Physical Principles of Electron Microscopy

Ray F. Egerton

Physical Principles of Electron Microscopy

An Introduction to TEM, SEM, and AEM

With 122 Figures

Springer

Ray F. Egerton
Department of Physics, University of Alberta
412 Avadh Bhatia Physics Laboratory
Edmonton, Alberta, Canada
T6G 2R3

Library of Congress Control Number: 2005924717

ISBN-10: 0-387-25800-0 Printed on acid-free paper.
ISBN-13: 978-0387-25800-0

Printed in the United States of America. (EB)

9 8 7 6 5 4 3 2 1

springeronline.com

To Maia

Contents

PREFACE

The telescope transformed our view of the universe, leading to cosmological theories that derive support from experiments involving elementary particles. But microscopes have been equally important, by helping us to understand both inanimate matter and living objects at *their* elementary level. Initially, these instruments relied on the focusing of visible light, but within the past 50 years other forms of radiation have been used. Of these, electrons have arguably been the most successful, by providing us with direct images down to the atomic level.

The purpose of this book is to introduce concepts of electron microscopy and to explain some of the basic physics involved at an undergraduate level. It originates from a one-semester course at the University of Alberta, designed to show how the principles of electricity and magnetism, optics and modern physics (learned in first or second year) have been used to develop instruments that have wide application in science, medicine and engineering. Finding a textbook for the course has always been a problem; most electron microscopy books overwhelm a non-specialist student, or else they concentrate on practical skills rather than fundamental principles. Over the years, this course became one of the most popular of our "general interest" courses offered to non-honors students. It would be nice to think that the availability of this book might facilitate the introduction of similar courses at other institutions.

At the time of writing, electron microscopy is being used routinely in the semiconductor industry to examine devices of sub-micrometer dimensions. Nanotechnology also makes use of electron beams, both for characterization and fabrication. Perhaps a book on the basics of TEM and SEM will benefit the engineers and scientists who use these tools. The more advanced student or professional electron microscopist is already well served by existing

textbooks, such as Williams and Carter (1996) and the excellent Springer books by Reimer. Even so, I hope that some of my research colleagues may find the current book to be a useful supplement to their collection.

My aim has been to teach general concepts, such as how a magnetic lens focuses electrons, without getting into too much detail – as would be needed to actually design a magnetic lens. Because electron microscopy is interdisciplinary, both in technique and application, the physical principles being discussed involve not only physics but also aspects of chemistry, electronics, and spectroscopy. I have included a short final chapter outlining some recent or more advanced techniques, to illustrate the fact that electron microscopy is a "living" subject that is still undergoing development.

Although the text contains equations, the mathematics is restricted to simple algebra, trigonometry, and calculus. SI units are utilized throughout. I have used *italics* for emphasis and **bold** characters to mark technical terms when they first appear. On a philosophical note: although wave mechanics has proved invaluable for accurately calculating the properties of electrons, classical physics provides a more intuitive description at an elementary level. Except with regard to diffraction effects, I have assumed the electron to be a particle, even when treating "phase contrast" images. I hope Einstein would approve.

To reduce publishing costs, the manuscript was prepared as camera-ready copy. I am indebted to several colleagues for proofreading and suggesting changes to the text; in particular, Drs. Marek Malac, Al Meldrum, Robert Wolkow, and Rodney Herring, and graduate students Julie Qian, Peng Li, and Feng Wang.

Ray Egerton
University of Alberta
Edmonton, Canada
regerton@ualberta.ca

January 2005

Chapter 1

AN INTRODUCTION TO MICROSCOPY

Microscopy involves the study of objects that are too small to be examined by the unaided eye. In the SI (metric) system of units, the sizes of these objects are expressed in terms of sub-multiples of the meter, such as the **micrometer** (1 μm = 10^{-6} m, also called a *micron*) and also the **nanometer** (1 nm = 10^{-9} m). Older books use the Angstrom unit (1 Å = 10^{-10} m), not an official SI unit but convenient for specifying the distance between atoms in a solid, which is generally in the range 2 – 3 Å.

To describe the wavelength of fast-moving electrons or their behavior inside an atom, we need even smaller units. Later in this book, we will make use of the **picometer** (1 pm = 10^{-12} m).

The diameters of several small objects of scientific or general interest are listed in Table 1-1, together with their approximate dimensions.

