1 Introduction

1.1 New Lamps for Old …

Electromagnetic radiation is well-established as an important probe of matter, be it in the form of free atoms or molecules, crystalline solids or soft matter including biological materials. In the long wavelength range, notably from the far infra-red to the extreme ultra-violet and beyond, spectroscopic studies reveal many aspects of the energetics that govern physical and chemical properties. In the short wavelength X-ray range, scattering experiments provide detailed structural information on an atomic scale. All of these studies have been transformed by one or both of two key developments initiated in the second half of the twentieth century, namely the discovery and development of facilities to exploit synchrotron radiation, and the invention and development of the laser. These two new light sources are highly complementary in character, leading in large part to different areas of exploitation.

Synchrotron radiation provides a continuum of wavelengths from the far infrared to hard X-rays/γ-rays from ~10^6 Å (100 μm) to ~10^{-1} Å (10 pm), with associated photon energies from ~10 meV to ~100 keV (see Fig. 1.1, which also shows the approximate ranges of some alternative radiation sources). It has a very high spectral brightness (number of photons per unit area per unit solid angle) due to the intrinsically narrow beam of the emitted radiation. When combined with suitable monochromators this produces a tuneable radiation source over a vast wavelength or energy range. Although it is strictly pulsed in time, the repetition rate is so high and the pulse length is sufficiently long that for very many studies it can be considered as quasi-continuous in time.

Laser radiation is also in the form of a narrow bright source, but is typically provided at a fixed wavelength, although limited tunability (even of up to a factor of ~10 in some cases, but usually much less) can be achieved in some systems. The wavelengths offered by a selection of different lasers are largely restricted to the range from the infrared to the relatively near ultra-violet, although in special cases even 'table-top' lasers (albeit highly specialised ones) have achieved soft-X-ray energies of ~100 eV. Although some lasers have a continuous time structure, many are pulsed, and indeed very short pulses in the fs range have opened up a whole new area of exploration of fast time-resolved studies and, by virtue of the associated huge peak power output, render multi-photon phenomena accessible. One further important difference is coherence; lasers are sources of fully coherent radiation,

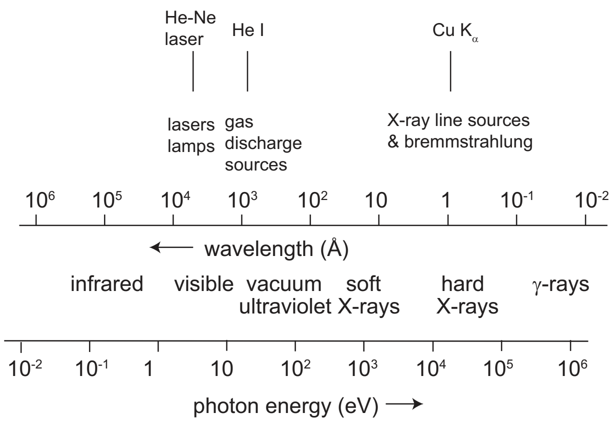

Figure 1.1 The electromagnetic spectrum covered by synchrotron radiation together with some laboratory sources. The range indicated for lasers corresponds to that of conventional 'off-the-shelf' table-top devices.

synchrotron radiation is not, although as discussed in Chapter 2 (and further in Chapter 8) certain types of experiments requiring coherence can be performed using synchrotron radiation. Of course, another key difference between lasers and synchrotron radiation is their physical scale and cost. Most conventional 'table-top' lasers are single-user devices located in the home laboratory of the experimentalist at a cost largely consistent with standard laboratory instruments. By contrast, synchrotron radiation is delivered to many (~20–30) simultaneous users in a central – typically national – facility with a capital cost as much as 10^5 or more times larger. This difference is almost certainly a major reason why the early development and exploitation of lasers was much more rapid than that of synchrotron radiation. The subsequent very significant developments in synchrotron radiation sources, firstly in creating dedicated user facilities, but subsequently in the creation of novel magnet structures that tailor the radiation output to the users' requirements, are described in Chapter 2.

