CHAPTER XIV. PROPAGATION OF ELECTRO-MAGNETIC WAVES

INTRODUCTORY

Frequency range of electro-magnetic waves

1. This chapter, dealing with the propagation of radio waves, may be divided into two portions, the first containing an outline of elementary theory leading to a brief and approximate treatment of the influence of ions and electrons upon the passage of an electro-magnetic wave, while the second gives a factual summary of the propagation of waves of different frequencies. In Chapter VII it was explained that when electrons are accelerated an electro-magnetic wave is produced. This wave consists of a sinusoidally varying electro-magnetic flux travelling with a velocity of approximately $\frac{3 \times 10^{10}}{\sqrt{\mu \kappa}}$ centimetres per second, its velocity thus depending upon the magnetic permeability and the dielectric constant of the medium. Reference was also made to the reception of light and heat energy from the sun, and it was stated that this energy is conveyed in the form of electro-magnetic waves. Heat and light radiation is therefore of the same nature as that used for wireless communication, differing only in frequency ; the whole spectrum of frequencies occupied by known forms of electro-magnetic radiation ranges from 10^{21} kc/s downwards. The known range of frequencies is shown in fig. 1, the frequencies being divided into bands according to the mechanism of production ; certain of these bands overlap because the particular frequencies can be generated by two different methods.

2. The highest frequencies so far detected are the so-called cosmic rays, which appear to be produced far out in space, and reach the earth from all directions. They are capable of penetrating very deeply into metallic bodies, although the latter are practically perfect reflectors for waves of lower frequencies. The next broad band embraces the gamma-rays emitted by radio-active substances such as uranium, and X-rays, which are produced by the bombardment of a metal plate by electrons moving with extremely high velocity. The X-ray tube is in fact nothing more than a special form of high-vacuum diode. The lower-frequency X-ray band overlaps the band embraced by ultra-violet light, which is emitted by bodies at extremely high temperatures. Both X-rays and ultra-violet rays affect a photographic plate, but do not give the sensation of vision. They are of medical value, but are capable of inflicting severe injury upon the delicate structure of the human eye.

3. The next band—the spectrum of visible light—is that with which we are most familiar When the whole visible spectrum is present, as in the radiation from the sun, the various frequencies combine in their action upon the eye and nervous system to give the sensation which we identify as white light. The lowest frequencies in this band give the sensation of red light, and the adjacent band, which embraces the radiation emitted by hot bodies at a temperature just below incandescence, is known as infra-red radiation. These rays are of course also emitted by incandescent bodies, e.g., of the total radiation from an electric lamp filament, about 99 per cent. is infra-red and only 1 per cent. visible light. It may be noted that the whole of the frequency range between about 10^{19} and 10^{9} kc/s can be produced in an indirect manner by electrical means. Radiation of frequency below about 10^{9} kc/s can be directly produced only by electrical oscillators, and can be detected only by electrical methods.

4. (i) Before dealing with the range of frequencies which for want of a better term we may call the radio-communication band, it may assist the correlation of ideas if it is pointed out that in general, we suppose all the forms of indirectly produced radiation to be initiated by large numbers of minute oscillators somewhat resembling miniature hertzian doublets (Chap. VII). Little is known about the actual mechanism of production of this kind of radiation, but a very rough idea of the production of, say, an oscillation in the ultra-violet region may be obtained

as follows. Consider a single hydrogen atom, the radius of which is about 10^{-8} centimetres. Suppose that owing to the sudden impact of a free electron, the electron belonging to the hydrogen atom is disturbed in its orbit and executes a to-and-fro vibration, superimposed upon its normal motion. We then have two equal electric charges of opposite sign, alternately appearing at either end of an element of space about 10^{-8} centimetres in length. This may be compared with the oscillation of the charge on the plates of the hertzian doublet, and the disturbed atom may, on this hypothesis, be regarded as an elementary half-wave aerial. The wavelength will then be 2×10^{-8} centimetres, and the frequency $\frac{3 \times 10^{10}}{2 \times 10^{-8}} = 1{\cdot}5 \times 10^{18}$ cycles per second or $1{\cdot}5 \times 10^{12}$ Mc/s ; thus the disturbance of an atom in the manner indicated would give rise to a momentary emission of ultra-violet light. It will of course be appreciated that the actual mechanism of emission is far more complicated than is here implied.

(ii) The frequencies actually used in radio-communication are also shown in fig. 1. As already stated, they cannot be detected by the unaided human senses. At the present time, the practical upper and lower limits may be taken as 3×10^{7} kc/s and 10 kc/s respectively. This band is sub-divided as follows:—

Below 30 kc/s	Very low frequencies (VL/F)
30–300 kc/s	Low frequencies (L/F)
300–3,000 kc/s	Medium frequencies (M/F)
3–30 Mc/s	High frequencies (H/F)
30–300 Mc/s	Very high frequencies (VH/F)
300–3,000 Mc/s	Decimetre waves (dc/W)
3,000–30,000 Mc/s	Centimetre waves (cm/W)

The present trend of development is towards the gradual abandonment of frequencies below about 100 kc/s, and exploration of the utility of the 300 to 30,000 Mc/s band. In the diagram, the radio-communication band has been extended to 10^{9} kc/s, the latter being the highest frequency produced in the laboratory by direct electrical methods. There is, however, no present prospect of the practical production and employment of such high frequencies.

ELEMENTARY PROPAGATION THEORY

5. Since light, heat and radio waves are all of the same nature, it is to be expected that the same general laws govern their propagation through space, but the presence of material molecules, or even of free electrons, in the region through which propagation takes place, has a considerable effect upon the wave, its exact nature depending to a great extent upon the frequency. In order to explain the general laws, it is convenient first to consider light waves, because they are easily produced and are perceptible to the human senses without the aid of extraneous apparatus, so that many of the laws are easily verified by simple experiments. A ray of light is defined as the path along which a wave travels, and a collection of light rays is called a beam. It is a matter of every-day observation that light travels in straight paths; for instance, if a small hole is made in the blind of a darkened room, the light entering by this hole will illuminate the dust particles in the atmosphere and the light reflected from these enters the eye. It is sometimes then said that a beam of light is seen, but actually the only light seen is that reflected from the particles, and these are observed to lie in straight lines. A natural consequence of the rectilinear propagation of light is the formation of shadows, because the presence of an obstacle in the path of the beam may prevent the illumination of particles on the side of the obstacle remote from the source of light. Bodies are said to be transparent, translucent or opaque according to the degree to which they allow the light to pass through them.

Definitions

6. When light comes into collision with a material body three effects may be observed.

(i) A portion of the light is reflected at (or near) the surface of the body.

(ii) A second portion may travel through the body, the rays being usually bent on passing through the boundary surface. The rays are then said to be refracted.