Table 1-1. Approximate sizes of some common objects and the smallest magnification M* required to distinguish them, according to Eq. (1.5).

Object	Typical diameter *D*	M* = 75μm / *D*
Grain of sand	1 mm = 1000 μm	None
Human hair	150 μm	None
Red blood cell	10 μm	7.5
Bacterium	1 μm	75
Virus	20 nm	4000
DNA molecule	2 nm	40,000
Uranium atom	0.2 nm = 200 pm	400,000

1.1 Limitations of the Human Eye

Our concepts of the physical world are largely determined by what we *see* around us. For most of recorded history, this has meant observation using the human eye, which is sensitive to radiation within the **visible region** of the electromagnetic spectrum, meaning wavelengths in the range 300 – 700 nm. The eyeball contains a fluid whose refractive index ($n \approx 1.34$) is substantially different from that of air ($n \approx 1$). As a result, most of the refraction and focusing of the incoming light occurs at the eye's curved front surface, the **cornea**; see Fig. 1-1.

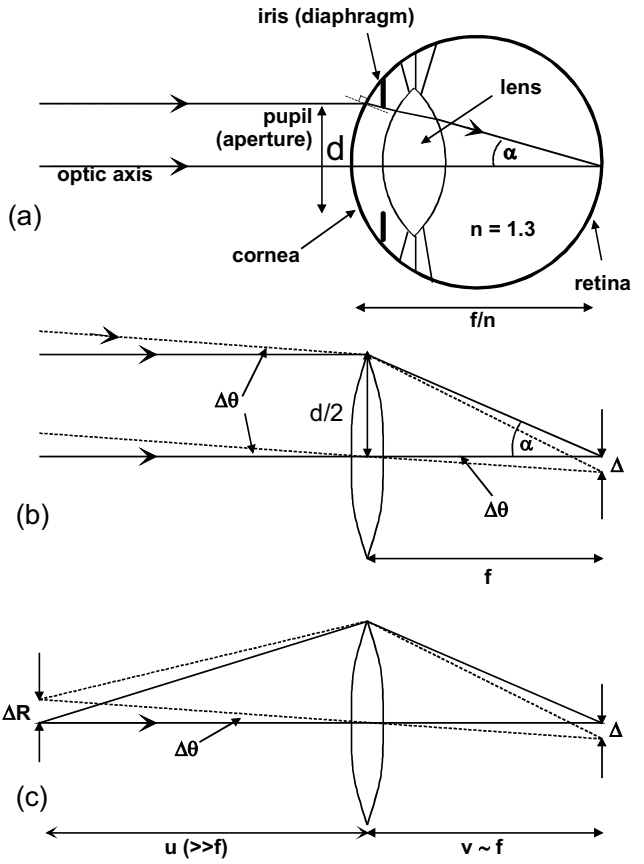

Figure 1-1. (a) A physicist's conception of the human eye, showing two light rays focused to a single point on the retina. (b) Equivalent thin-lens ray diagram for a *distant* object, showing *parallel* light rays arriving from opposite ends (solid and dashed lines) of the object and forming an image (in air) at a distance f (the focal length) from the thin lens. (c) Ray diagram for a *nearby* object (object distance $u = 25$ cm, image distance v slightly less than f).

In order to focus on objects located at *different* distances (referred to as **accommodation**), the eye incorporates an elastically deformable **lens** of slightly higher refractive index ($n \sim 1.44$) whose shape and focusing power are controlled by eye muscles. Together, the cornea and lens of the eye behave like a single glass lens of variable focal length, forming a **real image** on the curved **retina** at the back of the eyeball. The retina contains photosensitive **receptor** cells that send electrochemical signals to the brain, the strength of each signal representing the local intensity in the image. However, the photochemical processes in the receptor cells work over a limited range of image intensity, therefore the eye controls the amount of light reaching the retina by varying the diameter d (over a range 2 – 8 mm) of the **aperture** of the eye, also known as the **pupil.** This aperture takes the form of a circular hole in the **diaphragm** (or **iris**), an opaque disk located between the lens and the cornea, as shown in Fig. 1-1.

