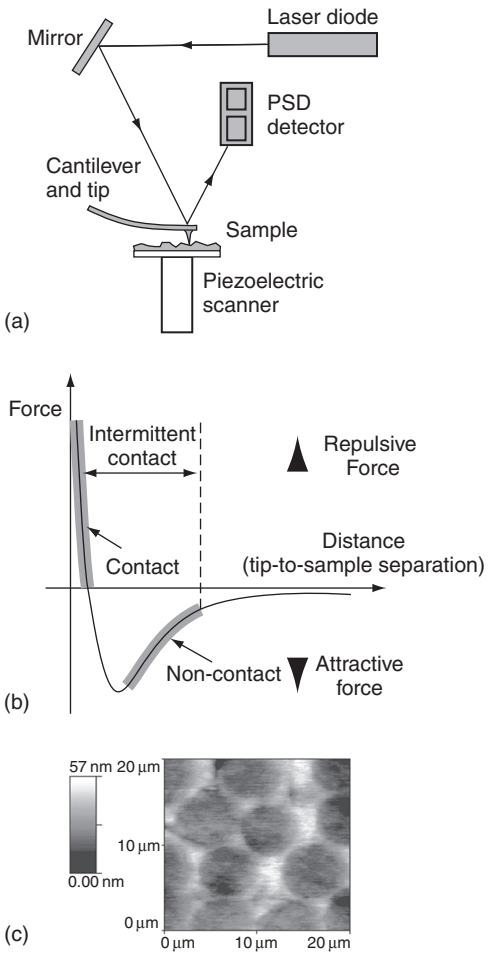

Figure 2.14 Schematic diagram of (a) an atomic force microscope, (b) the tip–sample interaction and (c) a typical AFM image of a carbon fibre–carbon composite. Image (c) courtesy of Professor Brian Rand, Institute for Materials Research, University of Leeds

but in AFM the magnitudes of atomic forces rather than tunnelling currents are monitored as a function of the probe position on the sample surface. AFMs can operate under UHV, ambient air conditions or even with the solid sample and tip submerged in a liquid cell, which is highly useful for biological systems (see Section 9.1.1).

The term 'atomic force microscopy' is used here in its loosest sense. Strictly, most images taken using an AFM are scanning force microscopy (SFM) images. Use of the phrase 'scanning force microscopy' is an acknowledgement that imaging with a scan length of up to 0.1 mm and pixel widths of nearly ~400 nm cannot be considered to be a probe of individual atomic forces. Indeed, the typical cantilever load of a few nanonewtons, if applied in contact mode through a single atom, would result in a pressure which exceeds the tensile strength of any known material by several orders of magnitude. In fact, contact mode tips have a large radii of curvature (about 50 nm) relative to

atomic dimensions. (The only circumstance in which single atom contact occurs is in the use of *non-contact* mode (see below), in which the tip oscillates above the sample, usually in an ultra-high vacuum environment, and the cantilever force is much lower.) In this text, it is perhaps unhelpful to use the term SFM rigorously, because the reader who dips into a particular chapter may be acquainted only with the term AFM, which enjoys much more widespread usage. After some soul-searching, we have therefore used 'AFM' in contexts where 'SFM' would be considered technically more appropriate.

The AFM tip is located at the free end of a cantilever that is 100 to 200 μm long and forces between the tip and the sample surface cause the cantilever to bend, or deflect. A detector is used to measure the cantilever deflection (usually just the displacement in the z direction although lateral force microscopy is also possible) as the tip is scanned over the sample, or alternatively, as the sample is scanned under the tip. The measured cantilever deflections allow a computer to generate a map of surface topography (see Figure 1.36).

The position of the cantilever is most commonly detected optically using a focused laser beam reflected from the back of the cantilever onto a position-sensitive (photo)-detector (PSD). Any bending of the cantilever results in a shift in the position of the focused laser spot on the detector, and the PSD can measure such displacements to an accuracy of < 1 nm. However, the ratio of the path length between the cantilever and the detector to the length of the cantilever itself, produces an additional mechanical amplification and, as a result, the overall system can detect sub-angstrom vertical movements of the cantilever tip. Other methods of detecting cantilever deflection rely on optical interference, or even an STM tip to read the cantilever deflection above a gold contact. If the cantilever is fabricated from a piezoresistive material, this obviates the need for a laser and PSD system as any deflection will result in strain, thereby changing the cantilever's resistivity, which can be monitored directly using an integrated electrical circuit.

Analogous to the case of STM, topographic AFM data sets are generated by operating in either constant-height or constant-force mode. In constant-height mode, the scanner height is fixed during the scan and the spatial variation of the cantilever deflection is recorded. Constant-height mode is employed for recording real-time images of changing surfaces, where high scan speeds are essential. It is also often used for taking atomic-scale images of atomically flat surfaces, where the cantilever deflections and thus variations in applied force are small. Conversely, in constant-force mode the deflection of the cantilever can be used as input to a feedback circuit that moves the scanner up and down in the z direction, responding to the topography by keeping the cantilever deflection (and thus also the total force applied to the sample) constant. Constant-force mode data sets are therefore generated from the scanner's motion in the z direction. In constant-force mode, the speed of scanning is limited by the response time of feedback circuit, but the total force exerted on the sample by the tip is well controlled. Constant-force mode is generally preferred for most applications.

Several forces typically contribute to the deflection of an AFM cantilever, the most common being relatively weak, attractive van der Waals forces as well as electrostatic repulsive forces. The general dependence of these forces upon the distance between the tip and the sample is shown schematically in Figure 2.14. When the tip is less than a few angstroms from the sample surface, the interatomic force between the cantilever and the sample is predominantly repulsive, owing to the overlap of electron clouds associated with atoms in the tip with those at the sample surface; this is known as the contact regime. In the non-contact regime, the tip is somewhere between ten to a hundred Angströms from

the sample surface, and here the interatomic force between the tip and sample is attractive, largely as a result of long-range attractive van der Waals interactions.

In contact-AFM mode, the AFM tip makes soft, physical contact with the sample. The tip is attached to the end of a cantilever, which has a lower spring constant than the effective spring constant holding the atoms of the sample together. Hence as the scanner gently traces the tip across the sample, the repulsive contact force (which is a very steep function of tip–sample separation) causes the cantilever to bend to accommodate changes in topography rather than forcing the tip atoms closer to the sample atoms. If the cantilever is very stiff (i.e., it has a larger spring constant) then the sample surface is likely to deform during contact and this can be used for nanopatterning of surfaces (Chapter 1).

In addition to the electrostatic repulsive forces described above, two other forces are generally present during contact-mode AFM operation. Firstly, an attractive capillary force (the magnitude of which is about 10^{-8} N depending on the tip–sample separation) exerted by the thin water layer often present between the tip and the sample in an ambient environment and, secondly, the force exerted by the cantilever itself. The force exerted by the cantilever can be either repulsive or attractive and depends on the deflection of the cantilever and upon its spring constant. Overall the total force that the tip exerts on the sample is the sum of the capillary plus cantilever forces (typically 10^{-8} to 10^{-6} N) and, in the case of contact-mode AFM, will be balanced by the repulsive electrostatic force.

In non-contact AFM mode, a stiff cantilever is vibrated near the surface of a sample close to its resonant frequency. The spacing between the tip and the sample is of the order of tens to hundreds of Angströms and this is similar to the vibration amplitude of the tip. As the tip approaches the surface, this is detected as a change in the resonant frequency or vibration amplitude of the tip oscillation, and this can provide sub-Angström vertical resolution in the topographic image. Generally the detection system monitors the resonant frequency or vibrational amplitude of the cantilever and keeps it constant with the aid of a feedback system that moves the scanner up and down, ensuring that the average tip-to-sample distance also remains constant.

Non-contact AFM mode provides a means of measuring sample topography with minimum contact between the tip and the sample, thus removing any possibility of contamination or degradation of the surface by the tip, as can occur during contact AFM scanning. Furthermore, the total force between the tip and the sample in the non-contact regime is very low, approximately 10^{-12} N, which is beneficial for studying soft or elastic samples. Generally both contact and non-contact topographic images of rigid samples will appear similar. However, the presence of a few monolayers of liquid (e.g., water) on the surface of a sample will result in non-contact AFM mode imaging the surface of the liquid layer, whereas operating in contact mode will penetrate the liquid layer to image the underlying surface.

A further AFM variant is intermittent-contact or 'tapping' AFM mode. This is similar to non-contact mode except the vibrating cantilever tip is brought closer to the sample so that at the bottom of its travel it just touches the sample. Tapping mode is less likely to damage the sample than contact AFM because it eliminates lateral forces (friction or drag) between the tip and the sample. Furthermore, in comparison with non-contact mode, tapping mode can image larger scan sizes that can include greater variations in sample topography.

AFM may also be used to measure the local elastic properties of a point on the surface or even a whole surface. Here the force on an AFM cantilever tip in the z direction is monitored as a function of the z position of the piezoelectric scanner tube.

This force versus distance curve will vary with changes in local elastic properties of the surface, or the presence of contaminants and lubricants.

Generally AFM can only be used to provide information about the surface of a sample; any subsurface information about the morphology and structure must be obtained by scattering techniques (such as SEM or reflectometry) or sputter depth profiling (dynamic SIMS). It is often possible to reveal structural information of mixtures by selectively dissolving one component with an appropriate solvent. AFM is performed on the surface of, say, a film of a two-component mixture before exposing that film to a solvent which dissolves only one component of the mixture. The change in morphology will tell us where that component resided in the film. However, on certain samples it is possible to reveal three-dimensional structural information with nanometre resolution. This technique, termed nanotomography by its inventor, involves imaging the surface of a film before etching only a small amount in a plasma. The same part of the film is imaged again using AFM before another etching step is performed. The images can be stacked together to produce the final three-dimensional volume image. Image reconstruction is complicated by the difficulty in obtaining the same scan location *exactly* each time, and also by the fact that different components etch at different rates, hence the need for a careful alignment of images (registration) similar to that used in magnetic resonance imaging. Figure 2.15 shows an example of a nanotomography scan of a block copolymer.