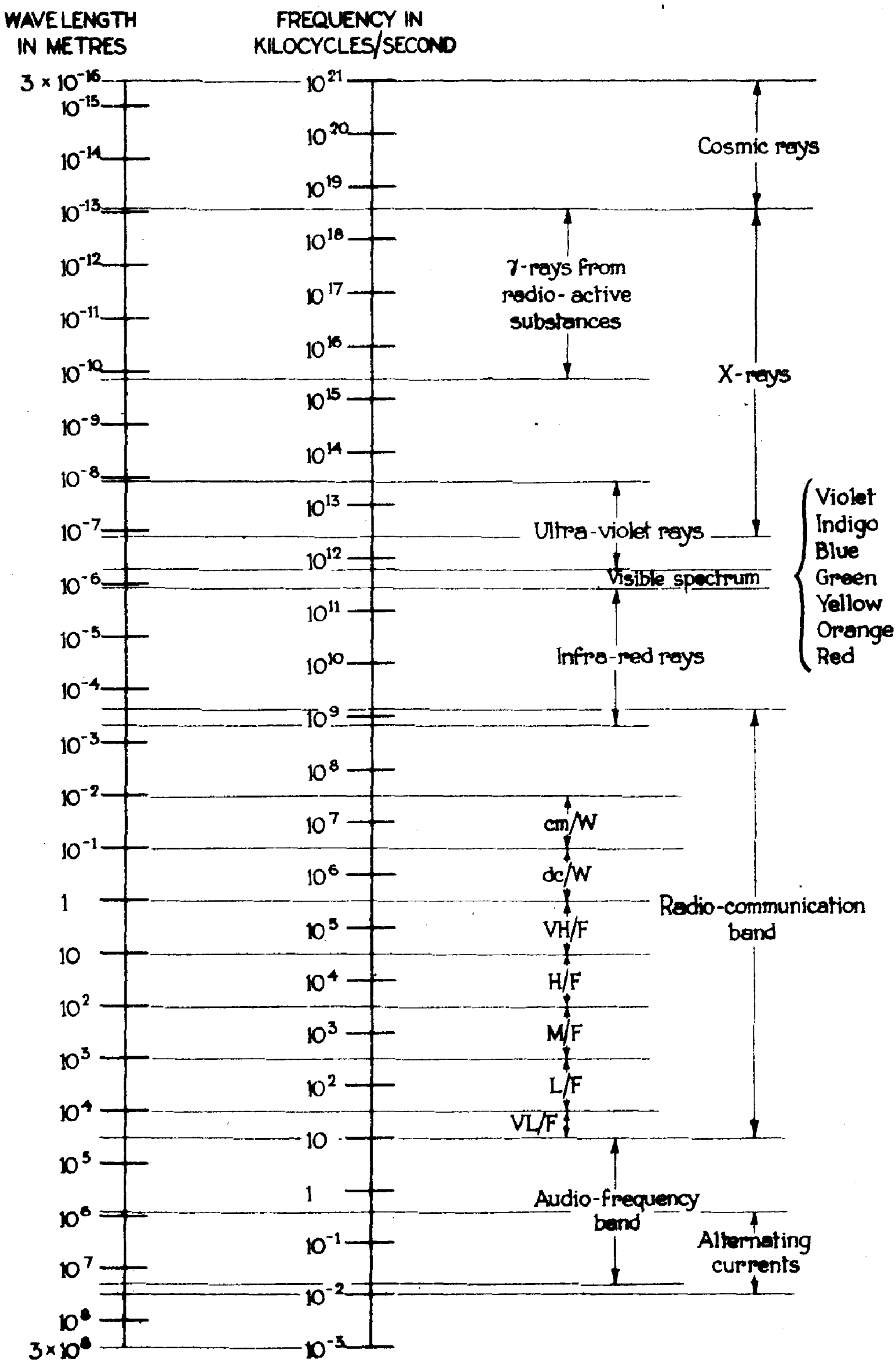


FIG. 1
CHAP. XIV

ELECTRO-MAGNETIC WAVES IN FREE SPACE

This Page Intentionally Blank

(iii) A portion of the refracted light is absorbed during its passage through the body.

In speaking of these phenomena, it is necessary to define

(*a*) The normal to the surface of the body. This is a line drawn through and perpendicular to the surface, at the point of incidence of a light ray.

(*b*) The incident ray, which is the path in which light travels before reaching the surface.

(*c*) The reflected ray, which is the path taken by the light after it has collided with the surface.

(*d*) The angle of incidence, or angle which the incident ray makes with the normal.

(*e*) The angle of reflection, or angle which the reflected ray makes with the normal.

(*f*) The angle of refraction, or angle which the refracted ray makes with the normal.

Reflection

7. The laws of reflection are as follows :—

(i) The angle of reflection is equal to the angle of incidence.

(ii) The incident ray, the reflected ray and the normal through the point of incidence all lie in the same plane.

These laws are illustrated in fig. 2.

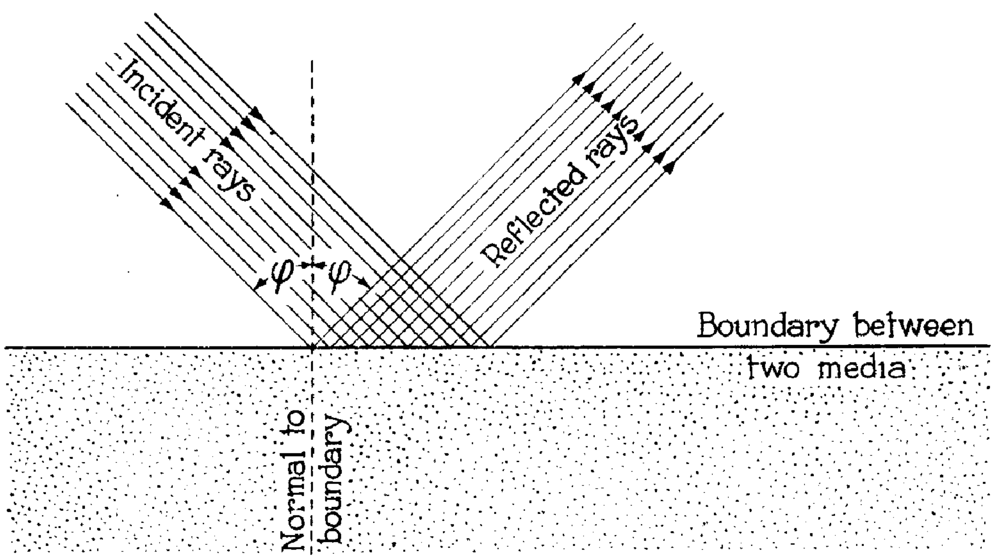


FIG. 2, CHAP. XIV.—Reflection of electro-magnetic wave.

Diffraction

8. In the early history of the wave theory of light, it was strongly opposed by Newton, because it appears that light travels in straight lines, whereas sound, which had previously been proved to be a wave motion in a material medium, was found to bend round obstacles. The scientist Fresnel appears to have been the first to point out that a light ray does in fact bend round an obstacle just as a sound wave does, the difference being merely one of degree. The bending of a wave round the edges of an obstacle is referred to as diffraction. It may be observed by allowing light to pass through a very small aperture and to fall on a screen. If an obstacle is interposed so as to cast a shadow upon the screen, the edges of the shadow are never perfectly defined. Close observation reveals that several alternately light and dark bands appear outside the shadow, and these get nearer to each other as they recede from the edge of the shadow, until they merge into the area of uniform illumination. Another example of diffraction may be

observed by viewing a distant source of light, such as a street lamp, through a mist or light fog, when the lamp itself appears to be surrounded by rings of coloured light. This is due to the bending of the light rays round the particles of water vapour, and it is found that the lower frequency (red) rays are bent to a greater degree than those of higher frequency. It is in fact capable of demonstration that the lower the frequency, the greater the degree of bending. Diffraction phenomena are therefore of considerable importance at radio-communication frequencies.

Refraction

9. The term refraction refers to the bending which takes place when a ray passes from one transparent medium to another. The effect is easily shown by the partial immersion of a stick in water, when the stick appears to be bent at the point where it meets the surface. The relation between the angle of incidence and the angle of refraction is known as Snell's Law, and is as follows :—

> When a ray of light is incident upon the surface of separation of two media, it is bent in such a manner that the ratio $\frac{\text{sine of the angle of incidence}}{\text{sine of the angle of refraction}}$ is a constant for all angles of incidence. This constant is called the refractive index of the two media, and is denoted by the symbol ν. In optics the symbol μ is usually used for the refractive index, but has the disadvantage of confusion with the magnetic permeability of the medium.

Physical meaning of ν

10. The refractive index has a real physical meaning apart from the geometrical one just described, for it depends upon the velocity of light in the two media. It is now convenient to refer to the wave-front of a wave, which is defined as an imaginary surface perpendicular to the direction of propagation of the wave. If the light is spreading outwards from a point source, the wave front at any point in space is a sphere, but if only a very small portion of the wave front is taken, at a distance of many wavelengths from the source, the wave front may be considered to be a plane surface. In fig. 3, let a, b and c be three rays forming part of a beam of

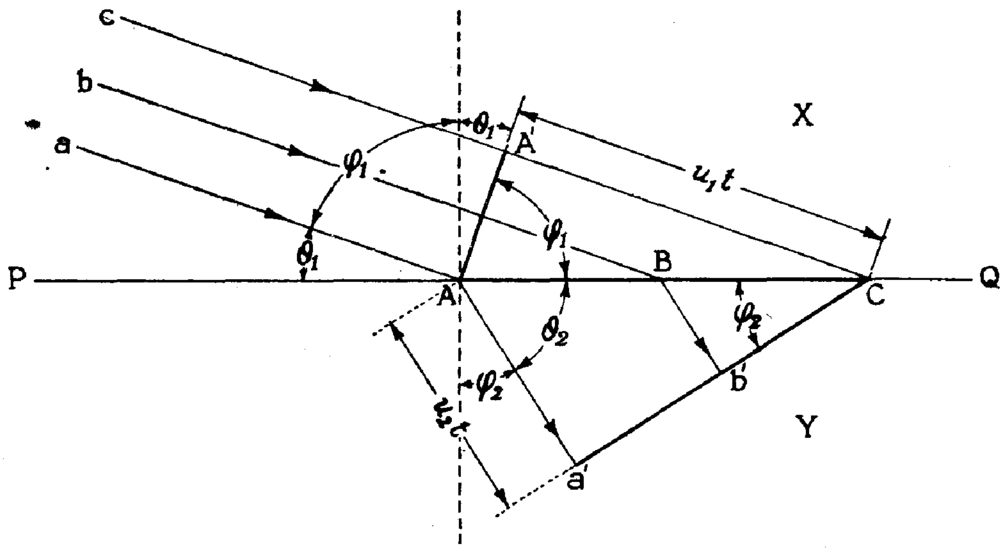


FIG. 3, CHAP. XIV.—Refraction of electro-magnetic wave.

light which is travelling with velocity u_1 in a medium X towards the surface P Q ; at the instant when the light travelling along the ray a reaches the surface at A the wave front is A A'. The light travelling along the ray c must therefore pass through a distance A' C before reaching the surface P Q at the point C. On reaching the surface at the point A the light of ray a enters the medium Y, in which its velocity is u_2. The light carried by ray c, travelling with velocity u_1, traverses the distance $A^1 C$ in a time $t = \frac{A'C}{u_1}$. The light travelling along the ray a, in the medium Y, will in this time move through a distance Aa′ where $Aa' = u_2 t$, or

$$A a' = u_2 \times \frac{A'C}{u_1}.$$

Hence
$$\frac{A'C}{Aa'} = \frac{u_1}{u_2}.$$

The angle of incidence is φ_1 and the angle of refraction φ_2, by the definitions given above. Since in the diagram $\varphi_1 + \theta_1 = 90°$, $\varphi_2 + \theta_2 = 90°$, it is seen that

$$\frac{A'C}{AC} = \sin \varphi_1$$

$$\frac{Aa'}{AC} = \sin \varphi_2$$

$$\frac{A'C}{Aa'} = \frac{\sin \varphi_1}{\sin \varphi_2} = \nu, \text{ by definition.}$$

It follows therefore that

$$\nu = \frac{u_1}{u_2}.$$

Hence the refractive index of any two media is equal to the ratio of the velocities in the respective media. If u_1 is greater than u_2, as in fig. 3, the refracted ray is bent towards the normal, while if u_2 is greater than u_1 it is bent in the opposite direction.