The complementary nature of the two different types of sources in terms of the range of tunability and time structure, but also the readily accessible wavelengths, means that they have been mostly exploited in quite different experimental techniques. Recent developments in free-electron lasers (FELs) with abilities to produce tuneable coherent radiation with extremely high instantaneous power at energies up to the true X-ray range, do occupy a middle ground that somewhat blurs the distinction between these two types of sources in terms of their underlying mechanisms, but it seems likely that many applications of FELs and synchrotron radiation sources will continue to be rather distinct. In this book conventional lasers, based on a lasing medium, will not be discussed further, and the focus will be on synchrotron radiation. However, a brief discussion of FELs and their applications is included in view of some significant similarities in the structure of the sources and some overlap in their fields of exploitation.

1.2 Particle Accelerators: A Brief Historical Introduction

The motivation for the initial development of particle accelerators (and particularly proton, rather than electron, accelerators) has its origins in particle physics, investigating the consequences of high-energy collisions between fundamental particles. The objective was to accelerate the charged particles to high energies and then extract them from the accelerator and direct them to a target. Of course the simplest form of charged particle accelerator simply involves allowing the particles to gain energy in a 'single push' by passage across a large voltage difference. In this sense a simple electron gun, installed in cathode ray tubes (for many years the basis of televisions and oscilloscopes), is a particle accelerator, leading to electron energies of up to a few tens of keV. Much higher voltage single-push accelerators using Van de Graaf generators and Cockroft/Walton cascade generators can achieve energies of hundreds of keV or even a few MeV.

To achieve the highest energies, however, one must increase the energy in multiple sequential steps. Two essentially different designs to achieve this emerged, namely linear accelerators ('linacs') in which the charged particles gain energy while travelling along a linear trajectory, and somewhat more compact devices in which the charged particles perform circular (or approximately circular) orbits, constrained by magnetic fields. Conceptually the simplest versions of these devices, that rely on the charged particles gaining energy by passing through a succession of appropriately applied voltage differences, are primarily of interest for the acceleration of ions (particularly hydrogen ions - protons) rather than the much lower-mass electrons, for reasons that will become clear below. In the earliest type of linac shown in Fig. 1.2(a) the particles pass through a sequence of conducting drift tubes. Within the drift tubes the particles experience no electric field but if the applied voltages to the drift tubes are periodic in time, and are phased such that each time the particle emerges from one tube it 'sees' an attractive voltage on the next tube, then it is accelerated in each gap. Notice that as the speed of the particles increases the length of the drift tubes must increase to ensure the passage between the tubes is periodic in time. However, as the particle energy becomes relativistic this effect becomes less significant. In fact for electrons their speed becomes extremely close to that of light, c, at quite modest energies (e.g. an electron with an energy of only 10 MeV has a speed of ~0.999c, whereas a proton of the same energy has a speed of only ~0.14c).

In the earliest circular accelerator, the cyclotron (Fig. 1.2(b)), the particles travel in a split conducting cavity comprising two 'Dees' located between the pole pieces of a large fixed-field electromagnet; the magnetic field forces the ions to travel in a circular orbit while at a fixed energy. In this case, too, a periodically-varying voltage difference is applied to the two Dees, phased such that each time the particle passes from one Dee to the other it is accelerated. As the energy increases, the bending radius of the trajectory in the fixed magnetic field increases proportionately, so the overall trajectory follows a spiral until the particles are deflected out of the accelerator. For protons at non-relativistic energies the transit duration of each full orbit is constant (i.e. phase-locked), so the accelerating field frequency is also fixed, but relativistic