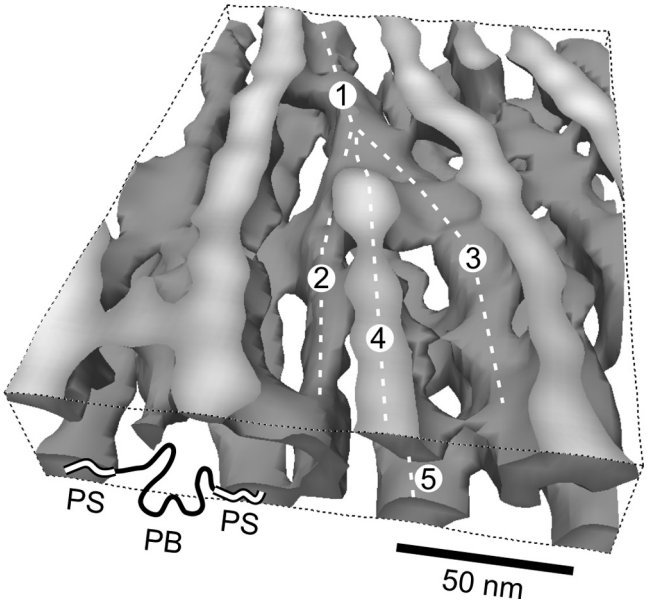

Figure 2.15 Nanotomography image of a triblock copolymer of polystyrene and polybutadiene (PS–PB–PS). The polystyrene component is shown in this image (clear regions indicate the location of polybutadiene). This method of imaging a material reveals information that would probably be impossible using other techniques; for example, notice how one of the polystyrene 'cylinders' (marked 1 in the figure) splits into four arms (marked 2, 3, 4 and 5). Reproduced with permission from R. Magerle, Nanotomography, *Phys. Rev. Lett.* **85**, 2749–2752 (2000). Copyright 2000 by the American Physical Society

2.5.3 Other scanning probe techniques

There are an increasing number of variations on the standard STM and AFM techniques that can provide additional information to surface topography. However, these may well be at a lower spatial resolution depending on the particular interaction involved. Force modulation microscopy works in contact mode and involves applying a periodic oscillation to the tip or the sample and monitoring the amplitude and sometimes phase lag of the cantilever oscillation; it provides information on the elastic properties of the sample. Lateral force microscopy (LFM) detects lateral deflections such as twisting of the cantilever that arise from forces on the cantilever parallel to the plane of the sample surface. LFM studies are useful for imaging variations in surface friction that can arise from surface inhomogeneities.

Magnetic force microscopy (MFM) is used to image the spatial variation of magnetic forces on a sample surface (such as domain structures in magnetic materials) and employs a tip coated with a ferromagnetic thin film. The system operates in non-contact mode, detecting changes in the resonant frequency of the cantilever induced by the dependence of the long-range magnetic field on tip-to-sample separation. The image also contains information on sample topography, particularly when the tip is close to the sample surface, however this can be separated from magnetic effects by recording images at different tip heights. Further details are given in Chapter 4.

In a similar fashion, local charge domains on a sample surface (such as a microprocessor) may be imaged using electrostatic force microscopy (EFM) where a voltage is applied between the tip and the sample whilst the cantilever vibrates in non-contact mode. Scanning capacitance microscopy (SCM) images spatial variations in capacitance, again with a voltage applied to the tip and a special circuit to monitor the capacitance between the tip and sample. SCM can be used to map the dielectric properties of a surface, which may be influenced by layer thickness or the presence of subsurface charge carriers.

Scanning thermal microscopy (SThM) measures the thermal conductivity of the sample surface using a resistively heated Wollaston wire as the probe. Changes in the current required to keep the scanned probe at constant temperature produce a thermal conductivity map of the surface (or thermal diffusivity as the temperature is oscillated). Conversely, changes in the probe resistance as a constant current is applied can generate local surface temperature maps. In addition, with the probe at a particular point on the surface, a localized microthermal analysis can be performed monitoring calorimetric or mechanical changes as a function of temperature (Section 2.9.4).

2.6 DIFFRACTION TECHNIQUES

2.6.1 Bulk diffraction techniques

As we have seen, diffraction techniques provide an insight into the crystallography of a material and, in some cases, the structure of a sample system or device. We have already mentioned electron diffraction in relation to TEM of thin specimens, and here we discuss bulk diffraction techniques which refer principally to X-ray diffraction (XRD) and neutron diffraction (ND) in which relatively large volumes of material, greater than

about 0.1 mm^3, need to be sampled. This need arises because, unlike electrons, X-rays and neutrons cannot easily be focused. They can usually only be collimated and the practical limit of collimation (in order to provide diffracted beams of sufficient intensity) is of the order of 0.5 mm diameter for X-rays and 10 mm diameter for neutrons. In practice, much wider beams than this are employed.

X-ray diffraction is of course of paramount importance in determining the structures of crystals. From the pioneering work of the Braggs in 1913–15, who solved the first crystal structures, to the determination of the structure of DNA by Watson, Crick, Franklin and Wilkins in 1953, XRD has unlocked the door to the transmission of the gene and the evolution of life itself. The application of XRD to nanocrystalline solids, powders, single-crystal thin films or multilayers may be less spectacular, but apart from the standard procedures for the determination of their crystalline structures (from an analysis of the directions and intensities of diffracted beams), XRD can provide information on nanocrystal size (strictly, coherence length) and microstresses and microstrains (from an analysis of line broadening), macrostrains (from an analysis of line shifts), repeat distances (or superlattice wavelengths) and total thicknesses in multilayer films (from an analysis of low-angle Bragg peaks) and orientation distributions or 'texture' (from an analysis of pole figures or orientation distribution functions).

In all these applications monochromatic X-rays are employed, either by the use of crystal monochromators or filters which pass only the strong characteristic $K\alpha_1$ components of the whole spectrum from an X-ray tube, or by crystal monochromatized radiation from a synchrotron source (and which has the advantage that the wavelength can be tuned to minimize unwanted fluorescent excitation in the sample). The original Laue technique, which used the whole white or bremsstrahlung radiation (plus a whole range of characteristic wavelengths in the spectrum) is inapplicable to polycrystalline solids because the two variables λ and θ (arising from many orientations) would, by simple inspection of Bragg's law (2.1), give rise in effect to very many superimposed diffraction patterns which would be quite impossible to sort out.

The X-ray technique of greatest importance is therefore the powder diffractometer in which the monochromatic beam is incident at an angle θ on a specimen of about 10 mm^2 in area and (for polycrystalline or powder specimens) a minimum thickness of about 20 μm, mounted on a support film that does not give rise to interfering reflections. The detector is set to receive reflections at an angle θ (the Bragg–Brentano symmetrical arrangement shown in Figure 2.16(a)) and this is varied over the angular range of interest (typically 1–6° for low-angle reflections and 6–80° for high-angle reflections), either by keeping the incident beam direction fixed and rotating the specimen and detector (the detector at twice the angular velocity) or by keeping the specimen fixed and rotating the incident beam and detector in opposite senses. In both cases this instrumental set-up preserves the symmetrical arrangement.

Reflected beams, it should be noted, are only recorded from those planes which lie parallel, or nearly parallel, to the specimen surface. In order to record reflections from planes which lie at an angle to the specimen surface, as required for texture or macrostrain determination, the specimen must be tilted or rotated to the required angle(s); i.e., away from the symmetrical Bragg–Brentano arrangement. The texture goniometer is essentially a specimen holder arrangement in which such tilts or rotations are carried out automatically: the incident beam and detector directions are fixed to receive a specific reflection (d spacing) of interest and the variations in intensity are recorded.

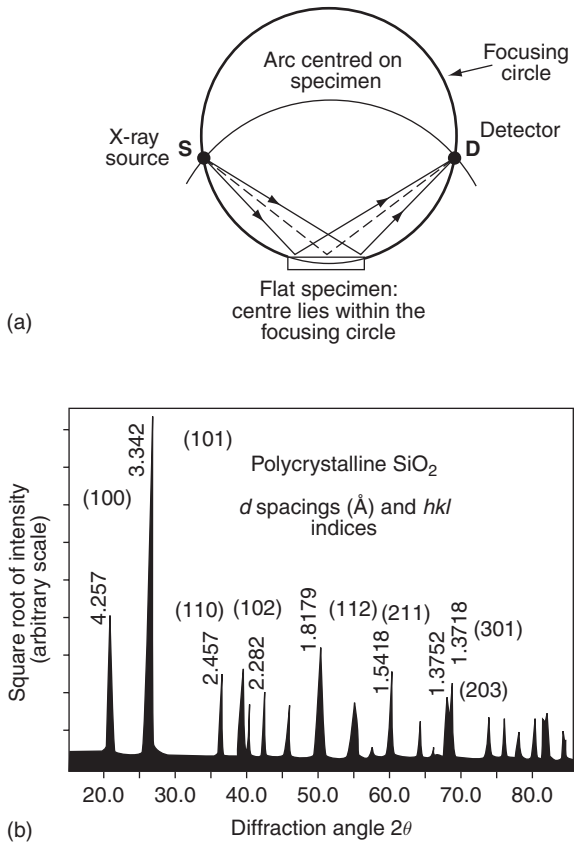

Figure 2.16 Schematic diagram of (a) the geometry in an XRD powder diffractometer and (b) an example of an XRD powder pattern from crystalline silicon oxide

In the diffractometer, the angles (d spacings) and intensities of the high-angle reflected beams serve as a 'fingerprint' for the crystal structure. This may be identified by comparing the XRD pattern (e.g., Figure 2.16(b)), with a database such as the Powder Diffraction File (PDF), which lists some 140 000 inorganic structures. The identification is not always easy, especially when the specimen contains two or more phases, or when the reflections are broadened or shifted as a result of small grain size or strain. The broadening β due to crystallite grain size t arises from the limited number of diffracting planes within a diffracting object and is given by the Scherrer equation

$$\beta = \frac{K\lambda}{t\cos\theta}, \tag{2.6}$$

where K is a shape factor ($\approx$1). This is analogous to the broadening of points in the reciprocal lattice that is encountered in electron diffraction of thin TEM samples. The

peak broadening due to the presence of internal elastic microstrains (varying between different crystals in the samples) is given by

$$\beta = \frac{-2\varepsilon}{\cot\theta}, \tag{2.7}$$

where ε is the elastic strain. The deconvolution of these contributions, as well as the instrumental broadening arising from such factors as detector slit width, is also not a simple task and a common procedure is to compare the broadening from, for example, large grain size and nanograin size specimens.