11. Although light waves have been used for illustrative purposes, the above reasoning is applicable to any form of electro-magnetic wave. It has already been stated that in a medium of permeability μ and permittivity $\varkappa$ the velocity of electro-magnetic waves is $\frac{c}{\sqrt{\mu \varkappa}}$ centimetres per second, where c is the velocity in free space. In the case of radio waves, we are not concerned with their propagation through any medium having a permeability differing appreciably from unity, and therefore it appears that the velocity, u, should be equal to $\frac{c}{\sqrt{\varkappa}}$ centimetres per second, and $\frac{c}{u} = \sqrt{\varkappa}$. But $\frac{c}{u}$ is by definition the refractive index ν, and thus the refractive index of a material should be equal to the square root of its permittivity.

12. As will be shown later, the permittivity of a given material is not truly a constant, but depends to some extent upon the frequency; consequently the velocity of electro-magnetic waves, except in free space, also depends upon the frequency. White light is composed of a mixture of waves of different frequencies, and if a narrow beam of white light, that is, sunlight, is allowed to fall obliquely upon a thick sheet of plate glass, it will be found that the light is resolved into seven different colours: red, orange, yellow, green, blue, indigo, and violet. These represent groups of different frequencies, of which the lowest (red) suffers least, and the highest (violet), the greatest refraction. If it is proposed to attempt to verify the above deduction that the refractive index of a material is equal to the square root of its electrical permittivity, therefore, it is essential that both the optical and the electrical measurement should be made at the same frequency—a matter of considerable difficulty. A further complication arises, in that no substance is a perfect insulant, and the conductivity must be taken into account in making the electrical measurement. Actually the refractive index is dependent upon the degree of ionization of the medium of propagation, as will be seen later. A medium in which the velocity varies with the frequency is said to be dispersive.

Phase and group velocity

13. In Chapter VII it was pointed out that no intelligence can be conveyed by a wave unless it is modulated in some manner. It has also been shown that a modulated wave may be regarded as the sum of a number of different components having various amplitudes and frequencies. In a dispersive medium, components of different frequencies travel with different velocities, and as a result the signal itself travels at a speed differing from that with which an unmodulated wave would be propagated in the same medium. To explain this, let us consider a transmission consisting only of two frequencies of the same amplitude. The two waves combine to form a composite group of waves, resembling in form the heterodyne beat discussed in earlier chapters, but the group itself will travel through space. Its velocity, however, will not be that of the individual waves and it is necessary to distinguish between the phase velocity and the group velocity in the particular medium. The phase velocity is that with which a single frequency of constant amplitude would be propagated, while the group velocity is that with which a signal is propagated.

14. The effect is illustrated in fig. 4, which represents the position at two successive moments

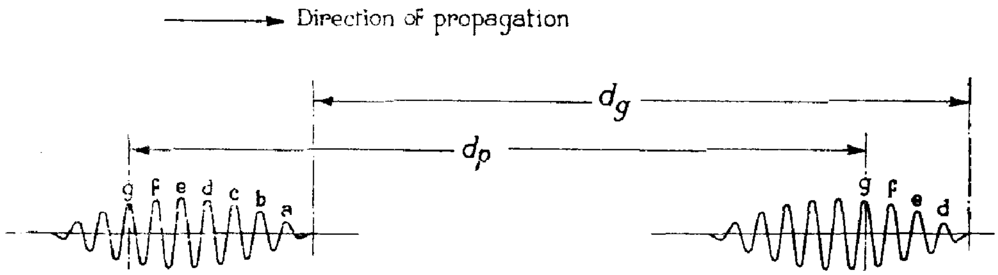


FIG. 4, CHAP. XIV.—Phase and group velocity.

of a group of electro-magnetic waves, which is moving through a cloud of electrons ; the distance travelled by the whole group during the interval is d_g and this distance, divided by the duration of the interval, is the group velocity, u_g. On the other hand we may fix our attention upon any given crest, say that marked g, and trace its progress. As the wave moves forward through the medium the vanguard a, b, c, is robbed of its energy in setting the electrons into oscillation, and this energy, less that which is dissipated by electronic collision, tends to maintain the waves as they die away. There is thus a continual eating away of the head and building up of the tail of the group of waves, with the result that the crest g gradually moves forward towards the head of the group ; the distance d_p through which this crest has travelled during the interval, divided by the interval, is the phase velocity, u_g. The phase velocity is obviously greater than the group velocity and it can be shown that, whereas u_p is greater than the velocity c of light in a vacuum, u_g is smaller than c. The phase and group velocities are in fact related by the equation $u_p\, u_g = c^2$.

Polarization

15. It has already been stated (Chap. VII) that electro-magnetic waves consist of a transverse vibration in space, the harmonically varying quantities being the electric and magnetic fields, which are perpendicular to each other and to the direction of propagation. The orientation of these fields is called the polarization of the wave. In the case of light radiation, e.g., by an incandescent solid, every molecule of which may be regarded as a rudimentary hertzian doublet, the vibrations take place in every conceivable plane and no definite polarization can be observed. During passage through certain media, for example crystalline quartz and Iceland spar, vibrations in other than particular planes may be suppressed, and on emergence, the wave is said to be plane-polarized. The wave emitted by a straight radiating conductor is polarized in such a manner that its electric field vector lies in a plane passing through the longitudinal axis of the conductor (Chap. VII). If the medium of propagation is free space, this state of polarization will remain unchanged. The presence of material molecules, or even of free electrons, may, however, affect the polarization of the wave during its passage.

16. In radio-communication practice, it is usual to define the polarization with reference to the surface of the earth. When the electric field vector lies in a plane perpendicular to the ground, the wave is said to be vertically polarized. It does not follow that its electric field vector is perpendicular to the ground. For example, if a wave reaches the earth from an aeroplane with an angle of incidence of 45°, and its magnetic field vector is horizontal, the electric field vector is perpendicular to the magnetic vector and also perpendicular to the direction of propagation. It is therefore in the vertical plane, but is tilted forward in the direction of propagation, making an angle of 45° with the ground. An alternative method of describing the polarization is adopted when the wave is travelling in one medium, and reaches the boundary surface of another. It is often then convenient to refer to the wave as being polarized in the plane of incidence with respect to the boundary surface, or perpendicular to this plane, as the case may be, referring to the electrical field vector in both instances. It should, however, be remembered that in most text-books dealing with the polarization of light, an older convention is adopted. A light wave is said to be polarized in the plane of incidence, when the magnetic field vector lies in this plane and the electric field vector is perpendicular thereto, which is opposite to the convention adopted in radio practice. It is easily seen that if a wave is plane-polarized at an angle with respect to the plane of incidence, the electric field vector can be resolved into two components, one in the plane of incidence and one perpendicular thereto, the two components being in phase with each other, and perpendicular to the direction of propagation.

