FIELDS AND WAVES
2002

10.0 ELECTROMAGNETIC WAVES

I will start this section with a quotation from *An introduction to the Meaning and Structure of Physics* by Leon N Cooper (Harper & Row, 1968):

> "Regarding electricity and magnetism, the situation in 1830 after Faraday's work was as follows:
>
> 1. Electric charges produce forces on each other that can be described by Coulomb's law and by the introduction of the electric field.
>
> 2. Wires carrying currents produce forces on each other that can be described using Ampere's law and by the introduction of magnetic fields.
>
> 3. No magnetic charges exist.
>
> 4. Changing magnetic fields produce electric fields - Faraday's law.
>
> 5. Electric charge is conserved: The total amount of charge in any region of space remains constant unless charges cross in or out of that region.
>
> It took about two generations to realise that these five statements are logically contradictory.
>
> [...] Maxwell tracked down the contradiction among the postulates of electromagnetism to Ampere's law. If Ampere's law was exact in the form it had been stated, then it was not consistent with the conservation of charge. This law relates the magnetic field produced solely to the current that flows, which in itself, if we were directed to it in the right way, might seem peculiar; electric fields can be produced both by charges and, according to Faraday's law, by changing magnetic fields. With a passion for symmetry we might then believe that magnetic fields could be produced both by currents and by changing electric fields. It was just by adding this possibility to Ampere's law that Maxwell was able to make it consistent with the conservation of electricity."

Maxwell's equations (1856 – 1862) summarise in mathematical form the content of the laws of electricity and magnetism. They had an unexpected but spectacular consequence. A changing electric field could produce a changing magnetic field, and a changing magnetic field could produce a changing electric field; and it followed that these changes could propagate as electromagnetic waves. And furthermore, the speed with which these waves propagate could be calculated, with the result that he found that they travel at the speed of light. It was then an inevitable inference that light was in fact a form of electromagnetic radiation.

The spectrum of electromagnetic radiation extends in wavelength both above and below the range of visible light (approximately 400 nm for violet to 720 nm for red). Electromagnetic radiation with wavelengths longer than the red light of the visible spectrum range from infrared (microns to millimetres) through microwaves (centimetres) to the wavelengths used for radar, television and radio and beyond. At the other extreme, we find ultraviolet radiation (down to around 1 nm), x-rays (down to about 10^{-11} m through to gamma rays (which are observed with wavelengths as short as 10^{-15} m). All these forms of electromagnetic radiation share much in common; they differ primarily through there methods of generation and detection. They all travel through empty space with the same speed,

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} = 10^{-8} \text{ m/s}.$$

10.1 RADIO AND TV

It was not until 1887 that Heinrich Hertz first verified Maxwell's prediction that electromagnetic waves could be generated by rapidly changing electric fields – in fact by sparks jumping across a gap between electrodes at a high voltage difference. Hertz detected the waves by showing that they produced sparks across a gap in a small copper loop. By demonstrating that the radiation produced by the sparks could be reflected and focused by a concave sheet of metal he went on to establish that he could establish a standing wave and so determine the wavelength of the radiation he had generated. Only a few years later Guglielmo Marconi began his pioneering study of radio waves as these became to be called. By 1895 he had succeeded in transmitting a signal over a distance of around 2 km, and in 1901 a signal was transmitted (using Morse code) across the Atlantic, with a regular service established the following year and a daily news service to trans-Atlantic liners in 1904.

It is clear that for such long range transmission the radio waves cannot travel directly from the transmitter to the receiver; the curvature of the Earth means that "line-of-sight" transmission is impossible. In fact at the wavelengths used by Marconi (of the order of around 30 m, so with a frequency around 10 MHz), the ionosphere, a layer of ionised gas in the upper atmosphere, acts to reflect the waves, and so allow for long range transmission. The short wave bands (high frequency – HF) used for long distance radio are in the range 20 – 60 m in wavelength. The reflected waves from the ionosphere provide what is called the "sky wave" signal. For frequencies above 30 MHz, the ionosphere ceases to be reflecting, and such high frequency waves pass out into space, and so their transmission is restricted to short range direct transmission (through the so- called "space wave"). These higher frequency waves (VHF and UHF) are used for transmission of hifi and local radio, mobile phone communication, and TV. On the other hand at frequencies much lower than around 50 MHz, transmission is possible through what is called the "ground wave", in which the radio waves travel close to the surface (essentially guided as in a wave-guide). This is of importance for the transmission of MW and LW radio – medium wave (around 300 m wavelength), and long wave (around 1500 m) respectively.

Radio waves carry a signal through being modulated. Usually the transmissions on long, medium and short wavelengths are amplitude modulated (AM), which means that the signal varies in amplitude, the frequency of the variation being in the audio range of 20–20,000 Hz, and so can be used to "carry" a sound-wave signal which can be extracted at the receiver. The VHF and UHF waves used for high-fidelity and TV broadcasts which need a more efficient use of the bandwidth available use FM or frequency modulation . For this, the amplitude is held steady, but the frequency is altered around the central frequency; the changes in the frequency then carry the signal. Analogue signals (in which the variation is continuous in time, but with a time-scale related to the audio frequencies to be broadcast or the much higher frequencies relevant for signalling television data) are fast giving way to digital signals (essentially a sequence of on-off transmissions) which can then be processed using error-correction techniques to yield much higher fidelity reception with minimal interference. Transmission and reception can be from a dipole antenna, or from an array which is designed to improve directionality, but may also be enhanced by using a reflector, often in the form of a parabolic dish. This is familiar of course for the dishes used to receive the signals transmitted from satellites – and also from astronomic sources.

10.2 POLARISATION

Electromagnetic waves are transverse. They are electric and magnetic fields which oscillate in space and time, with the direction of these fields perpendicular to the direction of propagation of the waves. The electric and magnetic fields are themselves perpendicular to one another. Because they are transverse waves, electromagnetic waves can be *polarized.* Plane polarised waves have of their electric field always in the same direction - and this is what is called the direction of polarisation of the wave. Some materials can be fabricated to act as polarisation analysers; they transmit light polarised in one direction, but block light polarised in the direction at right angles to this. They can then be used as polarising filters or analysers.

You may be familiar with the way that polarising sunglasses cut down the dazzle and glare from sunlight reflected from the surface of water. This is because the light reflected in this way is polarised, with the direction of polarisation being horizontal. So the polarising glass in the sunglasses is set in such a way that horizontally polarised light is blocked, whilst transmitting vertically polarised light. The result is that unpolarised light (light which contains a mixture of equal amplitudes of all directions of polarisation) will have its intensity reduced by a factor of two, but the glare from reflected light will be diminished even further. To understand more fully the way that light is polarised when reflected from the water surface, we consider what happens quite generally when light is incident on an interface between air (or a vacuum) and a transparent medium, such as water. As you know, the light is refracted at the interface, and the familiar Snell's Law gives the relationship between the angles of incidence and reflection:

$$\sin i / \sin r = n.$$

There is also reflection at the surface, with the reflected ray making the same angle with the normal to the surface as does the incident ray. But this partially reflected ray is polarised (and so also is the transmitted, refracted ray). The extent of polarisation depends on the angle of incidence. It becomes 100% when the angle of incidence is a critical angle called the Brewster angle. This angle is such that the transmitted, refracted, ray and the reflected ray are at right angles to one another. For this to be true,

$$\tan i = n.$$

Notes prepared by John M Charap on 18/03/02, revised 20/03/04