The low-angle reflections give information on the thickness of thin films and multilayers, as well as the mean size and size distributions of scattering objects such as micelles, colloidal particles and macromolecules. Since the scattered intensity is low, this is most commonly performed using high-intensity synchrotron radiation. The reflections, or rather diffuse diffraction maxima, from amorphous or semicrystalline materials such as glasses can provide information on short-range structural coherency and average spacings between atoms or ions.

The diffractometer is not appropriate or useful when only small quantities of material are available. In such situations, for high-angle work, the Debye–Scherrer powder camera and Seeman–Bohlin camera still retain their usefulness, particularly with the advent of improved area X-ray detectors. For very low angle work, the diffraction pattern (or rather radial diffracted intensity distribution) may be recorded from the beams transmitted through a small thickness or volume of material, the incident beam being very finely collimated using carefully aligned knife-edges; this is the basis for the small-angle X-ray scattering (SAXS) technique.

As mentioned in Section 2.1.2.5, neutrons are largely scattered by the nuclei, rather than the atomic electrons, and the scattering amplitudes vary in an irregular way with atomic number. The large scattering amplitudes of, for example, hydrogen and oxygen atoms, in comparison with heavy metal atoms, allow hydrogen and oxygen atoms to be located within the structure. Moreover, since the nuclear cross sections are very small, the interference effects, which in the case of X-rays lead to a decrease in scattering amplitude with angle, are also small and neutron scattering amplitudes do not decrease rapidly with angle. Finally, since neutrons possess spin, they interact additionally with the unpaired electrons in magnetic materials and hence neutron diffraction finds applications in the study of ferro-, antiferro- and ferrimagnetic materials.

Neutrons emerge from a high-flux nuclear reactor with a range of velocities corresponding (via de Broglie's equation) to wavelengths which peak typically in the range 0.1–0.2 nm; i.e., close to the characteristic Kα X-ray wavelengths from metal anodes of X-ray tubes. Single-wavelength beams are achieved, as with X-rays, through the use of crystal monochromators, hence the diffraction or reflection angles are similar to those for X-rays. Neutron diffraction is geometrically similar to X-ray diffraction except that the physical size of the apparatus needs to be much larger because of the lower intensities of neutron beams. In summary, therefore, neutron diffraction is very much a bulk technique, and apart from structural or magnetic changes which may occur in large assemblies of nanoparticles or nanocrystalline solids, it currently offers limited additional benefits for the analysis of nanostructures. Reflectometry experiments using neutrons are discussed in Section 2.8.4.

2.6.2 Surface diffraction techniques

In addition to grazing incidence (angle) techniques using X-rays and to some extent neutrons, which limits the penetration depth of the incident radiation, there are two important techniques for the structural analysis of surfaces, both of which employ electrons. These are reflection high-energy electron diffraction (RHEED) and low-energy electron diffraction (LEED), shown in Figure 2.17. Both methods can be used to determine the periodic, two-dimensional arrangement of atoms at surfaces. Since electrons are scattered much more strongly by matter than either X-rays or neutrons

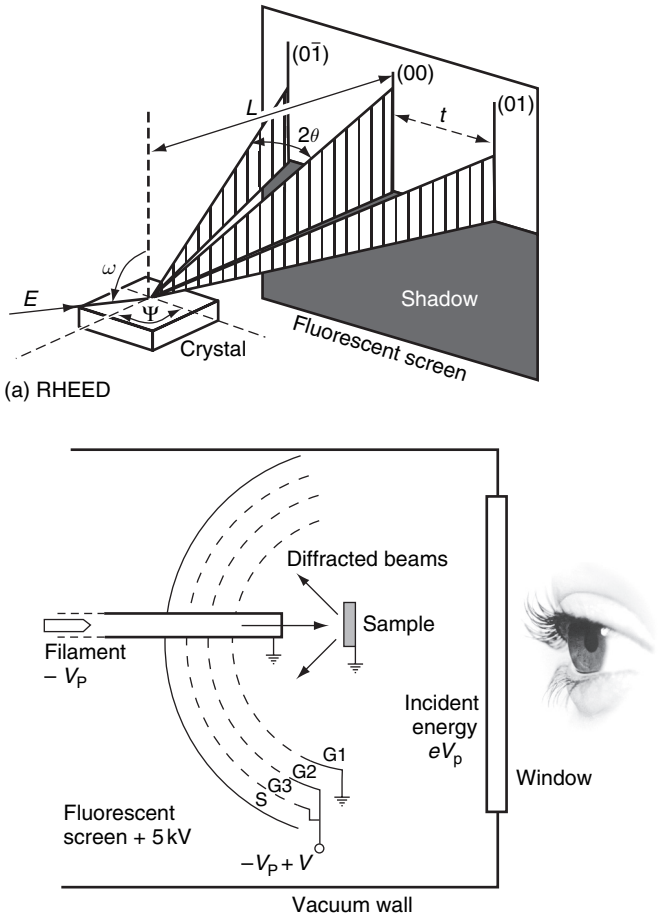

Figure 2.17 Schematic diagram of the experimental set-up for (a) RHEED and (b) LEED. In (a) the electron beam is of energy E and the direction relative to the sample is defined by the two angles ω and Ψ. The camera length L defines the observed spacing t of the diffraction streaks separated by twice the Bragg angle θ. In (b) the filament is held at a potential $-V_P$ and the diffracted beams are accelerated between the grid G_1 and the screen S. Grids G_2 and G_3 are used to reject inelastically scattered electrons

(Section 2.1.2.5) and penetrate less deeply, they constitute a very important probe of surface structure as, in general, the surface-specific signal can often be extremely weak relative to the bulk. To prevent surface contamination both techniques require UHV conditions and also methods for cleaning surfaces in situ, either by heating, low-energy ion sputtering or cleaving the sample in situ.

In RHEED a high-energy (typically 100 keV) beam of electrons is incident on a plane surface at a grazing angle, thus preventing excessive electron penetration. This experimental arrangement results in a set of diffraction streaks normal to the shadowed edge of the sample the form of which is dependent on Bragg's law and the electron beam direction. RHEED can determine the symmetry and dimension of the surface unit mesh (the repeat unit) by varying the incident beam direction. The grazing incidence geometry is very convenient for observing the change in surface structure whilst a material is being deposited on a surface. During monolayer deposition, such as occurs during MBE (Section 1.4.2.3), the intensities of the RHEED beams oscillate, which allows this technique to be used for monitoring atomic layer epitaxial growth.

In LEED a low-energy beam of electrons (between 10 and 1000 eV) is incident normal to a surface and diffracted. The elastically backscattered electron beams are accelerated towards a fluorescent screen or camera and the inelastically scattered electrons are rejected by grids held at a slightly positive potential. A single crystal surface will produce a spot pattern which can directly give information about the surface symmetry and the surface unit mesh. In principle LEED (and also RHEED) intensities can be modelled to obtain information on where atom types are located within the surface unit mesh.

2.7 SPECTROSCOPY TECHNIQUES

Generally in all types of spectroscopic measurements (e.g., optical, UV, IR, X-ray and electron spectroscopies) we are concerned with three types of processes: absorption (transmission), reflection or emission (luminescence). The different experimental methodologies are shown in Figure 2.18. Such experiments probe the energy differences between electronic, vibrational and, in some cases, rotational quantum energy states of atoms as well as the lifetime of excited states and their energy relaxation channels. The latter studies are often aided by using a pulsed radiation source.

Besides the spatial resolution of the radiation incident upon and propagating within the sample (which determines the volume of analysis), a further important experimental parameter is the wavelength or energy resolution of both the incident radiation and the detected radiation, which determines the overall spectral resolution. The intensity of the reflected, transmitted (absorbed) or emitted radiation may be recorded as a function of the wavelength of the incident radiation; alternatively, the wavelength of the incident radiation is fixed and the distribution in intensity of the reflected, transmitted or emitted signals are spectroscopically analysed as a function of wavelength.