17. If the wave front contains two mutually perpendicular electric fields of equal amplitude, but differing in phase by 90°, the wave is said to be circularly polarized. The conception is illustrated in fig. 5a, in which an electric charge Q is forced to vibrate under the action of two electric fields, one acting along the axis YOY, $\hat{\Gamma} \sin \omega t$, and one acting along the axis XOX, $\hat{\Gamma} \cos \omega t$. Under the action of the former force alone, the charge Q would execute sinusoidal

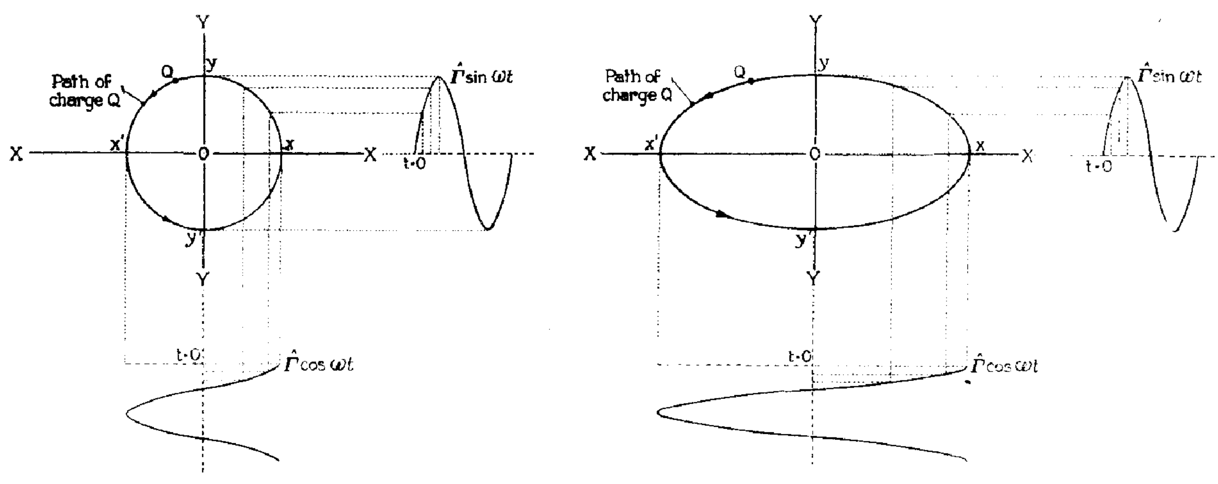


FIG. 5, CHAP. XIV.—Circular and elliptical polarization.

vibration between the points y, y′, while under the action of the latter alone, the charge Q would execute a sinusoidal vibration between the points x, x′. When both forces are applied simultaneously, the charge Q will travel round the circular path x′ y′ x y. At any instant, the charge Q must be moving in the direction of the resultant electric field, which therefore must be constantly rotating in space about the point O, with a frequency $\frac{\omega}{2\pi}$. If, however, the two fields are not of equal amplitude, although differing in phase by 90°, the conditions are those of fig. 5b.

The charge Q will now describe in space the elliptical path shown in the diagram, and the resulting force upon it must be rotating in space while undergoing a variation of amplitude according to its orientation. Such a field is said to be elliptically polarized.

Fresnel's equations

18. The first complete theory of reflection and refraction to give results in accordance with experiment was developed by Fresnel, who was unaware of the electro-magnetic nature of light, and considered its propagation to take place in an all-pervading ether having mechanical properties resembling those of an elastic solid. Fresnel's theory was therefore based on purely mechanical phenomena, and was expressed in the form of equations giving the proportions of reflected and refracted light to that incident on the boundary surface between two media. These are known as Fresnel's equations; they are also obtainable from the electro-magnetic theory of light propounded by Maxwell, but only by the use of mathematical reasoning of an advanced character. For this reason, they are here given without proof.

19. Suppose a parallel beam of plane-polarized electro-magnetic radiation to be incident at an angle φ upon the plane boundary surface between two non-conducting media the wave-front of the incident wave being also plane. Let the permittivity of the upper medium be $\varkappa_1$, that of the lower medium $\varkappa_2$, and the electric field strength of the incident wave be Γ_1. Then Γ_1 may be resolved into two components, A_1, in the plane of incidence and B_1 perpendicular thereto. Let the field strength of the refracted wave be Γ_2, and its components A_2 in, and B_2 perpendicular to, the plane of refraction, which is also, of course, the plane of incidence. The angle of refraction is denoted by θ. There may also be a reflected wave of field strength Γ_3, having components A_3 in, and B_3 perpendicular to, the plane of incidence. Fresnel's equations then give the values of A_2, B_2, A_3, B_3, in terms of A_1, B_1, and the angles of incidence and refraction. Before proceeding further, it is necessary to agree upon a convention as to the positive directions of the respective components; that usually adopted in connection with radio-communication is shown diagrammatically in fig. 6. The components in the plane of incidence, i.e., A_1, A_2, A_3, are regarded as positive upwards along the surface of the paper. The components B_1 and B_2 are regarded as positive outwards from the surface of the paper, while B_3 is positive inwards from the surface of the paper. To appreciate the reason for the convention with regard to B_3, we shall first assume its positive direction to be in the same direction as the other two. In the latter conditions, Fresnel's equations may be written

$$\left.\begin{aligned} A_2 &= A_1 \frac{2\sqrt{\varkappa_1}\cos\varphi}{\sqrt{\varkappa_1}\cos\theta + \sqrt{\varkappa_2}\cos\varphi} \\ B_2 &= B_1 \frac{2\sqrt{\varkappa_1}\cos\varphi}{\sqrt{\varkappa_2}\cos\theta + \sqrt{\varkappa_1}\cos\varphi} \end{aligned}\right\} \text{refracted wave}$$

$$\left.\begin{aligned} A_3 &= A_1 \frac{\sqrt{\varkappa_2}\cos\varphi - \sqrt{\varkappa_1}\cos\theta}{\sqrt{\varkappa_2}\cos\varphi + \sqrt{\varkappa_1}\cos\theta} \\ B_3 &= B_1 \frac{\sqrt{\varkappa_1}\cos\varphi - \sqrt{\varkappa_2}\cos\theta}{\sqrt{\varkappa_1}\cos\varphi + \sqrt{\varkappa_2}\cos\theta} \end{aligned}\right\} \text{reflected wave}$$

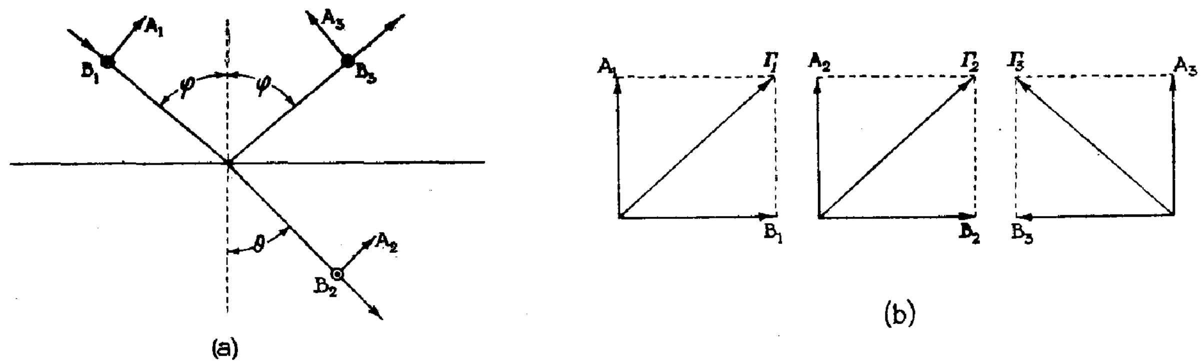


FIG. 6, CHAP. XIV.—Conventions used in Fresnel's equations.

20. In most practical work, the upper medium of fig. 6a is air, and $\varkappa_1$ may be put equal to unity. The component B_3 then becomes

$$B_3 = B_1 \frac{\cos\varphi - \sqrt{\varkappa_2}\cos\theta}{\cos\varphi + \sqrt{\varkappa_2}\cos\theta}.$$

The greatest possible value of $\cos\varphi$ and $\cos\theta$ is unity. Also, if $\varkappa_2$ is greater than unity, $\varkappa_2 \cos\theta$ is always greater than $\cos\varphi$, and $\frac{B_3}{B_1}$ is always negative. This implies that on reflection, the component perpendicular to the plane of incidence suffers a phase change of 180°, i.e., a complete reversal of phase. By reversing the assumed positive direction as in fig. 6, then, we preserve the continuity of phase and consider instead that the sign of the amplitude is reversed. With the convention of the diagram, therefore, we must write

$$B_3 = B_1 \frac{\sqrt{\varkappa_2}\cos\theta - \sqrt{\varkappa_1}\cos\varphi}{\sqrt{\varkappa_2}\cos\theta + \sqrt{\varkappa_1}\cos\varphi}.$$

This point has been dwelt on at some length because various writers use either of the above conventions, and it is always necessary to ensure that the convention is understood, e.g., before adding the fields due to direct and reflected waves.