The **spatial resolution** of the retinal image, which determines how *small* an object can be and still be separately identified from an adjacent and similar object, is determined by three factors: the size of the receptor cells, imperfections in the focusing (known as **aberrations**), and **diffraction** of light at the entrance pupil of the eye. Diffraction cannot be explained using a particle view of light (geometrical or ray optics); it requires a wave interpretation (physical optics), according to which any image is actually an interference pattern formed by light rays that take different paths to reach the same point in the image. In the simple situation that is depicted in Fig. 1-2 , a

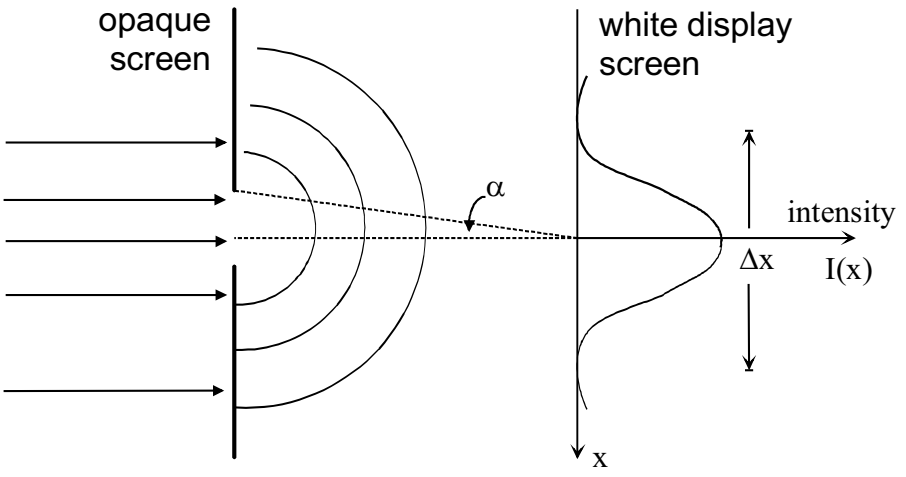

Figure 1-2. Diffraction of light by a slit, or by a circular aperture. Waves spread out from the aperture and fall on a white screen to produce a disk of confusion (Airy disk) whose intensity distribution $I(x)$ is shown by the graph on the right.

parallel beam of light strikes an opaque diaphragm containing a circular aperture whose radius subtends an angle α at the center of a white viewing screen. Light passing through the aperture illuminates the screen in the form of a circular pattern with diffuse edges (a **disk of confusion**) whose diameter Δx *exceeds* that of the aperture. In fact, for an aperture of small diameter, diffraction effects cause Δx actually to *increase* as the aperture size is reduced, in accordance with the **Rayleigh** criterion:

$$\Delta x \approx 0.6\,\lambda / \sin\alpha \tag{1.1}$$

where λ is the wavelength of the light being diffracted.

Equation (1.1) can be applied to the eye, with the aid of Fig. 1-1b, which shows an equivalent image formed in air at a distance f from a single focusing lens. For wavelengths in the middle of the visible region of the spectrum, $\lambda \approx 500$ nm and taking $d \approx 4$ mm and $f \approx 2$ cm, the geometry of Fig. 1-1b gives $\tan\alpha \approx (d/2)/f = 0.1$, which implies a small value of α and allows use of the small-angle approximation: $\sin\alpha \approx \tan\alpha$. Equation (1.1) then gives the diameter of the disk of confusion as $\Delta x \approx (0.6)(500\text{ nm})/0.1 = 3\ \mu$m.

Imperfect focusing (aberration) of the eye contributes a roughly *equal* amount of image blurring, which we therefore take as 3 μm. In addition, the receptor cells of the retina have diameters in the range 2 μm to 6 μm (mean value ≈ 4 μm). Apparently, evolution has refined the eye up to the point where further improvements in its construction would lead to relatively little improvement in overall resolution, relative to the diffraction limit Δx imposed by the wave nature of light.

To a reasonable approximation, these three different contributions to the retinal-image blurring can be combined **in quadrature** (by adding squares), treating them in a similar way to the statistical quantities involved in error analysis. Using this procedure, the overall image blurring Δ is given by:

$$(\Delta)^2 = (3\ \mu\text{m})^2 + (3\ \mu\text{m})^2 + (4\ \mu\text{m})^2 \tag{1.2}$$

which leads to $\Delta \approx 6$ μm as the blurring of the retinal image. This value corresponds to an *angular* blurring for distant objects (see Fig. 1-1b) of

$$\Delta\theta \approx (\Delta/f) \approx (6\ \mu\text{m})/(2\text{ cm}) \approx 3 \times 10^{-4}\text{ rad}$$
$$\approx (1/60)\text{ degree} = 1\text{ minute of arc} \tag{1.3}$$

Distant objects (or details within objects) can be separately distinguished if they subtend angles larger than this. Accordingly, early astronomers were able to determine the positions of bright stars to within a few minutes of arc, using only a dark-adapted eye and simple pointing devices. To see greater detail in the night sky, such as the faint stars within a galaxy, required a *telescope*, which provided *angular* magnification.