A further experimental variable in such spectroscopic experiments is the polarization of the incident and/or reflected, transmitted or emitted radiation. The absorption of electromagnetic radiation involves the resonant response of electrons in an atom or molecule to the oscillating electric field of the radiation, which is a vector quantity and is orthogonal to the propagation direction and the oscillating magnetic field vector of the

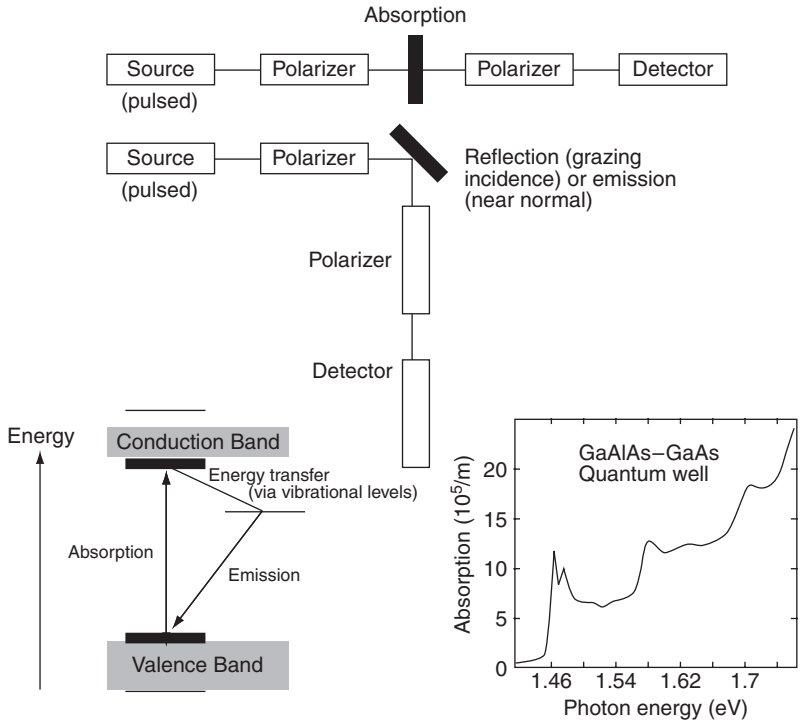

Figure 2.18 Schematic diagram of absorption, reflection and emission spectroscopic experiments together with an energy level diagram and a typical absorption spectrum from a quantum well

radiation. If the radiation is linearly polarized, we restrict the electric field vector of the radiation to one plane (plane polarized radiation) and this can be used to probe the directional properties of, in this case, the electronic transitions of atoms within the sample and hence the structure of the sample with respect to the plane of polarization. There are other forms of polarization, known as circular and elliptical polarization, which possesses a polarization vector that describes either a left-or right-handed helical path as a function of time or position and can be of use for studying optically active materials or magnetic transitions. Briefly, circular polarization involves the electric and magnetic fields of the polarization each being described by two equal orthogonal components which are out of phase (by 90°), while elliptical polarization involves the electric and magnetic fields of the polarization each being described by two unequal orthogonal components which are out of phase. In addition to electromagnetic radiation, it is also sometimes possible to perform similar experiments with electrons.

2.7.1 Photon spectroscopy

Generally the lateral spatial resolution of photon-based spectroscopies is limited by diffraction, and can attain probe sizes of at best a few micrometres, unless near-field techniques are used (Section 2.2.4.1), whereas X-ray synchrotron sources have demonstrated

spot sizes of 50–100 nm by focusing using Fresnel plates. Lateral resolution is therefore limited and essentially the techniques provide an ensemble-averaged signal from a nanostructured sample. Furthermore, most photon-based spectroscopies probe bulk rather than surface properties, however there do exist various techniques for localizing the signals from surfaces (e.g., grazing incidence studies) which are of use for analysing thin film nanostructures. Photon-based spectroscopies have been revolutionized in the X-ray regime by the use of synchrotron radiation and, in the optical/UV regime, by the use of lasers that provide coherent, intense and often tunable radiation (over a large range of wavelengths). Furthermore, lasers can often be pulsed over sub-picosecond timescales.

Despite the limited spatial resolution, extremely important information can be derived from photon-based spectroscopies. As discussed in Chapter 1, all electronic energies in nanostructures are affected by quantum confinement, in addition to the effects of the increased contribution of surfaces. These changes are most apparent for transitions between the highest occupied molecular orbitals (HOMO) in the valence band and the lowest unoccupied molecular orbitals (LUMO) in the conduction band of a solid, both of which are directly influenced by changes in the bonding. Such transitions often lie in the optical region of the electromagnetic spectrum, where λ is typically 400–700 nm (sometimes expressed as a wavenumber $(1/\lambda)$ in cm^{-1}). Photon-induced transitions follow a set of dipole selection rules for the quantum numbers of the initial and final electronic states of the transition, which causes only selected transitions to be observed.

Colloidal metallic nanoparticles have long been known to show colours that depend on their size and shape. Semiconductor nanoparticles also show a distinct quantum size effect, whereby the absorption edge shifts from the infrared to the visible region as the particle size decreases and the band gap correspondingly increases. Generally semiconductor nanoparticles are luminescent, however the surface properties of the particles can change the emission characteristics significantly via the introduction of defect states. Quantum well structures often exhibit distinct steps in absorption spectra, rather than a continuous increase as observed in the bulk, corresponding to the quantized energy levels associated with the localized well. This is shown schematically in Figure 2.18. There are further effects due to the fact that the electron–hole pair (an exciton) created during the energy absorption process may be constrained within the well, leading to an increase in the exciton binding energy and the appearance of sharp exciton peaks at the edges of the steps in the absorption spectrum.

Further effects due to nanodimensionality can be seen in the longer-wavelength, infra-red and far-infra-red ($1-20\,\mu\mathrm{m}$) portion of the electromagnetic spectrum which are associated with changes in vibrational (and hence thermal) properties arising from changes in structure and/or composition.

2.7.1.1 Optical measurements

The two simplest optical measurements are reflectivity and transmission/absorption. Both record either the reflectance or the transmittance/absorbance as a function of wavelength. For transmittance measurements in the wavelength range 185–3000 nm (UV, optical, IR region) there are a large number of commercially available dual-beam spectrophotometers.

Simplistically, absorption techniques involve a measurement of the attenuation of the intensity of a monochromatic light beam as it traverses the sample. At a given

wavelength or frequency, the relative absorption intensity I/I_0 generally follows the Beer–Lambert law:

$$I/I_0 = \exp(-Kx), \tag{2.8}$$

where K is the absorption or extinction coefficient and x is the path length through the sample. There are a large number of variations on this simple experiment, for instance using polarized light, external applied (electric or magnetic) fields or an external applied stress. The light source can be a tungsten lamp, a gas discharge lamp, a highly monochromatic continuous laser or even a pulsed laser. Some experiments probe the change in absorption following the absorption of a laser pulse. The absorption can be measured by monitoring the intensity of the transmitted beam directly (i.e., using the photoelectric effect to measure a current) or by measuring the absorbed energy in the form of heat deposited in the sample. Samples for optical spectroscopic measurements are often cooled to low temperatures (e.g., liquid helium at 4.2 K), essentially to depopulate vibrational energy levels and ease the interpretation of the spectra in terms solely of excited electronic states.

Modulation spectroscopy involves the measurement of small changes in optical properties (typically absorption or reflectivity) of a material that are caused by a modification of the measurement conditions. The modification is generally an external perturbation such as the application of an electric field, a light pulse, a magnetic field, a heat pulse or uniaxial stress. In some cases this can give rise to transitions to states that are not evident in normal absorption spectra. Two techniques of use for studying nanoparticles embedded in a matrix of another material are electro-absorption and photo-absorption. Here the wavelength is tuned to a region of the spectrum where the host matrix does not absorb radiation. In electro-absorption the absorptivity is altered by application of an external AC field, the so-called Stark effect. This can provide details on the nature of the excited state in nanocrystals depending on how the absorption is modified by the polarization of the incident radiation. Photo-absorption involves the measurement of the change in transmission resulting from a periodic light pulse that can cause filling and also splitting of electronic levels which will modify the absorption spectra. This can reveal the presence of non-linear optical effects, where the properties of the absorbing or reflecting species or particle are altered by the incident radiation.

Generally the optical properties of materials are expressed using a complex quantity known as the dielectric function, ε, which may be separated into real and imaginary parts, $\varepsilon = \varepsilon_1 + i\varepsilon_2$. The dielectric function represents the response of electrons in the solid to the electrical and magnetic field of the incident radiation. Reflectance and transmittance measurements allow determination of the index of refraction, n, and the extinction/absorption coefficient, K, of the sample. Under certain conditions these quantities can be used to derive the behaviour of the real and imaginary parts of the dielectric function as a function of frequency or wavelength. This allows identification of the contribution of resonant plasmon-type oscillations as well as interband transitions in the absorption or reflection spectrum. Measurement of the optical transmittance, and hence the absorption coefficient, allows a direct determination of the imaginary part of the dielectric function. When plane polarized light is reflected from a solid surface the reflected wave is, in general, found to be elliptically polarized.

Ellipsometry is a useful technique when one is interested in the real part of the dielectric function and involves the measurement of the ellipticity (the amplitudes of the two components of the oscillating electric field and their phase difference) as a function of incidence angle, the plane of polarization of the incident light and its wavelength.

As highlighted in Chapter 1, metallic nanoparticle systems show an absorption behaviour that is dependent on the size and shape of the particles (Figure 1.17). The original explanation was provided by Mie, who solved Maxwell's equations for the case of spherical particles. In the visible region this usually corresponds to excitation of surface plasmons by the incident electromagnetic radiation. Surface plasmons are collective oscillations of valence electrons associated with surface atoms, and while they are insignificant in the spectra obtained from bulk materials, they are an important phenomenon in nanometre-sized particles which have a large surface area to volume ratio. The technique forms the basis for the technique of surface plasmon resonance spectroscopy. If the particles are not spherical but rod-like, for example, then the surface plasmon absorption will split into tranverse and longitudinal modes corresponding to the electron oscillations perpendicular and parallel to the major axis of the rod.

2.7.1.2 Photoluminescence

Emission or luminescence studies are complementary to optical and IR absorption techniques in that they also provide details of the spectral distribution of electron energies and their polarization behaviour, in particular low energy states which may be obscured due to vibrational transitions. A comparison between the luminescence peak and the lowest absorption peak can provide information on the nature of the emitting state.

Photoluminescence (PL) measurements are made by exciting a sample with a laser tuned to a particular wavelength, collecting the scattered light from the front surface and dispersing it as a function of wavelength then measuring its intensity. The polarization dependence of the emitted intensity can be studied by inserting a linear polarizer in the detection pathway. A variation on this technique is photoluminescence excitation (PLE) spectroscopy, which monitors the intensity of a single luminescence band while the excitation wavelength is varied. Since the size of a nanoparticle affects the absorption wavelength, then PLE can be used to study size distributions. Imaging luminescence from individual nanoparticles has also been achieved via dilution of nanoparticle solutions to levels of 1nanocrystal/μm^2 and coating as thin films. The fluorescence is then filtered and imaged on a two-dimensional CCD camera. Spectroscopy from single nanocrystals can also be achieved in a similar fashion.