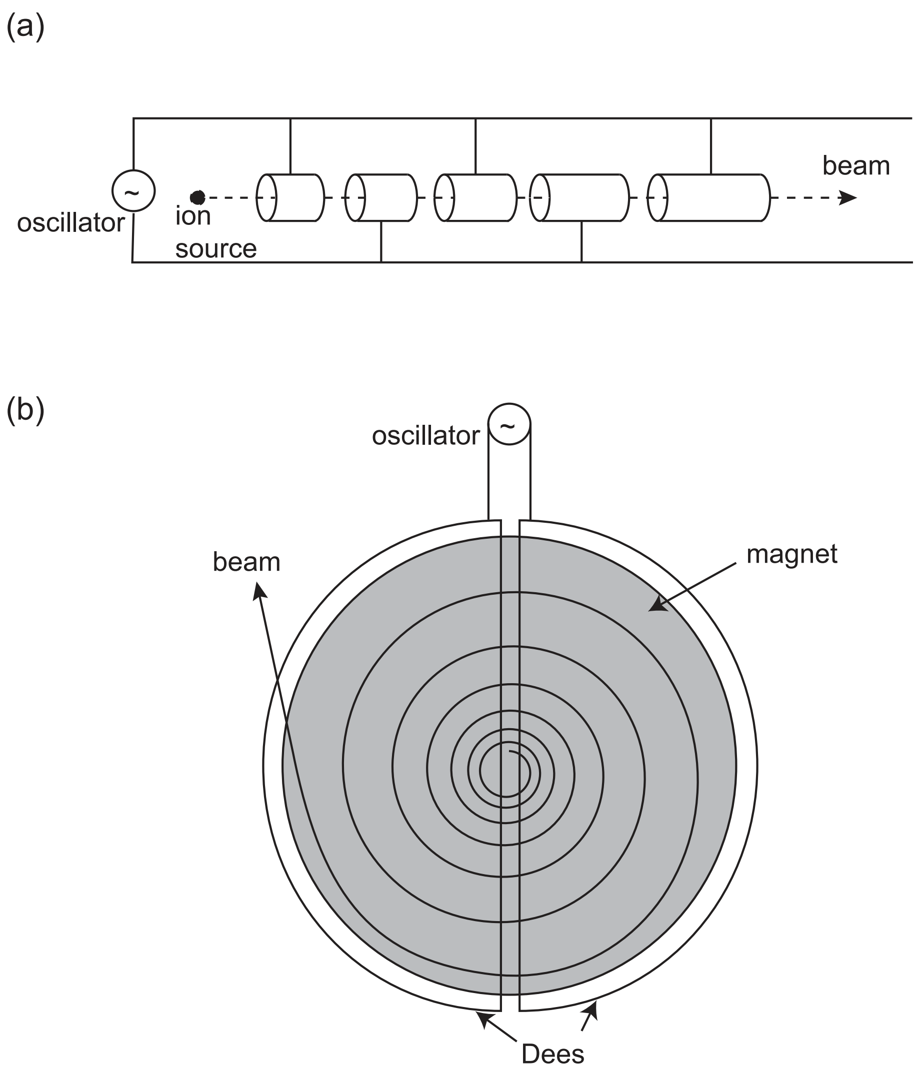

Figure 1.2 Basic features of two ion accelerators based on (a) linear and (b) circular geometries.

effects do limit performance eventually. While the cyclotron is far more compact than the linac, the energy to which the particles can be accelerated is also limited by the physical size of the device: specifically, the size of the magnetic pole pieces and the associated magnetic field strength.

Accelerators based on the circular motion of electrons are the devices of primary interest in the context of synchrotron radiation, but before focusing on this aspect it is appropriate to explain in slightly more detail how a linac is able to accelerate electrons at relativistic energies. Linacs are widely used to inject high (relativistic) energy electrons into the circular path of electron storage rings that are now the basis of synchrotron radiation sources, and are the essential ingredient of free electron lasers, both described in Chapter 2. Notice that the frequency of the oscillators in both linacs and circular accelerators, when operated with electrons, is in the (microwave) GHz range, with associated electromagnetic wavelengths of a few cms, and the way these radiofrequency (RF) waves propagate along a linac structure provides a rather different mode of acceleration structure to that described for ions in the context of Fig. 1.2(a). Specifically, in an electron linac the electrons 'surf' the RF travelling wave in a fashion somewhat similar to that of a surfer riding water waves. Fig. 1.3 shows in a simplified fashion how this can occur. A positive value of the longitudinal electric field E is defined here as one that accelerates the electrons in the direction of travel. The acceleration will be greatest if a travelling electron is located, relative to the travelling wave, near a peak in E but will experience less acceleration if it is located closer to the zero value of E, and indeed will be decelerated if it experiences a negative value of E. This leads to two consequences, namely a net acceleration of the 'surfing' electrons, but also a bunching of the electrons into groups located at similar relative positions on the travelling wave. Of course, in order for this to be effective, the travelling wave must match the average speed of the electrons.