Changing the shape of the lens in an adult eye alters its overall focal length by only about 10%, so the *closest* object distance for a focused image on the retina is $u \approx 25$ cm. At this distance, an angular resolution of 3×10^{-4} rad corresponds (see Fig. 1c) to a lateral dimension of:

$$\Delta R \approx (\Delta\theta)\, u \approx 0.075 \text{ mm} = 75 \text{ μm} \tag{1.4}$$

Because $u \approx 25$ cm is the smallest object distance for clear vision, $\Delta R = 75$ μm can be taken as the diameter of the *smallest* object that can be resolved (distinguished from neighboring objects) by the unaided eye, known as its *object* resolution or the **spatial resolution** in the object plane.

Because there are many interesting objects *below* this size, including the examples in Table 1-1, an optical device with **magnification factor** M (> 1) is needed to see them; in other words, a **microscope**.

To resolve a small object of diameter D, we need a magnification M^* such that the *magnified* diameter ($M^* D$) at the eye's object plane is greater or equal to the object resolution ΔR (≈ 75 μm) of the eye. In other words:

$$M^* = (\Delta R)/D \tag{1.5}$$

Values of this minimum magnification are given in the right-hand column of Table 1-1, for objects of various diameter D.

1.2 The Light-Optical Microscope

Light microscopes were developed in the early 1600's, and some of the best observations were made by Anton van Leeuwenhoek, using tiny glass lenses placed very close to the object and to the eye; see Fig. 1-3. By the late 1600's, this Dutch scientist had observed blood cells, bacteria, and structure *within* the cells of animal tissue, all revelations at the time. But this simple one-lens device had to be positioned very accurately, making observation very tiring in practice.

For routine use, it is more convenient to have a **compound microscope**, containing at least two lenses: an **objective** (placed close to the *object* to be magnified) and an **eyepiece** (placed fairly close to the *eye*). By increasing its dimensions or by employing a larger number of lenses, the magnification M of a compound microscope can be increased indefinitely. However, a large value of M does not guarantee that objects of vanishingly small diameter D can be visualized; in addition to satisfying Eq. (1-5), we must ensure that aberrations and diffraction *within the microscope* are sufficiently low.

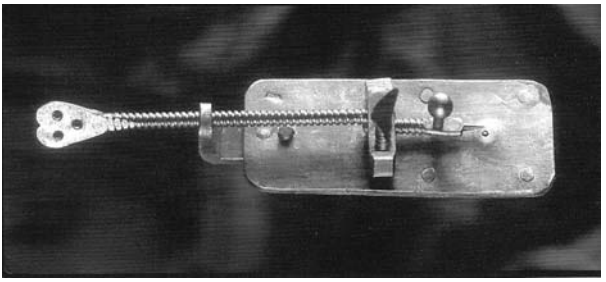

Figure 1-3. One of the single-lens microscopes used by van Leeuwenhoek. The adjustable pointer was used to center the eye on the optic axis of the lens and thereby minimize image aberrations. Courtesy of the FEI Company.

Nowadays, the aberrations of a light-optical instrument can be made unimportant by grinding the lens surfaces to a correct shape or by spacing the lenses so that their aberrations are compensated. But even with such aberration-corrected lenses, the spatial resolution of a compound microscope is limited by *diffraction* at the objective lens. This effect depends on the diameter (aperture) of the lens, just as in the case of diffraction at the pupil of the eye or at a circular hole in an opaque screen. With a large-aperture lens ($\sin \alpha \approx 1$), Eq. (1.1) predicts a resolution limit of just over half the wavelength of light, as first deduced by Abbé in 1873. For light in the middle of the visible spectrum ($\lambda \approx 0.5$ μm), this means a best-possible object resolution of about 0.3 μm.