21. The above equations completely determine the field strength of the reflected and refracted waves, when the incident field strength and the nature of the two media are known. Some important deductions may be made by the substitution of $\frac{\sin\varphi}{\sin\theta}$ for $\frac{\sqrt{\varkappa_2}}{\sqrt{\varkappa_1}}$, giving

$$A_2 = A_1 \frac{2\sin\theta\cos\varphi}{\sin(\varphi+\theta)\cos(\varphi-\theta)}$$

$$B_2 = B_1 \frac{2\sin\theta\cos\varphi}{\sin(\varphi+\theta)}$$

$$A_3 = A_1 \frac{\tan(\varphi-\theta)}{\tan(\varphi+\theta)}$$

$$B_3 = B_1 \frac{\sin(\varphi-\theta)}{\sin(\varphi+\theta)}.$$

The Brewster angle

22. First, if $\theta = \varphi$, both $\tan(\varphi-\theta)$ and $\sin(\varphi-\theta)$ are zero. This merely signifies that if the angle of refraction is equal to the angle of incidence, i.e., if $\varkappa_2 = \varkappa_1$, there can be no reflected wave, or $A_3 = B_3 = 0$. In the refracted wave $A_2 = A_1$, and $B_2 = B_1$. Second, if $\varkappa_2$ is not equal to $\varkappa_1$, θ is not equal to φ and $\sin(\varphi - \theta)$ never vanishes. Also, $\sin(\varphi+\theta)$ cannot exceed

unity. Provided, therefore, that the component B_1 has a finite value, the component B_2 of the reflected wave will always exist. In the case of the components in the plane of incidence this is not necessarily so, for A_3 is proportional to $\frac{\tan(\varphi - \theta)}{\tan(\varphi + \theta)}$ and although the numerator is never zero, $\tan(\varphi + \theta)$ becomes infinitely great, and $A_3 = 0$, if $\varphi + \theta = \frac{\pi}{2}$. When the latter relation is fulfilled, then, the reflected wave cannot possess a component in the plane of incidence. It will be observed that if $\varphi + \theta = \frac{\pi}{2}$, $\sin\theta = \cos\varphi$ and

$$\tan\varphi = \frac{\sin\varphi}{\cos\varphi} = \frac{\sin\varphi}{\sin\theta} = \nu = \sqrt{\frac{\varkappa_2}{\varkappa_1}}.$$

Under these conditions, the reflected and refracted components are perpendicular to each other as shown in fig. 7. The above equations embody what is known in optics as Brewster's Law, and the angle of incidence for which A_3 becomes zero is referred to as the Brewster angle φ_B. In the diagram, the upper medium is assumed to be air, $\varkappa_1 = 1$, and the lower to have a permittivity $\varkappa_2 = 3$. Then $\nu = \sqrt{\frac{\varkappa_2}{\varkappa_1}} = \sqrt{3} = \frac{0{\cdot}866}{0{\cdot}5} = \frac{\sin 60^\circ}{\sin 30^\circ}$, i.e., $\varphi = 60^\circ$, $\theta = 30^\circ$. The reflected ray is polarized perpendicularly to the plane of incidence, and its amplitude is equal to $B_1 \frac{\sin(\varphi - \theta)}{\sin(\varphi + \theta)} = B_1 \frac{\sin 30^\circ}{\sin 90^\circ} = 0{\cdot}5\, B_1$. If the incident ray contains a component A_1, the corresponding component of the refracted wave is

$$A_2 = A_1 \frac{2 \sin\theta \cos\varphi}{\sin(\varphi + \theta)\cos(\varphi - \theta)} = \frac{2 \sin 30^\circ \cos 60^\circ}{\sin 90^\circ \cos 30^\circ} = 1\,;$$

thus the whole of the A component is refracted. The component B_2 is equal to

$$B_1 \frac{2 \sin 30^\circ \cos 60^\circ}{\sin 90^\circ} = 0{\cdot}5\, B_1,$$

i.e., the field strength of the refracted component B_2 is equal to that of the reflected component B_1.

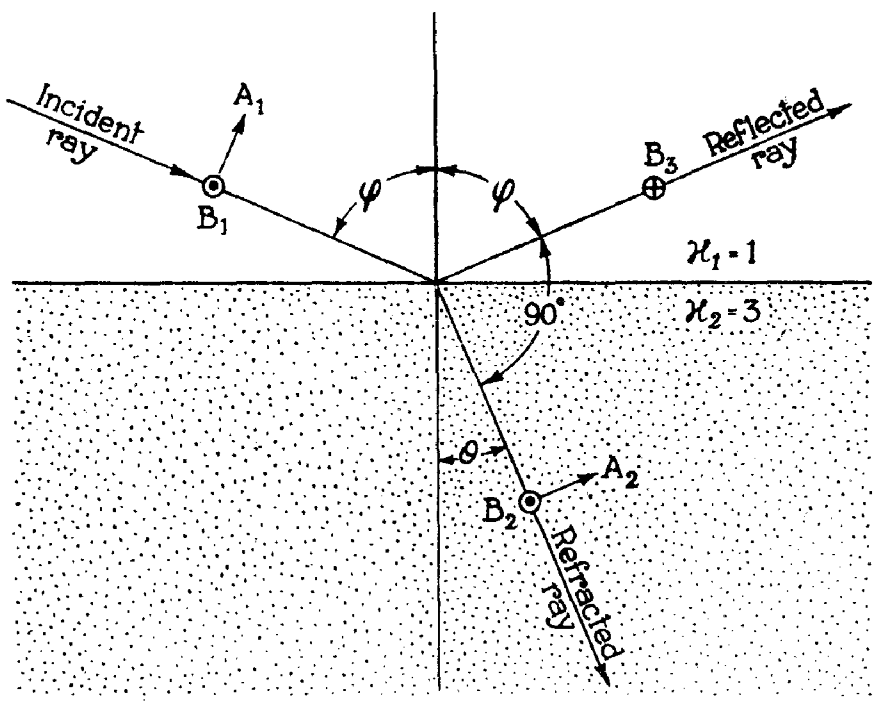
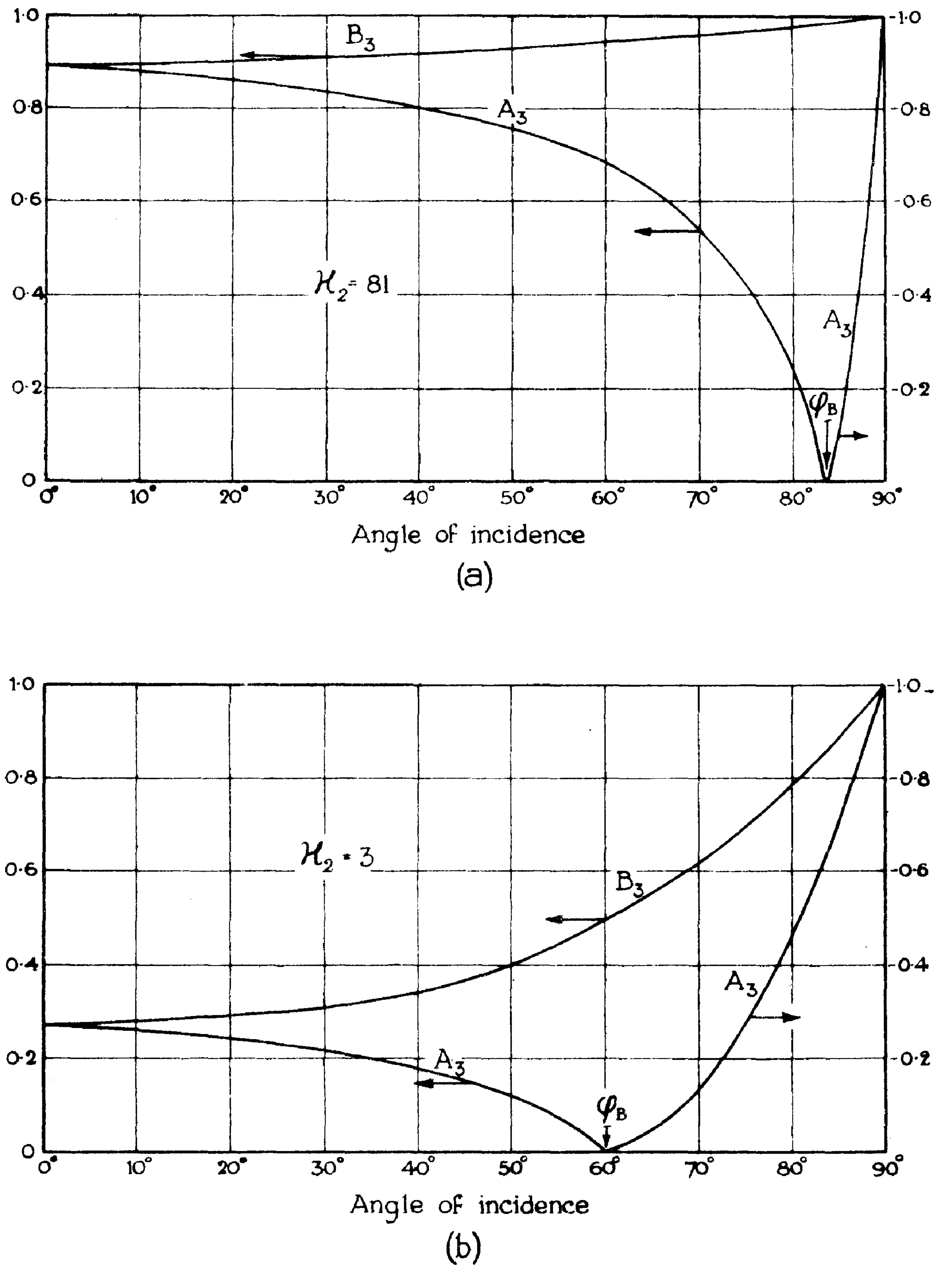


FIG. 7, CHAP. XIV.—Conditions at the Brewster angle of incidence.