Extremely important information comes from the time dependence and efficiency of luminescence. Once an excited state has been created, it can either emit or decay non-radiatively into thermal energy; alternatively the excitation may migrate to a different site within the lattice, which may itself either emit or decay non-radiatively. The migration process is known as energy transfer and can reveal how the two sites are coupled. An impurity, defect or surface site may act as a trap for the mobile excitation. The use of pulsed laser sources and gated detection in time-resolved PL can provide details of the energies and lifetimes of these trap states. For example, nanoparticles with different surface modifications (e.g., surfactants and stabilizers) will show different time-resolved PL spectra.

2.7.1.3 Infrared and Raman vibrational spectroscopy

The energy of most molecular vibrations corresponds to the infrared region of the electromagnetic spectrum and these vibrations may be detected and measured in an infrared (IR) or Raman spectrum, which typically covers a wavelength range of 2–16 μm (more commonly expressed as 400–5000 cm^{-1}).

For IR the spectrometer system consists of an IR source emitting throughout the whole frequency range; the beam is split, with one beam passing either through the sample (transmission) or alternatively reflected (for studies of surfaces) while the other beam is employed as a reference beam. Samples may be a vapour or solution contained in a special NaCl cell, a liquid between NaCl plates, or a solid ground together with excess KBr and pressed into a disc. Fourier transform IR (FTIR) spectroscopy works by allowing the transmitted (or reflected) beam to recombine with the reference beam after a path difference has been introduced. The interference pattern so produced is Fourier transformed to produce the spectrum of material being analysed as a function of wavelength or wavenumber of the incident radiation.

If a molecular vibration is excited in the sample, the molecule absorbs energy of the particular frequency and this is detected as absorption relative to the reference beam. As shown in Figure 2.19, different functional groups have characteristic vibration frequencies arising due to stretching, bending, rocking and twisting of bonds which allows the particular functional group to be identified. These may change slightly when the functional group is incorporated into a solid. However, not all possible vibrations are excited in IR spectroscopy, as a dipole moment must be created during the vibration. In molecules possessing a centre of symmetry, vibrations symmetrical about the centre of symmetry are IR-inactive.

The attenuated total reflection spectrometer (ATR-FTIR) is an extension of the FTIR technique and consists of an infrared-transparent block; silver chloride, zinc selenide and germanium are suitable materials. An infrared beam passes through this block at a glancing angle and is totally reflected several times at the top and bottom before exiting at the other end. The sample can be adhered to the block, changing the IR spectrum. In total internal reflection there is no power transmission outside of the block but this does not imply that the IR wave is completely confined within the block; in fact, the beam penetrates a very small distance into the contacting sample layer (this is the evanescent wave) and this can be utilized for infrared absorption. In the IR case, the evanescent wave penetrates about a micron into any sample clamped onto the sample cell. If this sample has strong IR absorption lines, these will be revealed in the final spectrum. In fact, ATR-FTIR is not restricted to solid materials attached to the ATR block; the block can also be immersed in solution to study adsorption to the block. This is a particularly useful way of studying nanoscale conformational events such as protein folding and unfolding, which might take place at surfaces.

A complementary vibrational spectroscopic technique is Raman spectroscopy; this usually employs a laser source and here the scattered light is analysed in terms of its wavelength, intensity and polarization. The parent line is simply formed by absorption and re-emission of the light, whereas other weaker lines involve not only absorption/emission but also vibrational excitation or de-excitation (Raman lines). The vibrational frequency is just the difference between the parent line and the Raman line. Generally IR-inactive vibrations are Raman active, which makes the techniques complementary for analysis of functional groups or bonds (chemistry). Additionally, low-frequency

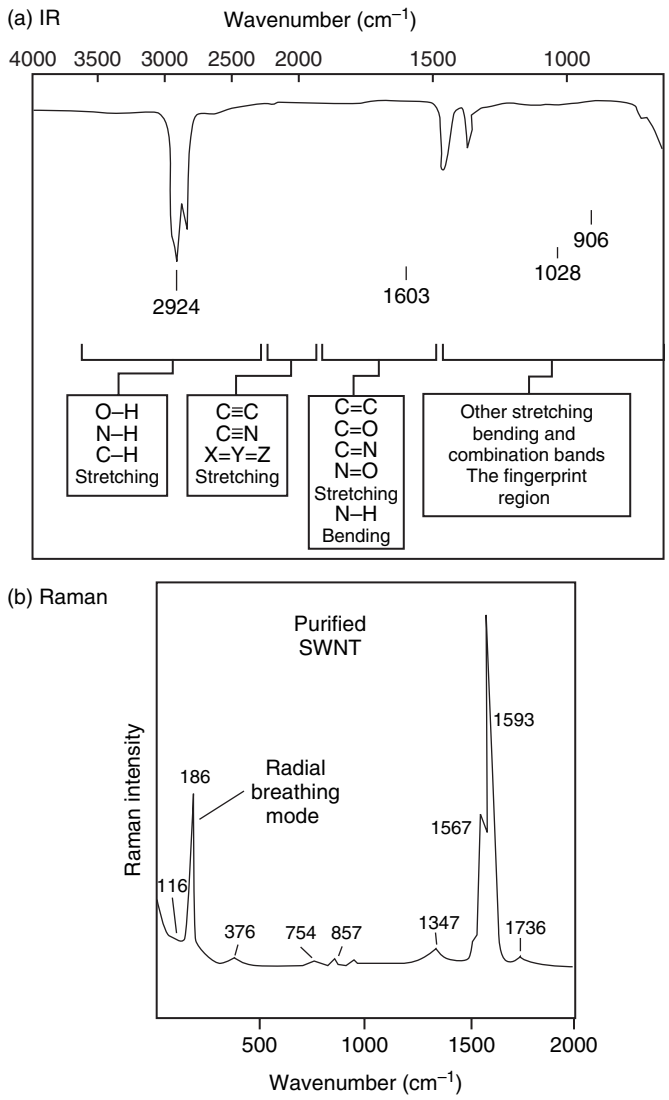

Figure 2.19 (a) Diagram of the IR spectral regions for different chemical bonds in organic materials. (b) A schematic example of a Raman spectrum from single-walled carbon nanotubes; the radial breathing mode at 186 cm^{-1} is characteristic of the entire nanostructure

Raman scattering can probe elastic vibrations of an entire nanoparticle, which most often shows a distinct size dependence; it is used extensively to analyse single- and multi-walled carbon nanotubes.

Both IR and Raman microscopy are possible using a specially designed confocal scanning laser microscope (CSLM) that provides a laser spot size, hence lateral spatial resolution, of around 1 μm. The confocal nature of the microscope allows different depths of up to 2 μm below the surface to be sampled. Typically the laser produces polarized red light, the scattered light is filtered to remove the parent line and the remaining Raman lines are analysed with a monochromator.

2.7.1.4 X-ray spectroscopy

X-rays from a high-intensity source with a controllable wavelength, such as a synchrotron, can also be used to perform absorption or reflectivity studies. X-ray absorption spectroscopy (XAS) involves the ionization of atoms by high-energy photons causing electron transitions to occur from deep inner shell atomic levels to the empty conduction band. X-ray ionization edges occur at energies characteristic of both the element and the inner shell involved and show two types of fine structure: X-ray absorption near-edge structure (XANES), which involves strong intensity oscillations within about 40 eV of the absorption edge onset, and weaker oscillations at higher energies known as extended X-ray absorption near-edge structure (EXAFS). XANES is directly related to the unoccupied density of electronic states and has characteristic features and energy positions for different bonding environments of the atoms in the solid, whereas EXAFS oscillations can be used to determine bond lengths and coordination numbers in terms of a radial distribution function around the ionized atom. This is shown schematically in Figure 2.20. Both these techniques have electron equivalents in the technique of EELS in the TEM (Section 2.7.3.3).

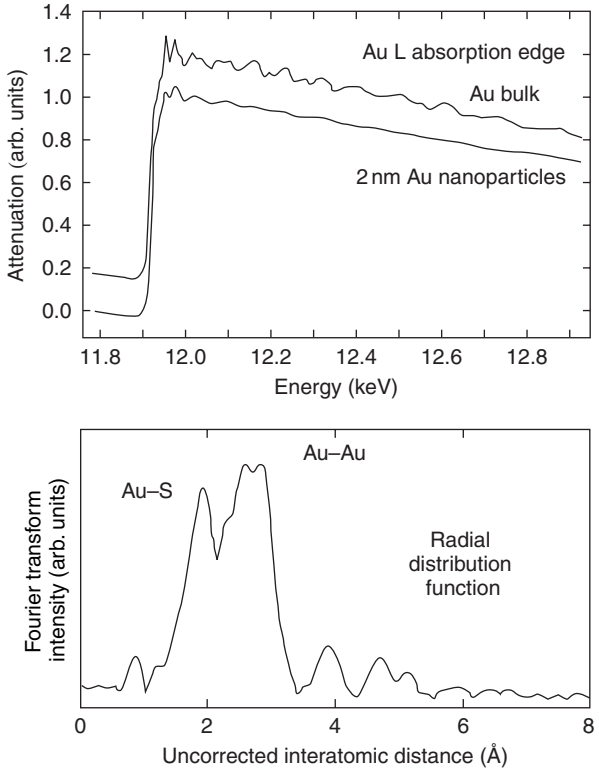

Figure 2.20 Schematic comparison between the L X-ray absorption edge in bulk gold and nanoparticulate gold together with the radial distribution function obtained via Fourier transformation of the EXAFS oscillations

Surface studies (often called SEXAFS or NEXAFS) can be performed using glancing angle incident radiation, whereas polarized X-ray sources can provide information on the orientation of species (using linearly polarized radiation) as well as magnetic dichroism (using circularly polarized radiation).