This is a substantial improvement over the resolution (≈ 75 μm) of the unaided eye. But to achieve this resolution, the microscope must magnify the object to a diameter at least equal to ΔR, so that *overall* resolution is determined by microscope diffraction rather than the eye's limitations, requiring a microscope magnification of $M \approx$ (75 μm)/(0.3 μm) = 250. Substantially larger values ("empty magnification") do not significantly improve the sharpness of the magnified image and in fact reduce the **field of view**, the area of the object that can be simultaneously viewed in the image.

Light-optical microscopes are widely used in research and come in two basic forms. The **biological** microscope (Fig. 1-4a) requires an optically transparent specimen, such as a thin slice (section) of animal or plant tissue. Daylight or light from a lamp is directed via a lens or mirror through the specimen and into the microscope, which creates a real image on the retina of the eye or within an attached camera. Variation in the light intensity (**contrast**) in the image occurs because different parts of the specimen *absorb* light to differing degrees. By using **stains** (light-absorbing chemicals attach themselves preferentially to certain regions of the specimen), the contrast can be increased; the image of a tissue section may then reveal the

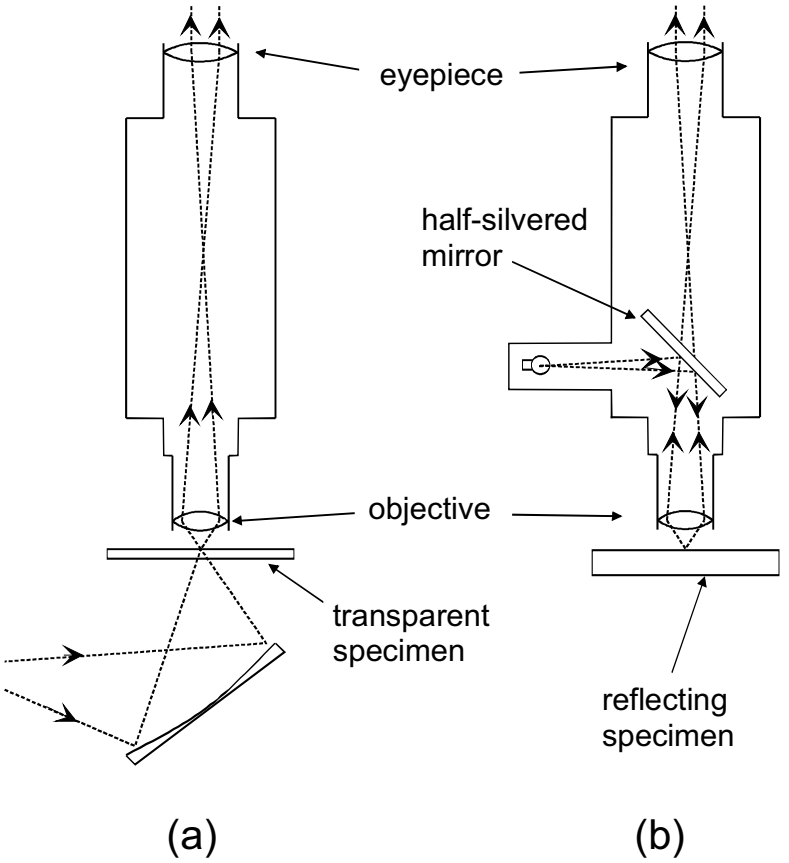

Figure 1-4. Schematic diagrams of (a) a biological microscope, which images light transmitted through the specimen, and (b) a metallurgical microscope, which uses light (often from a built-in illumination source) reflected from the specimen surface.

individual components (organelles) within each biological cell. Because the light travels *through* the specimen, this instrument can also be called a **transmission** light microscope. It is used also by geologists, who are able to prepare rock specimens that are thin enough (below 0.1 μm thickness) to be optically transparent.

The **metallurgical** microscope (Fig. 1-4b) is used for examining metals and other materials that cannot easily be made thin enough to be optically transparent. Here, the image is formed by light *reflected* from the surface of the specimen. Because perfectly smooth surfaces provide little or no contrast, the specimen is usually immersed for a few seconds in a chemical **etch**, a solution that preferentially attacks certain regions to leave an uneven surface whose reflectivity varies from one location to another. In this way,