AMPLITUDE OF REFLECTED WAVE

FIG. 8
CHAP. XIV

This Page Intentionally Blank

23. In radio-communication we are rarely concerned with the refracted ray, except in connection with the passage of a wave through an ionized region of the atmosphere, and in this case, since the latter has a density of ionization which varies with height, the refractive index is not constant and it is difficult to draw any useful general conclusions. Further consideration of Fresnel's equations will therefore be confined to the reflected ray. The ratios $\frac{A_3}{A_1}$ and $\frac{B_3}{B_1}$ for various angles of incidence are shown in fig. 8a for the particular case when $\varkappa_1 = 1$, $\varkappa_2 = 81$; this represents the state of affairs which would arise if the second medium were pure fresh water. If the wave is incident vertically upon the surface, and $A_1 = B_1 = 1$, both A_3 and B_3 components have an amplitude of about 0·9. As φ increases the component B_1 increases gradually, becoming unity for grazing incidence. The amplitude of the component A_3 gradually decreases, becoming zero at the Brewster angle φ_B. Since $\varphi_B = \tan^{-1}\sqrt{\frac{\varkappa_2}{\varkappa_1}}$ (= $\tan^{-1} 9$, in the particular instance illustrated), φ_B is seen to be 83·6° approximately. For angles of incidence greater than φ_B, the A component suffers a phase reversal upon reflection so that the ratio $\frac{A_3}{A_1}$ is negative. Fig. 8b shows the ratios $\frac{A_3}{A_1}$ and $\frac{B_3}{B_1}$ for $\varkappa_1 = 1$, $\varkappa_2 = 3$. It is seen that at vertical incidence the strength of each component is only 0·27 of the strength of the incident field.

Dispersion

24. Let us now consider the passage of an electro-magnetic wave through an ionized medium. We may obtain an idea of the phenomena involved from ordinary A.C. theory. In Chapter V it was shown that if an E.M.F. $\mathcal{E} \sin \omega t$ is applied to a circuit consisting of an inductance L and capacitance C in parallel, the supply current is

$$i_s = \left(\omega C - \frac{1}{\omega L}\right) \mathcal{E} \cos \omega t$$

If $\omega C > \frac{1}{\omega L}$, the current leads on the applied voltage by 90°, and lags by the same angle if $\frac{1}{\omega L} > \omega C$. Now consider a condenser consisting of two parallel plates each having an area of one square centimetre, placed one centimetre apart, the space being filled by a dielectric of permittivity $\varkappa$, and let a single free electron of charge e E.S. units and mass m C.G.S. units be situated in the space between the plates. If the space is traversed by the electric field of a wave of (peak) strength $\hat{\Gamma}$ E.S. units and frequency $\frac{\omega}{2\pi}$, the electron will be set in oscillatory motion at this frequency, its instantaneous velocity being

$$u_1 = \frac{e}{\omega m} \hat{\Gamma} \sin\left(\omega t - \frac{\pi}{2}\right)$$

i.e., the velocity will undergo sinusoidal variations lagging by $\frac{\pi}{2}$ radians or 90° on the variations of the electric field. It is, for the present, assumed that the motion is entirely undamped, i.e., that no energy is dissipated by the vibration of the electron. The peak value of the velocity will be

$$\mathcal{U}_1 = \frac{e}{\omega m} \hat{\Gamma} \text{ (centimetres per second).}$$

25. A charge e moving with a velocity of u_1 centimetres per second is equivalent to the current i_1 (E.S.U.) in an element of space one centimetre in length, or $i_1 = eu_1$. The peak value of the current will be

$$\mathcal{I}_1 = e\mathcal{U}_1 = \frac{e^2}{\omega m}\hat{\Gamma} = \frac{\hat{\Gamma}}{\omega \frac{m}{e^2}}.$$

This is clearly analogous to the current flowing in an inductance L, which is given by the equation

$$\mathcal{I}_L = \frac{\mathcal{E}}{\omega L}$$

observing that in both cases the current lags on the impressed force by 90°. The free electron in the dielectric of the condenser acts as though it were an inductance

$$L' = \frac{m}{e^2}\text{ (E.S.U.)}$$

in parallel with the capacitance C, and the effective admittance of the parallel combination is

$$\omega C' = \omega C - \frac{1}{\omega L'}$$

As the condenser in this particular instance has all its geometrical dimensions equal to unity, its capacitance is $\frac{\varkappa}{4\pi}$ E.S.U., i.e.,

$$\omega C' = \omega C - \frac{1}{\omega L'} = \frac{\omega\varkappa}{4\pi} - \frac{e^2}{\omega m}$$

$$\therefore C' = \frac{\varkappa}{4\pi} - \frac{e^2}{\omega^2 m} = \frac{1}{4\pi}\left(\varkappa - \frac{4\pi e^2}{\omega^2 m}\right)$$

26. (i) The effective permittivity of the dielectric is therefore not $\varkappa$ but $\varkappa - \frac{4\pi e^2}{\omega^2 m}$; if instead of a single electron we consider the effect of N electrons in the unit volume of dielectric, the current will be Neu_1 E.S. units and the effective dielectric constant becomes

$$\varkappa' = \varkappa - \frac{4\pi N e^2}{\omega^2 m}$$

Since the permittivity of a vacuum is unity, the above consideration shows that the permittivity of a vacuous space containing N electrons per cubic centimetre is $1 - \frac{4\pi N e^2}{\omega^2 m}$. The phase velocity, u_p, of propagation of electro-magnetic waves in entirely empty space being $\frac{c}{\sqrt{\varkappa}}$ centimetres per second, it follows that, in space occupied only by free electrons

$$u_p = \frac{c}{\sqrt{\varkappa'}} = \frac{c}{\sqrt{1 - \frac{4\pi N e^2}{\omega^2 m}}},$$

i.e., the phase velocity is greater (and the group velocity less) than the velocity c in free space

(ii) If positive ions (i.e., gaseous atoms minus one or more electrons) are also present, these will also affect the value of κ' and therefore the phase and group velocities, but since the mass of such an atom is at least 1840 times as great as that of the electron the effect on the permittivity is very small. In the above expression N represents the number of free electrons per cubic centimetre since $e = 4 \cdot 77 \times 10^{-10}$ E.S. units and $m = 8 \cdot 8 \times 10^{-28}$ gram,

$$\frac{e^2}{m} = \frac{4 \cdot 77 \times 10^{-10}}{8 \cdot 8 \times 10^{-28}}$$

$$= 2 \cdot 6 \times 10^8$$

It is convenient to refer to the frequency f rather than the angular velocity ω of the wave, giving

$$u_p = \frac{c}{\sqrt{1 - \dfrac{4\pi N \times 2 \cdot 6 \times 10^8}{4\pi^2 f^2}}}$$

$$= \frac{c}{\sqrt{1 - 8 \cdot 27 \times 10^7 \dfrac{N}{f^2}}},$$

thus the phase velocity of the wave in an ionized medium depends upon the frequency; the higher the frequency, the smaller is the effect of the ionization, and we realize why the refractive index of a medium, as measured by observations at very high frequency, i.e., visible light waves, may differ from that obtained by measurement of the permittivity at radio-communication frequencies. Since u_p is greater than c, a wave, in crossing the boundary between free space and an ionized medium, will be refracted away from the normal.

Effect of ionization gradient

27. Now let us suppose that in a certain region the ionization increases progressively. To take a concrete case, let the line A C in fig. 9 represent the lower boundary of a medium in which the number of electrons per cubic centimetre increases uniformly in the upward direction. Let a ray P A enter the medium with a velocity c, at an angle θ_i. Just inside the boundary, if the refractive index is ν, the phase velocity will be

$$u_p = \frac{c}{\nu},$$

and $\nu \sin \theta_r = \sin \theta_i =$ constant.

The ray is therefore bent at the point of entry; during its passage through a further small element of distance in the medium, the value of ν changes, but so also does the angle of the ray with reference to the boundary surface. At any point where the direction of propagation makes an angle θ with the normal to the boundary, $\nu \sin \theta$ remains constant. It follows that the effect of the increasing electronic density, or reduction of refractive index, will be to increase the angle θ, which may eventually become 90°. The ray then commences to pass through successive strata of progressively decreasing ionization, and is therefore bent downwards until it finally passes through the boundary at an angle equal to θ_r, with which it entered. At the highest point B, where the electron density is, say, N_B electrons per cubic centimetre, $\theta_B = 90°$, $\sin \theta_B = 1$, and

$$\nu_B = \sin \theta_i = \frac{c}{u_p}$$

$$= \sqrt{1 - \frac{8 \cdot 27 \times 10^7 N_B}{f^2}}$$

$$\text{or } \nu_B = 0 \text{ if } N_B = \frac{f^2}{8 \cdot 27 \times 10^7}.$$

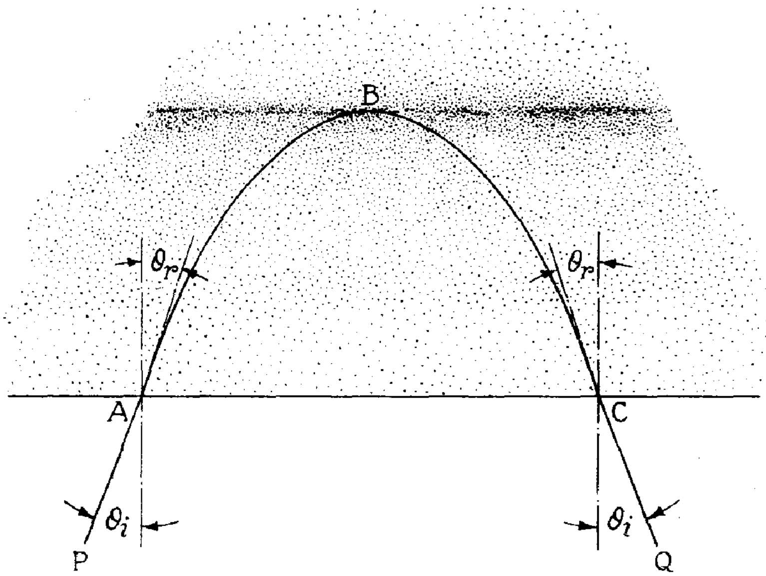


FIG. 9, CHAP. XIV.—Refraction in an ionized region.