2.7.2 Radio frequency spectroscopy

Atomic nuclei contain protons and neutrons, both of which possess an intrinsic spin just like the electron (i.e., the spin quantum number s). In a nucleus the proton and neutron spins combine to give an overall nuclear spin (I). Certain nuclei, such as ^{1}H, ^{2}H, ^{7}Li, ^{11}B, ^{13}C, ^{14}C, ^{14}N, ^{15}N, ^{17}O, ^{19}F, ^{23}Na, ^{27}Al, ^{29}Si, ^{31}P and ^{35}Cl, have a nuclear spin that is non-zero and therefore they possess a magnetic moment and behave like a bar magnet. If a nucleus of spin 1/2 (e.g., ^{13}C) is placed in a magnetic field, it may orient itself with the field (in a low energy state) or it may align itself against the field (in a higher energy state). This is known as Zeeman splitting and the energy difference between these two states is proportional to the applied magnetic field multiplied by the magnetic moment of the nucleus. At equilibrium, the lower energy state will be slightly more populated, however in a field of 10–100 kGauss, transitions between the two energy states can be induced by radio frequency (RF) radiation ($\Delta E = h\nu$, where $\nu = 50-500$ MHz). Such transitions will absorb radiation and then re-emit upon relaxation; overall, radiation will be absorbed since there are initially very slightly more nuclei in the lower energy state. This forms the basis for the technique of nuclear magnetic resonance (NMR) spectroscopy, essentially the absorption of RF radiation by nuclei in a strong magnetic field, which can provide information on the local environment of the nuclei and hence the local bonding of atoms in a sample.

The experimental set-up for NMR involves an RF transmitter, a strong homogeneous magnetic field and an RF receiver. In a large molecule or solid, the presence of different chemical bonds may lead to the different magnetic nuclei of a particular atom species experiencing slightly different electronic environments. Hence the actual resonance frequency may vary between different nuclei and, to accommodate this, either the frequency or the magnetic field is swept through a range of values typically spanning the parts per million (ppm) range. As the detected signal is usually quite low, the radio frequency is often pulsed, which causes it to behave like a range of frequencies. As the signal decays following the pulse, it is Fourier transformed to obtain the separate resonance frequencies that are related to the differing electronic and hence chemical environments. After many such RF pulses, the spectral noise will average out, so improving statistics.

The actual value of the resonance frequency depends on a number of factors. Firstly, there is the chemical shift interaction caused by the fact that different nuclei have different chemical environments. Here the individual nuclei are shielded from the applied field to different extents by the electron cloud surrounding the nucleus, providing a means for identifying the local atomic environment; the actual value of the chemical shift is often referenced to a particular molecule, such as H or Si in tetramethylsilane (TMS) for ^{1}H or ^{29}Si NMR. In a molecule containing OH, for example, the electronegative oxygen atom essentially pulls the electron away from the hydrogen atom, which is largely unshielded, giving rise to a smaller chemical shift than would be

expected for other bonds (e.g., C–H) where the electron will behave so as to oppose the applied magnetic field. Secondly, the coupling of spins between neighbouring nuclei within a molecule, or even between different molecules, can change the observed NMR frequencies, leading to splitting and broadening of peaks. The magnitudes of these effects are dependent on the orientation of the magnetic nuclei with respect to the magnetic field.

The measured NMR frequencies and splittings can be used to characterize the chemical environments of the constituent atoms in a molecule, while the individual signal intensities are proportional to the number of nuclei in each particular environment. As a result, this technique is frequently used on solution species. However, in a disordered solid or a polycrystalline material, the orientation of the magnetic nuclei relative to the field is largely random. This anisotropy affects the chemical shifts and the dipolar couplings; it leads to a range of resonance peaks due to the large range of possible molecular or crystallite orientations. In low-viscosity liquids these problems disappear because the rapid molecular motion averages all these anisotropies to zero. However, in solids, translational movement is not possible and the resonances all overlap, leading to very broad peaks. This anisotropy problem can be overcome by spinning the sample around an axis inclined at an angle of 54.7° (the magic angle which trisects the three Cartesian axes) to the axis of the magnetic field. The powdered solid sample is typically contained in a plastic or ceramic rotor, and the rotor is spun by an air turbine around an axis oriented at exactly the magic angle using a spinning rate of a few kilohertz, which is comparable to the frequency spread of the anisotropy. The technique is known as magic angle spinning NMR (MASNMR) and produces a single, much narrower isotropically averaged NMR peak from solids, as in spectra obtained from solutions. A schematic MASNMR spectrum is shown in Figure 2.21. A variety of pulse sequences and magnetization changes (referred to as cross-polarization techniques) are often used to enhance signals from certain nuclei within a system. The MASNMR signal is obviously an average over the whole sample volume.

Dynamical information can also be obtained by NMR spectroscopy by measuring relaxation times. The interaction of a nucleus with its environment or with other (identical) nuclei can be obtained by measuring the longitudinal (spin–lattice) or transverse (spin–spin) relaxations, respectively. Nuclei precess about the applied magnetic field, $\mathbf{B}_0$, at their *Larmor* frequency. By applying a circularly polarized magnetic field orthogonal to $\mathbf{B}_0$ at the Larmor frequency, the nuclear spins will essentially follow a gyroscopic motion precessing about two different axes simultaneously, which allows one to isolate the nuclei of interest. This is because, if one removes the circularly polarized field after only 90°, the nucleus precesses around an axis away from the other nuclei. This orthogonality means that the behaviour of these nuclei can be monitored without interference from all the other nuclei with different Larmor frequencies precessing about $\mathbf{B}_0$. Of course, these nuclei will relax back to precession about $\mathbf{B}_0$, but then that is generally the sort of information that one is interested in. This $\pi/2$ spin flip is of central importance to dynamical measurements, but it must be used in parallel with one π spin flip (spin–lattice experiments) or multiple π spin flips (spin–spin experiments). By isolating individual nuclei it is possible to map their location throughout the sample, enabling a two-dimensional image to be constructed. Indeed it is possible to scan such two-dimensional sections through a solid and reconstruct 3D images. This is essentially a tomographic technique which forms the basis for magnetic resonance imaging (MRI).

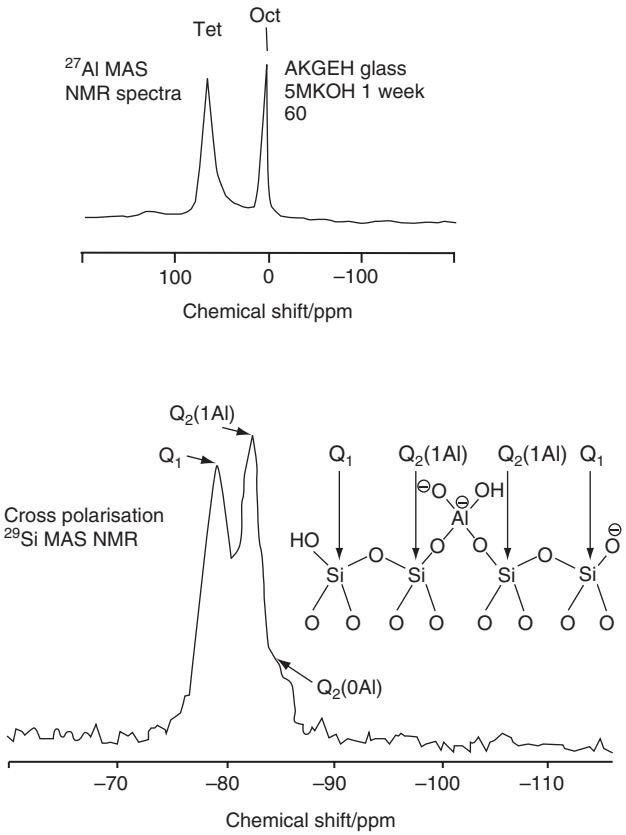

Figure 2.21 Schematic diagram of ^{27}Al and cross-polarized ^{29}Si MASNMR spectra from a Portland cement sample revealing the different chemical environments; inset is the derived chemical structure of the calcium silicate hydrate chains within the cement nanostructure. Data courtesy of Dr Ian Richardson and Dr Adrian Brough, School of Civil Engineering, University of Leeds

In biology and medicine, NMR images are constructed from ^{1}H NMR spectra of proton-bearing species such as water and fatty acids. However, the spatial resolution of these techniques is limited to at best microns or tens of microns, depending on the magnetic homogeneity of the sample. It is also possible to use the spin–spin relaxation to measure molecular diffusion coefficients using NMR. This is often called spin-echo NMR.

Another related spectroscopy worthy of mention is Mössbauer spectroscopy, which employs a γ-ray source containing an isotope with the same nuclear energy as the species of interest, for example, a ^{57}Co source is used to analyse ^{57}Fe, so as to excite transitions between nuclear states in the sample. The source is moved relative to the fixed sample so as to vary the incident γ-ray wavelength via the Doppler effect. Transitions between the ground and excited states are modified by interactions between the nucleus and the surrounding electronic and magnetic fields, termed hyperfine interactions, and these can be analysed to determine the valence states of

atoms in the sample. This technique is restricted to elements possessing isotopes with suitable nuclear transitions such as ^{57}Fe, ^{119}Sn and ^{197}Au.

Finally, electron paramagnetic resonance (EPR) is a microwave spectroscopy which employs a variable magnetic field to analyse crystals containing molecules or ions with unpaired electronic spins. The magnetic field causes Zeeman splitting of electronic energy levels, and transitions between these levels are induced by the absorption of microwave radiation.

2.7.3 Electron spectroscopy

Initially we concentrate on electron-induced spectroscopies in the electron microscope, either the SEM or TEM, in particular the analysis of the X-rays emitted by the sample as a result of electron beam ionization of atoms in the sample. True electron spectroscopy is discussed in Section 2.7.3.3.