As remarked above, for energies above a relatively low threshold value (at which the electrons are injected into this travelling wave), the electrons have a speed very close to c, and will gain energy mostly by gaining relativistic mass. Nevertheless, it is clear that the travelling wave must have a *phase velocity* that is close to, but slightly less than, c. The simplest situation would seem to be when the waves travel down a conducting tube, but in such a hollow tube waveguide, their phase velocity, $v_{ph} = \omega/k$ is actually greater than c, although their group velocity ($d\omega/dk$), the speed at which information can be transferred, is less than c and thus consistent with the requirements of relativity. In order to create a linear electron accelerator based on this travelling RF wave approach it is therefore necessary to modify the waveguide to slow down the phase velocity. This is achieved by 'loading' the waveguide cavity with a series of periodic obstacles in the form of discs with central apertures or 'irises' (Fig. 1.4).

Electron linacs based on this principle use a succession of these cavities, coupled to individual RF power sources, to achieve the multi-GeV energies required for injection into synchrotrons or electron storage rings (see below), and also for FELs. Notice that

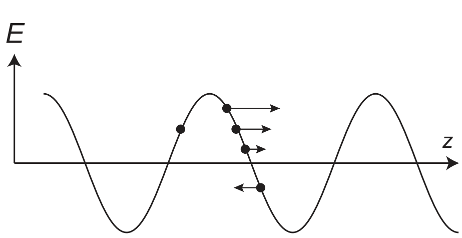

Figure 1.3 Schematic diagram showing the instantaneous force felt by electrons (represented by black discs) located at different positions relative to a travelling RF wave.

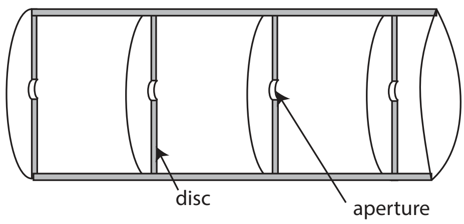

Figure 1.4 Schematic cutaway diagram of a disc-loaded cavity.

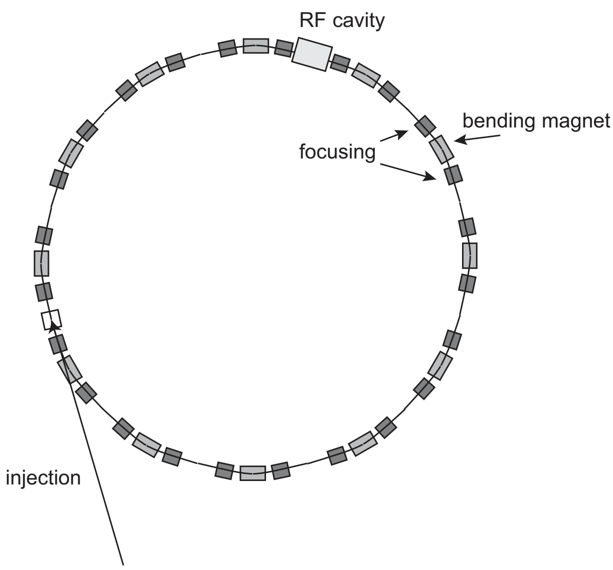

Figure 1.5 Schematic view of a synchrotron showing the main components in a simplified fashion.

essentially the same device is also used prior to high energy acceleration as a ‘buncher’ to produce electron bunches with the desired time structure and energy spread.