If the electron density exceeds this figure, the direction of travel of the ray will be completely reversed. It will be noted that a region having a refractive index of zero is impenetrable by an electro-magnetic wave.

Relation between ionization and conductivity

28. If the wave is travelling through a space containing matter, e.g., an ionized gas, collisions occur between electrons (or ions) and the gas molecules, and it can be shown that if p such collisions take place per second, the permittivity of the medium is

$$\varkappa' = \varkappa - \frac{4\pi Ne^2}{(p^2 + \omega^2)m}$$

The effect of these collisions is to cause a dissipation of energy, and we may infer that the collision frequency p will enter into any expression giving the specific resistance, or alternatively, the conductivity, of the medium. The conductivity of a gaseous medium is

$$\sigma = \frac{pNe^2}{(p^2 + \omega^2)m} \text{ (E.S.U.)}$$

The change of permittivity and the conductivity of the medium are evidently closely related physically. As already shown, the reduction in permittivity is due to a velocity component 90° out of phase with the field, while the conductivity is represented by a velocity component in anti-phase with the field causing the motion. The passage of a wave through a material of finite conductivity therefore takes place with a progressive decrease of amplitude, and the wave is said to be attenuated. The conductivity is sometimes expressed in E.S.U., sometimes in E.M.U. and sometimes as the reciprocal of the specific resistance ϱ (ohms per centimetre cube). The following conversion factors are therefore given.

$$\sigma \text{ (E.S.U.)} = \sigma \text{ (E.M.U.)} \times (9 \times 10^{20})$$
$$= \sigma' \times 9 \times 10^{11}$$
$$\text{where } \sigma' = \frac{1}{\varrho}.$$

Effect of conductivity

29. (i) We may obtain some idea of the effect of conductivity in a medium in which an electro-magnetic wave is propagated, by first considering the motion of an electron which is situated between the parallel plates of a condenser of unit dimensions, the permittivity of its

dielectric being $\varkappa'$. We may assume that this dielectric has a number of free electrons, the conductivity being a measure of the ability or otherwise of these electrons to execute sinusoidal vibration in the presence of an electric field. The effect of a lack of freedom is to reduce the amplitude of vibration and to cause the velocity of the electrons to lag by less than 90° on the electric field producing the motion.

(ii) Let the conductivity of the medium be σ E.S. units and a sinusoidal P.D. of v E.S. units be set up between the plates of the condenser. As a result of this P.D. an alternating current will flow. This current will consist of a displacement component

$$i_d = j\omega C v$$

and a conduction component

$$i_c = gv$$

so that the total current is (in E.S.U.)

$$i = (g + j\omega C)v.$$

Since the condenser is of unit dimensions, $C = \frac{\varkappa'}{4\pi}$ and g $= \sigma$, hence

$$i = \left(\sigma + j\omega \frac{\varkappa'}{4\pi}\right)v$$

$$= \frac{j\omega}{4\pi}\left(\varkappa' - j\frac{4\pi}{\omega}\sigma\right)v$$

$$= \frac{j\omega}{4\pi}\left(\varkappa' - j\frac{2\sigma}{f}\right)v.$$

30. (i) In effect, therefore, the permittivity of the conductive medium is $\varkappa' - j\frac{2\sigma}{f}$ instead of $\varkappa$ as in a perfect dielectric. It is now necessary to apply this reasoning to the special case where the electric field between the plates is set up by an electro-magnetic wave, which is propagated parallel to them in such a manner that the electric field vector of the wave is in the same direction as the P.D. just discussed. It must be observed that since the only object of the plates is to distribute the charge over a given volume of dielectric, we may now ignore their existence, thus removing complications due to the presence of the highly conductive boundary surfaces.

(ii) The equation of a travelling wave of frequency $\frac{\omega}{2\pi}$ is given in Chapter VII. In the present instance we may write for the electric field

$$\gamma = \hat{\Gamma} \cos \frac{2\pi}{\lambda}\left(u_p t - x\right),$$

where u_p is the phase velocity. As already shown, in a perfect dielectric $u_p = \frac{c}{\sqrt{\varkappa}}$. Eliminating λ, then,

$$\gamma = \hat{\Gamma} \cos \omega\left(t - \frac{x}{u_p}\right)$$

and for a conductive dielectric, instead of $u_p = \frac{c}{\sqrt{\varkappa'}}$, we must write $u_p = \frac{c}{\sqrt{\varkappa' - j\frac{2\sigma}{f}}}$ so that

$$\gamma = \hat{\Gamma} \cos \omega\left(t - \frac{x}{c}\sqrt{\varkappa' - j\frac{2\sigma}{f}}\right).$$

To interpret this equation, we use Demoivre's theorem (Chapter V). It is there shown that $\cos\theta$ is the real part of $\varepsilon^{j\theta}$. Since the real and imaginary parts of a complex number are entirely independent we may perform any operation on $\varepsilon^{j\theta}$, and take the real part of the result to give the effect of the operation upon $\cos\theta$. In the present example then, if

$$\gamma = \hat{\Gamma} \cos \omega \left(t - \frac{x}{c}\sqrt{\kappa' - j\frac{2\sigma}{f}}\right)$$

$$= \text{real part of } \hat{\Gamma}\, \varepsilon^{j\omega\left(t - \frac{x}{c}\sqrt{\kappa' - j\frac{2\sigma}{f}}\right)}$$

and $\sqrt{\kappa' - j\frac{2\sigma}{f}} = \nu - j\alpha.$

$$\gamma = \hat{\Gamma}\, \varepsilon^{j\omega\left(t - \frac{x}{c}\nu + j\frac{x}{c}\alpha\right)}$$

$$= \hat{\Gamma}\, \varepsilon^{j\omega\left(t - \frac{x}{c}\nu\right)} \varepsilon^{-\frac{\omega x}{c}\alpha}$$

$$= \varepsilon^{-\frac{\omega x\alpha}{c}}\, \hat{\Gamma} \cos \omega\left(t - \frac{x}{c}\nu\right).$$

Hence the phase velocity is $\frac{c}{\nu}$. The exponential factor $\varepsilon^{-\frac{\omega x\alpha}{c}} = \varepsilon^{-\frac{2\pi\alpha}{\lambda}x}$ gives the rate at which the amplitude decreases as the wave progresses.

31. If both κ' and $\frac{\sigma}{f}$ are known it is easy to find numerical values for ν and α. Since

$$\sqrt{\kappa' - j\frac{2\sigma}{f}} = \nu - j\alpha,$$

$$\kappa' - j\frac{2\sigma}{f} = \nu^2 - 2j\nu\alpha - \alpha^2$$

$$\therefore \kappa' = \nu^2 - \alpha^2$$

$$\frac{2\sigma}{f} = 2\nu\alpha$$

$$(\kappa')^2 = \nu^4 - 2\nu^2\alpha^2 + \alpha^4$$

$$\left(\frac{2\sigma}{f}\right)^2 = 4\,\nu^2\alpha^2$$

$$(\kappa')^2 + \left(\frac{2\sigma}{f}\right)^2 = \nu^4 + 2\nu^2\alpha^2 + \alpha^4$$

$$\sqrt{(\kappa')^2 + \left(\frac{2\sigma}{f}\right)^2} = \nu^2 + \alpha^2$$

Hence

$$2\nu^2 = \sqrt{(\kappa')^2 + \left(\frac{2\sigma}{f}\right)^2} + \kappa'$$

$$2\alpha^2 = \sqrt{(\kappa')^2 + \left(\frac{2\sigma}{f}\right)^2} - \kappa'$$

i.e.,

$$\nu = \sqrt{\sqrt{\left(\frac{\kappa'}{2}\right)^2 + \left(\frac{\sigma}{f}\right)^2} + \frac{\kappa'}{2}}$$

$$\alpha = \sqrt{\sqrt{\left(\frac{\kappa'}{2}\right)^2 + \left(\frac{\sigma}{f}\right)^2} - \frac{\kappa'}{2}}.$$

These formulæ have an important bearing upon the depth of penetration of a radio wave into the earth. If $\varkappa'$ is small compared with $\frac{2\sigma}{f}$, we may write

$$\alpha = \sqrt{\frac{\sigma}{f} - \frac{\varkappa'}{2}}$$

Thus if $\sigma = 10^8$ E.S.U., $f = 1$ Mc/s, $\varkappa' = 10$, $\frac{\sigma}{f} = 100$ and

$$\alpha = \sqrt{95}$$
$$= 9\cdot75.$$

The rate at which the amplitude of the wave is reduced is given by

$$\Gamma_x = \Gamma_0 \, \varepsilon^{-\frac{\omega x \alpha}{c}}$$

The amplitude will be reduced to $\frac{1}{\varepsilon} = 0\cdot367$ of its original value in passing through a distance x, where

$$\frac{\omega x \alpha}{c} = 1$$

$$x = \frac{c}{\omega\alpha}$$

$$= \frac{3 \times 10^{10}}{2\pi \times 10^6 \times 9\cdot75}$$

$$= 440 \text{ centimetres.}$$

Also $\nu = \sqrt{105} = 10\cdot25$. The phase velocity in the ground is less than one-tenth of that in free space.