2.7.3.1 X-ray emission in the SEM and TEM

As discussed in Section 2.3.3.3, following ionization of atoms in a sample by an electron beam, one possible de-excitation process is X-ray emission. The energy of the X-ray photon emitted when a single outer electron drops into the inner shell hole is given by the difference between the energies of the two excited states involved. A set of dipole selection rules determine which transitions are observed. Due to the well-defined nature of the various atomic energy levels, it is clear that the energies and associated wavelengths of the set of emitted X-rays will have characteristic values for each of the atomic species present in the specimen. By measuring either the energies or wavelengths of the X-rays emitted from the top surface of the sample, it is possible to determine which elements are present at the particular position of the electron probe; this is the basis for energy-dispersive and wavelength-dispersive X-ray analysis (EDX and WDX). WDX spectrometers use crystal monochromators to disperse the emitted X-ray spectrum in terms of diffraction angle and hence wavelength. Detectors then move along the arc of a circle centred on the specimen and collect the spectrum serially. EDX detectors collect X-rays in a near-parallel fashion and rely on the creation of electron–hole pairs in a biased silicon crystal; the number of electron–hole pairs and hence current is directly proportional to the energy of the incident X-ray. Fast electronics allows separate pulses of X-rays to be discriminated and measured. EDX detectors often have some form of window which, depending on the material, may reduce sensitivity to light elements ($Z < 11$). Compared to EDX, WDX is slow but has increased resolution and hence sensitivity to all elements, particularly light elements. WDX is confined to dedicated analytical SEMs known as electron probe microanalysers (EPMA), whereas EDX detectors may be fitted as an add-on attachment to most SEMs and TEMs.

Figure 2.22 shows a typical electron-generated X-ray emission spectrum from molybdenum oxide. The Mo Kα, Kβ and Lα X-ray lines as well as the O Kα line are superimposed on the bremsstrahlung background. The bremsstrahlung X-rays are not characteristic of any particular atom but depend principally on specimen thickness. To a first approximation, peak intensities are proportional to the atomic concentration of the

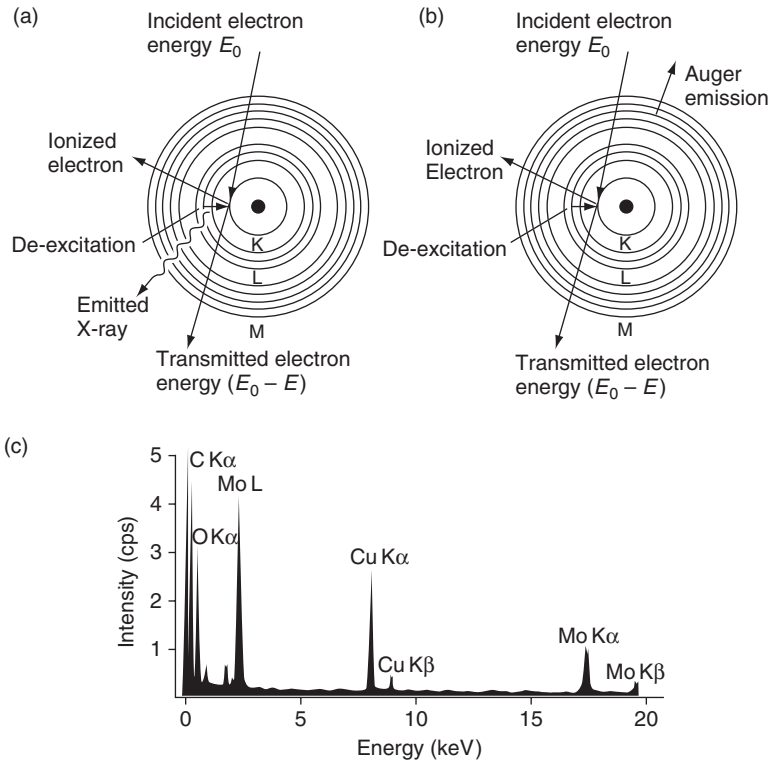

Figure 2.22 Schematic diagram of (a) de-excitation by X-ray emission, (b) de-excitation by Auger electron following ionization by an electron of incident energy E_0. (c) Schematic energy-dispersive X-ray emission spectrum from a thin specimen of molybdenum oxide on a carbon support film; the peaks are labelled with standard X-ray notation as discussed in the text; the peaks due to copper are from the specimen holder

element and, with careful measurement, EDX and WDX can detect levels of elements down to 0.1 at%.

When EDX is used in the TEM as opposed to the SEM, the reduced sample thickness and probe size lead to much higher spatial resolution. At 100 kV, a 100 nm thick sample typically gives a beam broadening of the order of 10 nm. EDX analysis will collect all X-rays produced isotropically within the beam-broadened volume, and this is of obvious importance for the analysis of nanostructures. EDX maps of multilayers and nanoscale precipitates have demonstrated resolutions of at best a few nanometres in ultrathin sample areas.

2.7.3.2 Cathodoluminescence in the SEM and STEM

Cathodoluminescence (CL) is generated by the recombination of electron–hole pairs produced by electron beam induced ionization of valence electrons in a semiconducting or insulating sample. The exact CL energies (and therefore wavelengths) depend on the width of the band gap of the material. Any changes in temperature, crystal structure

(e.g., different polymorphs), impurity levels or defect concentrations will modify the band gap and thus the CL wavelength. Spectroscopic analysis of the CL emission can therefore provide information on the microstructure of the sample. The emitted photon intensity is low and therefore low scan rates and high probe currents (large probe diameters) are required, limiting the resolution to 1–10 μm in the SEM. In the STEM the resolution can be improved considerably, making the technique viable for nanostructural analysis.

2.7.3.3 Electron energy loss spectroscopy

Electron energy loss spectroscopy (EELS) in a TEM or STEM involves analysis of the inelastic scattering suffered by the electron beam via measurement of the energy distribution of the transmitted electrons. The technique allows high-resolution elemental mapping and the measurement of local electronic structures for the determination of the local chemical bonding, such as that present at an interface or defect.

The transmitted electrons are dispersed according to their energy losses using a magnetic sector spectrometer (see Figure 2.8). EELS spectra can be recorded in parallel using a scintillator and diode array detection system, with the TEM in either diffraction or image mode; spatially resolved spectra are obtained using a small, highly focused probe in a STEM. The transmitted EELS signal is highly forward-peaked and the collected signal can be well defined using a suitable aperture that can significantly improve the ultimate spatial resolution of the analysis compared to EDX, where beam broadening will limit the ultimate (lateral) spatial resolution for microanalysis. EELS spectra can be obtained from individual atomic columns.

The various energy losses observed in a typical EELs spectrum are shown in Figure 2.23a, which gives the scattered electron intensity as a function of the decrease in kinetic energy (the energy loss E) of the transmitted fast electrons and essentially represents the response of the electrons in the solid to the disturbance introduced by the incident electrons. In a specimen of thickness less than the mean free path for inelastic scattering (roughly 100 nm at 100 keV), by far the most intense feature is the zero-loss peak at 0 eV energy loss, which contains all the elastically and quasi-elastically (i.e., phonon) scattered electron components. Neglecting the effect of the spectrometer, the full width at half maximum (FWHM) of the zero-loss peak is usually limited by the energy spread inherent in the electron source. In a TEM the energy spread will generally lie between $\sim$0.1 and 3 eV, depending on the type of emitter, and this parameter often determines the overall spectral resolution.

The *low-loss* region of the EELs spectrum, extending from 0 to about 50 eV, corresponds to the excitation of electrons in the outermost atomic orbitals which are often delocalized due to interatomic bonding and extend over several atomic sites. This region therefore reflects the solid-state character of the sample. The smallest energy losses (10–100 meV) arise from phonon emission, but these are usually subsumed in the zero-loss peak. The dominant feature in the low-loss spectrum arises from collective, resonant oscillations of the valence electrons known as plasmons. The energy of the plasmon peak is governed by the density of the valence electrons, and the width by the rate of decay of the resonant mode. In a thicker specimen ($\geq$100 nm) there are additional peaks at multiples of the plasmon energy, corresponding to the excitation of more than

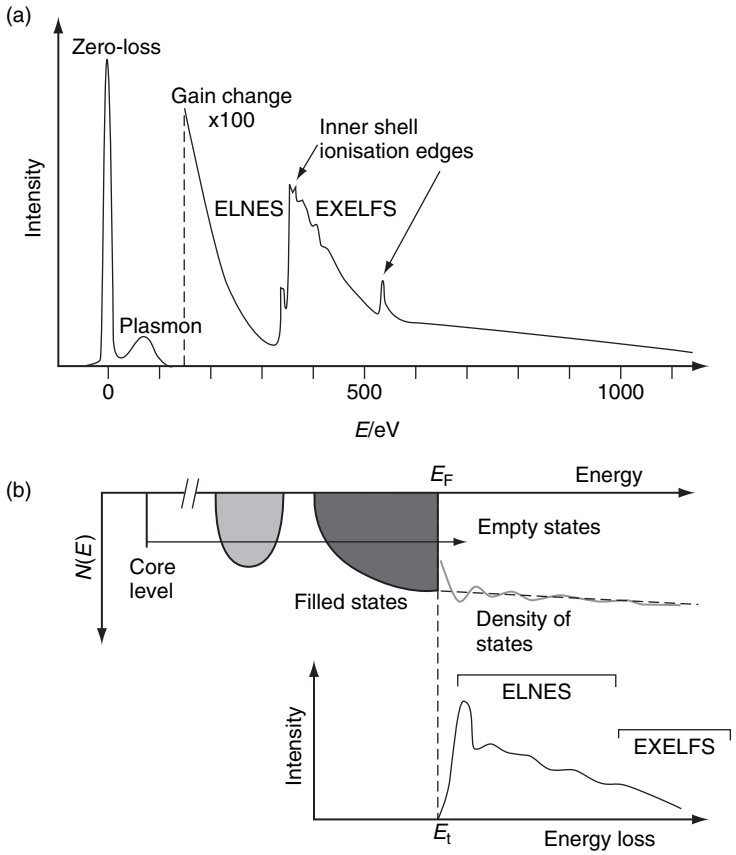

Figure 2.23 Schematic diagram of (a) an EELS spectrum and (b) the ELNES intensity which reflects the unoccupied DOS above the Fermi level

one plasmon; the intensities of these multiple plasmon peaks follow a Poisson distribution. A further feature in the low-loss spectra of insulators are peaks, known as interband transitions, which correspond to the excitation of valence electrons to low-energy unoccupied electronic states above the Fermi level. These single-electron excitations may lead to a shift in the energy of the plasmon resonance. The low-loss region is used mainly to determine the specimen thickness and to correct for the effects of multiple inelastic scattering when performing quantitative microanalysis on thicker specimens. In a more detailed analysis the overall shape of the low-loss region may be related to the dielectric response function of the material, which allows a correlation with optical measurements, including reflectivity and band gap determination in insulators and semiconductors.