The fact that accelerated electrons quickly achieve highly relativistic speeds means that the simplest circular accelerator, the cyclotron, cannot maintain phase stability to high energies if operated at a fixed frequency. The synchrocyclotron, a development of essentially the same device but operated on single bunches of electrons at variable frequency, addressed this problem, but was still restricted in ultimate energy by the need for huge magnetic pole pieces to accommodate the spiralling trajectory of the accelerating electrons in the fixed magnetic field. The solution to these problems was the synchrotron, in which the electron trajectories are no longer strictly circular; the huge single magnet is replaced by a series of bending magnets located at the ‘corners’ of a polygonal racetrack (Fig. 1.5). The electron trajectory in such a device (through a very large but thin doughnut-shaped vacuum vessel) is nominally independent of energy as it increases; this fixed trajectory is maintained constant as the energy is increased by increasing the field strength of the bending magnets. Acceleration is achieved by installing one or more RF cavities in straight sections between the bending (dipole) magnets. Unlike the mechanism of the acceleration by travelling waves in a linac described above, in this case the RF cavity produces a standing wave, but the phasing of the time variation of the RF field with the arrival of an electron bunch is such as to achieve acceleration and maintain a near constant energy of electrons within the bunch. Additional magnetic structures (particularly quadrupoles – see Chapter 2) ensure the electrons all stay within a narrow path around the ring. In particular, the synchrotron exploits the idea of 'strong' or 'alternating gradient' focusing, in which the circulating beam is locally refocused at many points around the ring. It transpires that this technique allows the construction of much smaller aperture accelerators than earlier ones exploiting weak focusing. Initial injection of the electrons is from a lower-energy accelerator, most commonly nowadays a linac.

1.3 Synchrotron Radiation Basics

The key characteristics of synchrotron radiation are its extreme brightness or brilliance, its broad spectral continuum and its polarisation. In this section the basic physics leading to these properties is explained and quantified briefly.

Any acceleration or deceleration of a charged particle leads to the emission of electromagnetic radiation. It is this bremsstrahlung ('braking radiation') that provides the continuum background of X-radiation in a conventional laboratory X-ray source; energetic electrons (typically ~40 keV) strike a metallic target and emit bremsstrahlung as they lose energy in the solid. A rather special case of electromagnetic radiation from an accelerated charged particle occurs if the particle is forced into a circular orbit by a central force that constantly accelerates the particle along the direction of this central force. In this case the radiation emitted has the frequency of the circular motion and within the frame of reference of the charged particle one has dipolar emission; the angular dependence of the amplitude of the emission is given by the cosine of the angle between the instantaneous direction of travel and the direction of observation (see Fig. 1.6). It is this emission of radiation by an electron orbiting in a central Coulomb potential, and the resulting continuous energy loss, that leads to the downfall of a purely classical 'planetary' model of the hydrogen atom; the additional postulates of the Bohr theory were the first steps towards the proper quantum mechanical description. However, a quite different mechanism for producing circular motion of a charged particle is the force experienced by the moving particle due to the presence of a constant magnetic field perpendicular to the direction of motion. This is exactly the situation in a cyclotron or a synchrocyclotron as described in the previous section. It also occurs in astronomical plasmas of ionised gases and electrons in the proximity of sources of magnetic fields, the emission being known as cyclotron radiation. This magnetism-induced central-force acceleration is also a feature of the bending magnets used in a synchrotron to guide the electrons around the accelerator, although the nature

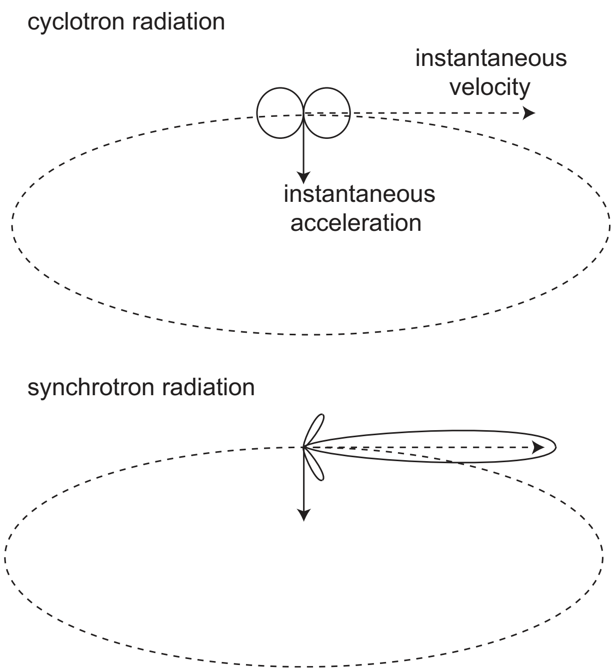

Figure 1.6 Schematic diagram showing the angular dependence of the radiation emitted from an electron travelling in a circular trajectory in a magnetic field at sub-relativistic energies (cyclotron radiation) and at relativistic energies (synchrotron radiation).