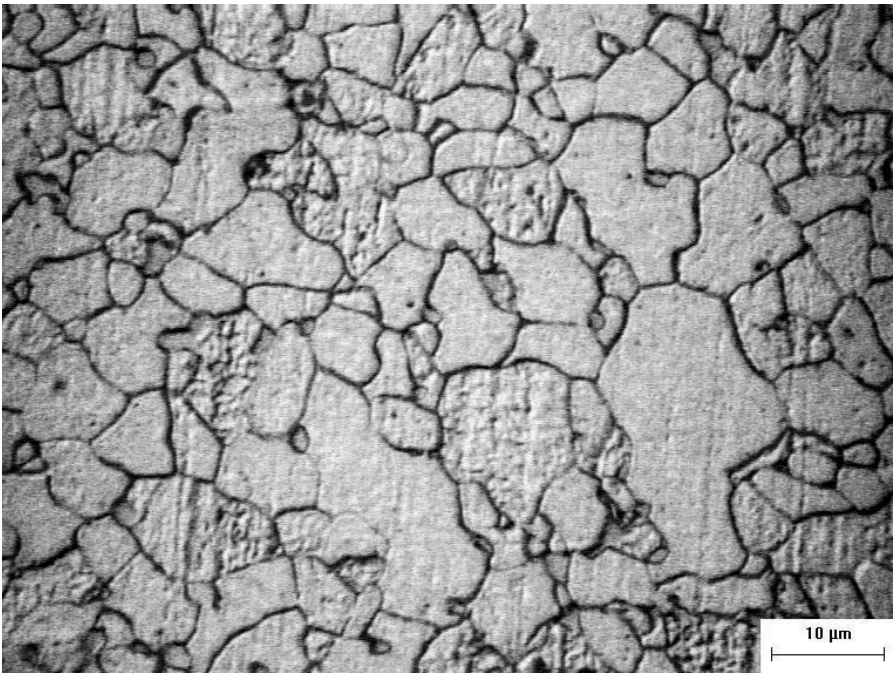

Figure 1-5. Light-microscope image of a polished and etched specimen of X70 pipeline steel, showing dark lines representing the grain boundaries between ferrite (bcc iron) crystallites. Courtesy of Dr. D. Ivey, University of Alberta.

the microscope reveals the microstructure of crystalline materials, such as the different phases present in a metal alloy. Most etches preferentially dissolve the regions between individual crystallites (grains) of the specimen, where the atoms are less closely packed, leaving a grain-boundary groove that is visible as a dark line, as in Fig. 1-5. The metallurgical microscope can therefore be used to determine the grain shape and grain size of metals and alloys.

As we have seen, the resolution of a light-optical microscope is limited by diffraction. As indicated by Eq. (1.1), one possibility for *improving* resolution (which means *reducing* Δx, and therefore Δ and ΔR) is to decrease the wavelength λ of the radiation. The simplest option is to use an **oil-immersion** objective lens: a drop of a transparent liquid (refractive index n) is placed between the specimen and the objective so that the light being focused (and diffracted) has a reduced wavelength: λ/n. Using cedar oil (n = 1.52) allows a 34% improvement in resolution.

Greater improvement in resolution comes from using **ultraviolet** (UV) radiation, meaning wavelengths in the range 100 – 300 nm. The light source

can be a gas-discharge lamp and the final image is viewed on a phosphor screen that converts the UV to visible light. Because ordinary glass strongly absorbs UV light, the focusing lenses must be made from a material such as quartz (transparent down to 190 nm) or lithium fluoride (transparent down to about 100 nm).

1.3 The X-ray Microscope

Being electromagnetic waves with a wavelength shorter than those of UV light, x-rays offer the possibility of even better spatial resolution. This radiation cannot be focused by convex or concave lenses, as the refractive index of solid materials is close to that of air (1.0) at x-ray wavelengths. Instead, x-ray focusing relies on devices that make use of *diffraction* rather than refraction.

Hard x-rays have wavelengths below 1 nm and are diffracted by the planes of atoms in a solid, whose spacing is of similar dimensions. In fact, such diffraction is routinely used to determine the atomic structure of solids. X-ray microscopes more commonly use **soft x-rays**, with wavelengths in the range 1 nm to 10 nm. Soft x-rays are diffracted by structures whose periodicity is several nm, such as thin-film multilayers that act as focusing mirrors, or **zone plates**, which are essentially diffraction gratings with circular symmetry (see Fig. 5-23) that focus monochromatic x-rays (those of a single wavelength); as depicted Fig. 1-6.