The *high-loss* region of the EELs spectrum extends from about 50 eV to several thousand electron volts and corresponds to the excitation of electrons from localized orbitals on a single atomic site to extended, unoccupied electron energy levels just above the Fermi level of the material. This region therefore reflects the atomic

character of the specimen. As the energy loss progressively increases, this region exhibits steps or edges superimposed on the monotonically decreasing background intensity; the intensity at 2000 eV is typically eight orders of magnitude less than at the zero-loss peak and so, for clarity, in Figure 2.23 a gain change has been inserted in the linear intensity scale at 150 eV. These edges correspond to excitation of inner shell electrons and are known as ionization edges. The various EELS ionization edges are classified using the standard spectroscopic notation; e.g., K excitation for ionization of 1s electrons. Since the energy of the edge threshold is determined by the binding energy of the particular electron subshell within an atom, the atomic type may be easily identified. The signal under the ionization edge extends beyond the threshold, since the amount of kinetic energy given to the excited electron is not fixed. The intensity or area under the edge is proportional to the number of atoms present and this allows the technique to be used for quantitative analysis. EELS is particularly sensitive to the detection and quantification of light elements ($Z < 11$) as well as transition metals and rare earths.

If electrons are scattered via inelastic collisions with K-shell electrons of free atoms (e.g., gases) the core-loss edges are sharp, sawtooth-like steps displaying no features. Other core-loss excitations in free atoms display a variety of basic edge shapes that are essentially determined by the degree of overlap between the initial and final state wavefunctions. In solids, however, the unoccupied electronic states near the Fermi level may be appreciably modified by chemical bonding, leading to a complex density of states (DOS), and this is reflected in the electron energy loss near-edge structure (ELNES), which modifies the basic atomic shape within the first 30–40 eV above the edge threshold. The ELNES effectively represents the unoccupied DOS (above the Fermi level) in the environment of the atom(s) being ionized, as shown in Figure 2.23b, hence it gives information on the local structure and bonding. Beyond the near-edge region, superimposed on the gradually decreasing tail of the core-loss edge, a region of weaker, extended oscillations is observed, known as the extended energy loss fine structure (EXELFS). As in EXAFS (Section 2.7.1.4), the period of the oscillations may be used to determine bond distances, while the amplitude reflects the coordination number of the particular atom.

Energy-filtered TEM (EFTEM) involves the selection of a specific energy loss value, or narrow range of energy losses, from the transmitted electron beam via use of an energy-selecting slit after the spectrometer. Using only zero-loss (elastically) scattered electrons to form images and diffraction patterns increases contrast and resolution, allowing easier interpretation than with unfiltered data. Chemical mapping may be achieved by acquiring and processing images formed by electrons which have undergone inner shell ionization events. An alternative approach, used when mapping is performed with an EELs spectrometer and detector attached to an STEM, is to raster the electron beam across the specimen and record an EELS spectrum at every point (x,y) – this technique being known as 'spectrum imaging'. The complete data set may then be processed to form a 2D quantitative map of the sample using either standard elemental quantification procedures or the position and/or intensity of characteristic low-loss or ELNES features. The inherently high spatial resolution of the EELS technique has allowed maps of elemental distributions to be formed at subnanometre resolution. In addition, the possibility of plasmon or ELNES chemical bonding maps has also been demonstrated.

2.8 SURFACE ANALYSIS AND DEPTH PROFILING

One basic form of surface analysis, highly relevant for nanostructured materials, is the determination of the specific surface area (in $m^2\,g^{-1}$) of a sample. This is usually performed by measurement of the volume of gas adsorbed onto a specific mass of sample, as a function of gas pressure or more usually relative gas pressure (pressure/saturation vapour pressure). This measurement is known as an adsorption isotherm. Samples are carefully outgassed under vacuum and cooled to low temperatures, prior to controlled physical adsorption of a suitable gas, such as nitrogen, which is small enough ($\sim$0.4 nm) to access even the smallest nanopores. The form of the resulting adsorption isotherm depends on the nature of the adsorbent and the adsorbate and may be analysed using a number of theoretical models, all of which make assumptions about the nature of the gas–solid interaction and the types of pores present. In essence the models attempt to identify when one monolayer of gas molecules has been adsorbed at the surface so as to determine the surface area. In addition to the total surface area, by modelling the data it is also possible to calculate the average pore size and the pore size distribution within a porous material, although it is important to realise that such experiments will only access open (not closed) porosity within a sample.

In addition to the basic adsorption methods, surface chemical analysis techniques directly probe the elemental composition and chemical state of the outermost atomic layers (i.e., 0.1–10 nm below the surface) of all types of solid materials. The principal techniques used for probing surface chemistry are based on electron, mass and vibrational spectroscopy. In terms of nanostructured materials and devices, the surface is obviously an important consideration in terms of properties and material performance, as well as the nature of the interaction between the sample and the environment. Properties of adhesion, colour, catalytic activity, biocompatibility and lubrication, for example, can all be strictly controlled by the surface of a material. However, if samples are of reduced dimension laterally this can present particular problems as high lateral spatial resolution on the surface is then also required.

Due to the extreme surface sensitivity of most surface analysis methods, the amount of information available decays rapidly with degrading vacuum since, as highlighted above, ambient gases readily adsorb onto the surface of the specimen. Most surface analysis methods therefore require analysis to be performed under UHV conditions (10^{-10} mbar) where the pressure is an order of magnitude lower than for, say, electron microscopy. For this reason instrument design has evolved around a substantial evacuation system involving a two-chamber apparatus incorporating metal-on-metal vacuum seals for maintaining a good vacuum during analysis. The first preparation chamber is for sample introduction and high-vacuum experimentation (e.g., gas adsorption experiments, sample heating or ion beam etching), while the second chamber is the analysis chamber containing the various radiation sources and detection systems. Figure 2.24 is a schematic diagram of a surface analysis system.

Depth profiling provides surface analysis coupled with subsurface information. Techniques such as neutron reflectometry can provide information about the structure of samples many microns below the surface, without the often prohibitive UHV requirements. Other techniques such as ion beam analysis can also provide similar information. The key to depth profiling is the ability to have a probe which can interrogate deep into the sample. Neutrons, X-rays, and high-energy ions are generally

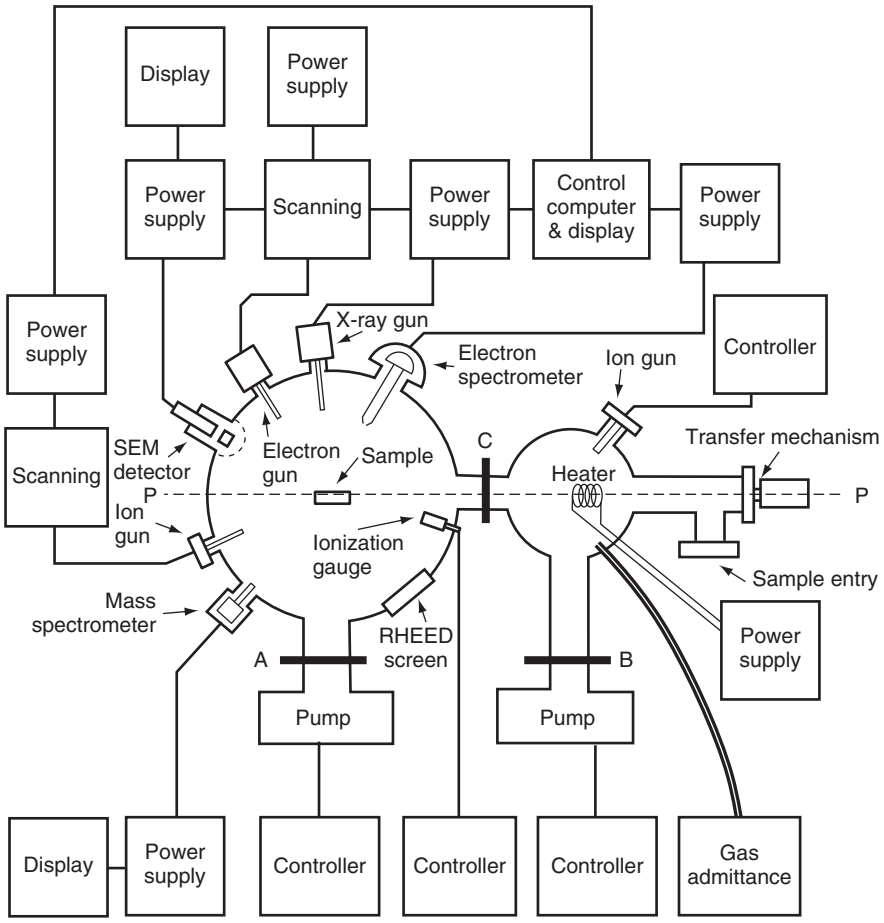

Figure 2.24 Schematic diagram of a surface analysis system incorporating XPS, Auger and SIMS

very effective at this. We shall consider depth profiling with ions and neutrons below, but first we turn to the surface-specific electron spectroscopy and mass spectrometry techniques.

2.8.1 Electron spectroscopy of surfaces

X-ray photoelectron spectroscopy (XPS) and Auger electron spectroscopy (AES) have shown the greatest applicability to the widest range of materials surfaces. XPS and AES are conveniently linked as techniques since they are both dependent on the analysis of low-energy electrons emitted from surfaces, in the range 10–3000 eV. As a result, both techniques can use the same electron spectroscopy instrumentation, although they tend to use different excitation sources.

In XPS the sample is irradiated with a beam of usually monochromatic, low-energy X-rays. Photoelectron emission results from the atoms in the specimen surface, and the