of the radiation emitted is determined by the bending radius in these magnets, rather than the much larger radius of the whole synchrotron. Nevertheless, this bending radius is ~10 m, so if the speed of the electrons is effectively the speed of light, c, (3×10^8 ms^{-1}) then the frequency of the cyclotron radiation would be ~10^7 Hz (in the radio frequency range) and of little interest as a photon source.

However, because the energy of the electrons is such that their speed is very close to c, there are very strong relativistic effects that modify the properties of the emission, which is then known as synchrotron radiation rather than cyclotron radiation. In describing these relativistic effects it is sometimes convenient to express the electron velocity v relative to the speed of light in a vacuum, c by the parameter $\beta = v/c$ although even more useful is the parameter, γ, defined as

$$\gamma = 1\Big/\sqrt{(1-\beta^2)} = E/m_0c^2, \tag{1.1}$$

where m_0 is the rest mass of the electron and E is the total energy, mc^2, where m is the relativistic mass of the electron. Notice that as m_0c^2 is 0.511 MeV, an approximate numerical expression for γ for an electron is simply $\gamma \approx 2{,}000\ E$ with E expressed in GeV.

One relativistic effect is the consequence of applying the Lorentz transformation to transfer the dipolar angular distribution of the radiation in the rest frame of the electron, into the laboratory frame (Fig. 1.6). The angular spread of the radiation is reduced to a total angle, $\Delta\theta$, with a value of approximately $2/\gamma$.[1] For example, with an electron energy of 5 GeV, this angular spread is only ~10^{-4} rad or 0.1 mrad (0.0006°). This is the origin of the highly directional character of synchrotron radiation and a major factor leading to its extreme brightness.

High electron energies also have a major impact on the amount of radiated power. In the non-relativistic case the total power radiated when an electron is accelerated was derived by Larmor at the end of the nineteenth century (Larmor, 1897) as

$$S = \frac{e^2}{6\pi\varepsilon_0 m_0 c^3}\left(\frac{\mathrm{d}p}{\mathrm{d}t}\right)^2, \tag{1.2}$$

where e is the electronic charge, ε_0 is the dielectric constant of the vacuum and $p = m_0 v$ is the electron momentum. At highly relativistic energies, however, one needs a Lorentz invariant form of Equation (1.2). The time transforms according to $\mathrm{d}t \to \mathrm{d}\tau = \frac{1}{\gamma}\mathrm{d}t$ while the momentum p must be replaced by the momentum 4-vector P_μ of relativistic electrodynamics

$$\left(\frac{\mathrm{d}P_\mu}{\mathrm{d}\tau}\right)^2 \to \left(\frac{\mathrm{d}p}{\mathrm{d}\tau}\right)^2 - c^2\left(\frac{\mathrm{d}E}{\mathrm{d}\tau}\right)^2. \tag{1.3}$$

In the case of an electron in a circular trajectory, with the acceleration directed to the centre of the circle, the energy of the electron is constant, so the second term in Equation (1.3) is zero. The total power radiated then becomes

$$S = \frac{e^2 c}{6\pi\varepsilon_0 (m_0 c^2)^2}\left(\frac{\mathrm{d}p}{\mathrm{d}\tau}\right)^2 = \frac{e^2 c\gamma^2}{6\pi\varepsilon_0 (m_0 c^2)^2}\left(\frac{\mathrm{d}p}{\mathrm{d}t}\right)^2, \tag{1.4}$$