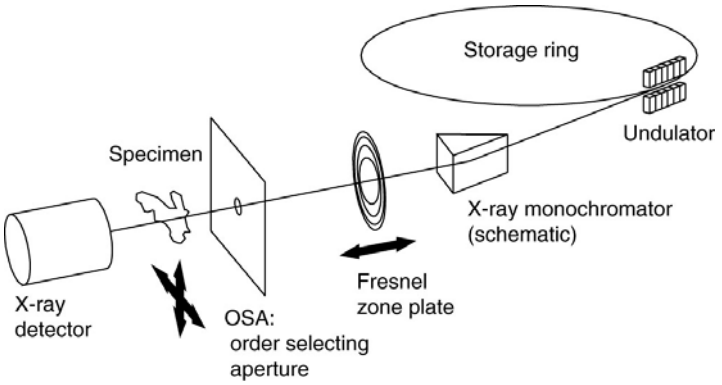

Figure 1-6. Schematic diagram of a scanning transmission x-ray microscope (STXM) attached to a synchrotron radiation source. The monochromator transmits x-rays with a narrow range of wavelength, and these monochromatic rays are focused onto the specimen by means of a Fresnel zone plate. The order-selecting aperture ensures that only a single x-ray beam is focused and scanned across the specimen. From Neuhausler *et al*. (1999), courtesy of Springer-Verlag.

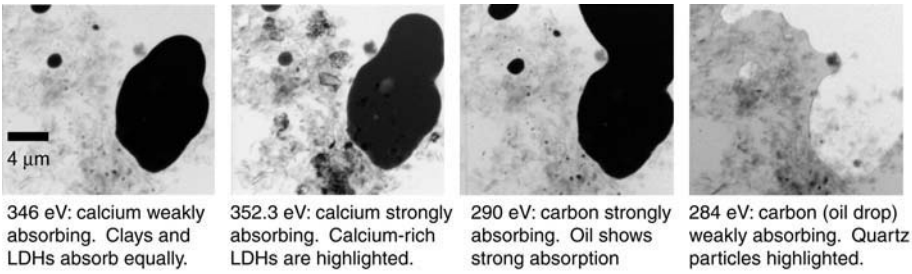

Figure 1-7. Scanning transmission x-ray microscope (STXM) images of a clay-stabilized oil-water emulsion. By changing the photon energy, different components of the emulsion become bright or dark, and can be identified from their known x-ray absorption properties. From Neuhausler *et al.* (1999), courtesy of Springer-Verlag.

Unfortunately, such focusing devices are less efficient than the glass lenses used in light optics. Also, laboratory x-ray sources are relatively weak (x-ray diffraction patterns are often recorded over many minutes or hours). This situation prevented the practical realization of an x-ray microscope until the development of an intense radiation source: the **synchrotron**, in which electrons circulate at high speed in vacuum within a **storage ring**. Guided around a circular path by strong electromagnets, their centripetal acceleration results in the emission of **bremsstrahlung x-rays**. Devices called undulators and wobblers can also be inserted into the ring; an array of magnets causes additional deviation of the electron from a straight-line path and produces a strong bremsstrahlung effect, as in Fig. 1-6. Synchrotron x-ray sources are large and expensive (> $100M) but their radiation has a variety of uses; several dozen have been constructed throughout the world during the past 20 years.

An important feature of the x-ray microscope is that it can be used to study hydrated (wet or frozen) specimens such as biological tissue or water/oil emulsions, surrounded by air or a water-vapor environment during the microscopy. In this case, x-rays in the wavelength range 2.3 to 4.4 nm are used (photon energy between 285 and 543 eV), the so-called **water window** in which hydrated specimens appear relatively transparent. Contrast in the x-ray image arises because different regions of the specimen absorb the x-rays to differing extents, as illustrated in Fig. 1-7. The resolution of these images, determined largely by zone-plate focusing, is typically 30 nm.

In contrast, the specimen in an *electron* microscope is usually in a dry state, surrounded by a high vacuum. Unless the specimen is cooled well below room temperature or enclosed in a special "environmental cell," any water quickly evaporates into the surroundings.

1.4 The Transmission Electron Microscope

Early in the 20th century, physicists discovered that material particles such as electrons possess a wavelike character. Inspired by Einstein's photon description of electromagnetic radiation, Louis de Broglie proposed that their wavelength is given by

$$\lambda = h/p = h/(mv) \tag{1.5}$$

where $h = 6.626 \times 10^{-34}$ Js is the Planck constant; p, m, and v represent the momentum, mass, and speed of the electron. For electrons emitted into vacuum from a heated filament and accelerated through a potential difference of 50 V, $v \approx 4.2 \times 10^6$ m/s and $\lambda \approx 0.17$ nm. Because this wavelength is comparable to atomic dimensions, such "slow" electrons are strongly diffracted from the regular array of atoms at the surface of a crystal, as first observed by Davisson and Germer (1927).

Raising the accelerating potential to 50 kV, the wavelength shrinks to about 5 pm (0.005 nm) and such higher-energy electrons can penetrate distances of several microns (μm) into a solid. If the solid is crystalline, the electrons are diffracted by atomic planes inside the material, as in the case of x-rays. It is therefore possible to form a **transmission electron diffraction** pattern from electrons that have passed *through* a thin specimen, as first demonstrated by G.P. Thomson (1927). Later it was realized that if these transmitted electrons could be focused, their very short wavelength would allow the specimen to be imaged with a spatial resolution much better than the light-optical microscope.

The focusing of electrons relies on the fact that, in addition to their wavelike character, they behave as negatively charged particles and are therefore deflected by electric or magnetic fields. This principle was used in cathode-ray tubes, TV display tubes, and computer screens. In fact, the first electron microscopes made use of technology already developed for radar applications of cathode-ray tubes. In a **transmission electron microscope** (TEM), electrons penetrate a *thin* specimen and are then imaged by appropriate lenses, in broad analogy with the *biological* light microscope (Fig. 1-4a).

Some of the first development work on electron lenses was done by Ernst Ruska in Berlin. By 1931 he had observed his first transmission image (magnification = 17) of a metal grid, using the two-lens microscope shown in Fig. 1-8. His electron lenses were short coils carrying a direct current, producing a magnetic field centered along the optic axis. By 1933, Ruska had added a third lens and obtained images of cotton fiber and aluminum foil with a resolution somewhat better than that of the light microscope.

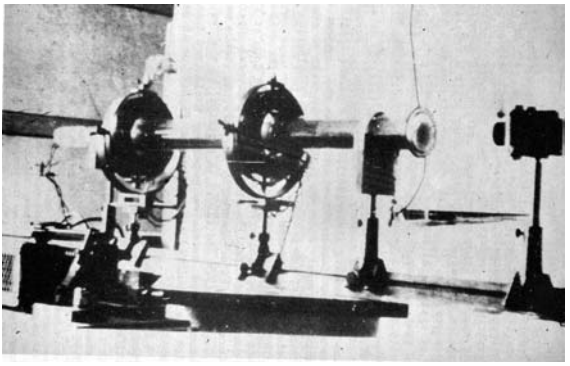

Figure 1-8. Early photograph of a horizontal two-stage electron microscope (Knoll and Ruska, 1932). This material is used by permission of Wiley-VCH, Berlin.

Similar microscopes were built by Marton and co-workers in Brussels, who by 1934 had produced the first images of nuclei within the interior of biological cells. These early TEMs used a horizontal sequence of lenses, as in Fig. 1-8, but such an arrangement was abandoned after it was realized that precise alignment of the lenses along the optic axis is critical to obtaining the best resolution. By stacking the lenses in a vertical column, good alignment can be maintained for a longer time; gravitational forces act *parallel* to the optic axis, making slow mechanical distortion (creep) less troublesome.

In 1936, the Metropolitan Vickers company embarked on commercial production of a TEM in the United Kingdom. However, the first regular production came from the Siemens Company in Germany; their 1938 prototype achieved a spatial resolution of 10 nm with an accelerating voltage of 80 kV; see Fig. 1-9.

Some early TEMs used a gas discharge as the source of electrons but this was soon replaced by a V-shaped filament made from tungsten wire, which emits electrons when heated in vacuum. The vacuum was generated by a mechanical pump together with a diffusion pump, often constructed out of glass and containing boiling mercury. The electrons were accelerated by applying a high voltage, generated by an electronic oscillator circuit and a step-up transformer. As the transistor had not been invented, the oscillator circuit used vacuum-tube electronics. In fact, vacuum tubes were used in high-voltage circuitry (including television receivers) until the 1980`s because they are less easily damaged by voltage spikes, which occur when there is high-voltage discharge (not uncommon at the time). Vacuum tubes were also used to control and stabilize the dc current applied to the electron lenses.