while for circular motion one can also write

$$\frac{\mathrm{d}p}{\mathrm{d}t} = p\omega = p\frac{v}{R}, \tag{1.5}$$

where R is the bending radius. Moreover, for large values of γ, $E = pc$, so

$$S = \frac{e^2 c}{6\pi\varepsilon_0 (m_0 c^2)^4}\frac{E^4}{R^2}, \tag{1.6}$$

a result first obtained by Liénard (1898). Of course, Liénard's work predated the theory of relativity, but the equations of electromagnetic theory are invariant under transformations between moving systems so his original equation remains valid at relativistic energies This equation highlights two key results. Firstly, the radiated energy scales as E^4, so while the radiated energy is very small in a low

[1] In some presentations the whole angular spread is defined as $1/\gamma$. In fact the exact value of this angle depends on the wavelength of the radiation emitted but also on how the angle is defined, for example as the Gaussian width σ or, indeed, 2σ (see Equation (1.18) and the associated discussion). Here $1/\gamma$ is taken to be the half-angle. The arguments based on either definition are necessarily approximate.

energy cyclotron, it can be very significant at the higher energies found in a synchrotron. Indeed, in a synchrotron one can show from Equation (1.6) that the energy loss of an electron passing just once around the synchrotron is, in keV, equal to 88.5 E^4/R with E expressed in GeV and R in m. For example, if E is 3 GeV and R is 10 m the energy loss by a single electron in each circuit around the ring is 717 keV. The second significant result contained in Equation (1.6) is that the radiated power scales as $1/m_0{}^4$. This clearly favours a low mass charged particle as the optimum source of synchrotron radiation. Specifically, as the mass of a proton is 1,836 times that of an electron, the radiated power from a proton at the same total energy as an electron is 1.13×10^{13} times smaller. Machines designed to deliver synchrotron radiation to users have thus always used accelerated electrons or positrons.

The second key feature of synchrotron radiation is its frequency spectrum. To understand this property one must note that because the emission from electrons orbiting with a frequency ω_0 has a very narrow angular spread of only $\Delta\theta \approx 2/\gamma$, the radiation illuminates a detector ('observer') for a duration of only $\Delta t \approx \Delta\theta/\omega_0 = 2/\gamma\omega_0$, so the Fourier spectrum of this pulse of radiation must contain frequencies of at least ~$1/\Delta t$ or $\gamma w_0/2$. We must, however, also take proper account of relativistic time dilation and the Doppler effect. A particularly simple way of taking account of both of these effects is to consider the trajectory of the electron and the emitted radiation within the time period during which an observer is illuminated (Fig. 1.7). The time interval between the two ends of the detected electromagnetic pulse is just the difference in the time of flight of the electron, speed v, and the photon, speed c, as the electron travels from A to B

$$\Delta t = t_{el} - t_{ph} = \frac{R\Delta\theta}{v} - \frac{R\sin\theta}{c} = \frac{R}{c}\left(\frac{\Delta\theta}{\beta} - \Delta\theta + \frac{\Delta\theta^3}{3!}\ldots.\right) \tag{1.7}$$

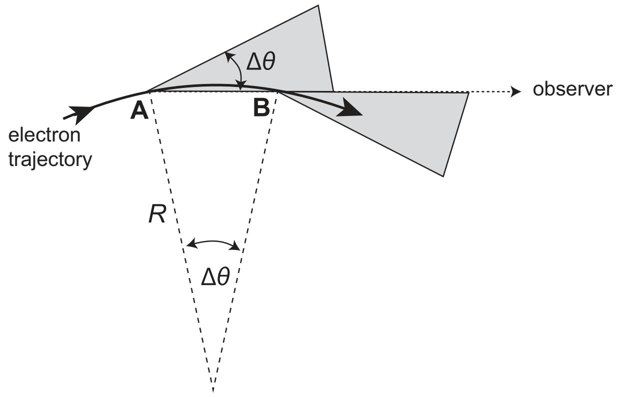

Figure 1.7 Schematic figure showing the range of the electron trajectory in a circular orbit from **A** to **B** during which a fixed observer can observe the (shaded) fan of synchrotron radiation with an angular width of $\Delta\theta$.