Effect of magnetic field

32. When the ionized region through which an electro-magnetic wave is propagated is also occupied by a magnetic field, the character of the wave may be considerably modified during its passage. An attempt will be made to show the manner in whi.h these effects are produced. We may first consider a charge of Q (E.M.U.) moving with constant velocity U through a uniform

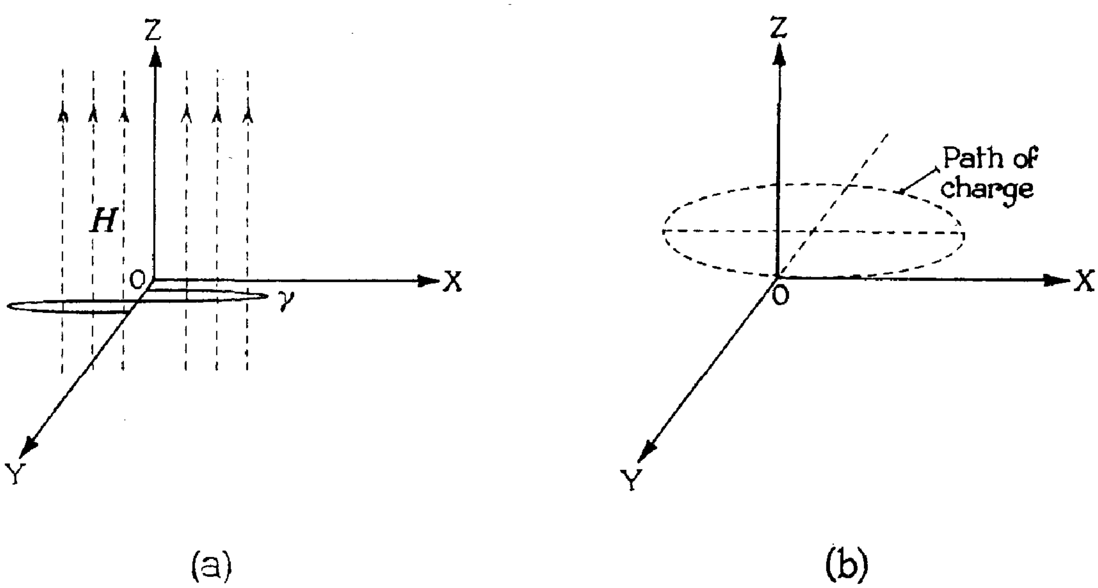


FIG. 10, CHAP. XIV.—Effect of magnetic field.

magnetic field of strength H (E.M.U.). Since a moving charge is equivalent to an electric current, there will be a force F acting upon the charge, its direction being perpendicular to the directions both of H and U, and its magnitude will be $F = Q\,H\,U$ dynes. Let the charge be carried by a mass of m grams, and at a time $t = 0$ to be situated at the origin O of the rectangular co-ordinates O X, O Y, O Z, where H is along O Z, as in fig. 10a. Suppose a stationary sinusoidal electric field $\gamma = \hat{\Gamma} \sin \omega t$ to be applied along the axis O X. At $t = 0$, let the instantaneous velocity of the charge along O X be $u_x = \frac{dx}{dt}$ and along O Y be $u_y = \frac{dy}{dt}$. Then the charge will be subject to forces F_x and F_y where

$$F_x = Q\,(\hat{\Gamma} \sin \omega t - H\,u_y)$$
$$F_y = Q\,(H\,u_x).$$

Since force = mass × acceleration, we may write

$$m\frac{d^2x}{dt^2} = Q\left(\hat{\Gamma} \sin \omega t - H\frac{dy}{dt}\right) \qquad \text{.. .. ..} \qquad (a)$$

$$m\frac{d^2y}{dt^2} = QH\frac{dx}{dt} \qquad \text{.. .. ..} \qquad (b)$$

From (b)

$$m\frac{d^3y}{dt^3} = QH\frac{d^2x}{dt^2}$$

$$\frac{d^2x}{dt^2} = \frac{m}{QH}\frac{d^3y}{dt^3}$$

inserting this value for $\frac{d^2x}{dt^2}$ in (a)

$$\frac{m^2}{QH}\frac{d^3y}{dt^3} + QH\frac{dy}{dt} = Q\,\hat{\Gamma} \sin \omega t$$

$$m\frac{d^3y}{dt^3} + \frac{Q^2H^2}{m}\frac{dy}{dt} = \frac{Q^2H}{m}\hat{\Gamma} \sin \omega t$$

This equation can be immediately integrated, giving

$$m\frac{d^2y}{dt^2} + \frac{Q^2H^2}{m}y = -\frac{Q^2H}{\omega m}\hat{\Gamma} \cos \omega t + K$$

The constant K depends on the initial conditions and may be made zero by a suitable choice of these. The solution of an equation of the form

$$P\frac{d^2y}{dt^2} + Ry = S \cos \omega t$$

is given in standard text-books on differential equations. The complete solution will consist of a forced, i.e., sustained oscillation and a damped oscillation. The forced oscillation along O Y will take place in accordance with the equation

$$y = \frac{Q^2H}{\omega\,(\omega^2m^2 - Q^2H^2)}\hat{\Gamma} \cos \omega t,$$

and the motion along the axis O X is given by

$$x = \frac{Qm}{\omega^2m^2 - Q^2H^2}\hat{\Gamma} \sin \omega t.$$

Comparing these results with those of para. 17 it will be seen that the path of the charge will be an ellipse, as shown in fig. 10b.

33. (i) The point of immediate interest in the above expressions is that, theoretically, the amplitudes of vibration along both O X and O Y become infinite if $\omega^2 m^2 = Q^2 H^2$ or

$$\omega = \frac{QH}{m}.$$

Actually, of course, the amplitude would gradually build up from zero and would take an infinite time to reach an infinite amplitude. The position is in fact analogous to that reached in considering the current in an acceptor circuit with zero resistance. The presence of the magnetic field introduces into the motion of the charge a kind of resonance effect. If the charge, hitherto denoted by Q (E.M.U.) is that of an electron, the resonant frequency is $\frac{\omega}{2\pi} = \frac{eH}{mc}$ where e is the charge of an electron in E.S.U.

(ii) When the sinusoidal electric field, instead of being stationary, is due to the propagation of an electro-magnetic wave, the effects are somewhat complicated, depending upon the direction of the magnetic field relative to the direction of propagation and to the polarization of the wave on entering the field. Assuming the wave to be plane polarized on entry, if the electric field vector Γ is parallel to the magnetic field vector H, the latter has no effect upon the propagation characteristics. If, however, Γ is perpendicular to H, various effects are produced, depending upon the direction of propagation. If the latter is parallel to H, the plane polarized wave is split into two circularly polarized components, which undergo different degrees of absorption and refraction. If both the direction of propagation and the electric field vector are perpendicular to H, e.g., if the former is along O X and the latter along O Y, it is split into two plane-polarized waves, which again undergo unequal degrees of absorption and refraction. In general, the propagation is along neither of the axes, but in some intermediate direction, and the emergent wave is elliptically polarized.

Fresnel's equations for conductive media

34. Fresnel's equations may be applied to the problem of finding the respective amplitudes of the reflected and refracted waves at the boundary surface between two conductive media, by substituting the complex permittivity of each medium for the simple permittivity appropriate to a perfectly non-conducting substance. If the incident wave is in air or free space, its permittivity is unity and its conductivity zero. Let the permittivity of the second medium be $\varkappa$, and its conductivity σ. Its complex permittivity is then $\varkappa' = \varkappa - j\frac{2\sigma}{f}$, where f is the frequency of the incident disturbance. Then

$$A_3 = A_1 \frac{\varkappa' \cos\varphi - \sqrt{\varkappa' - \sin^2\varphi}}{\varkappa' \cos\varphi + \sqrt{\varkappa' - \sin^2\varphi}},$$

$$B_3 = B_1 \frac{\sqrt{\varkappa' - \sin^2\varphi} - \cos\varphi}{\sqrt{\varkappa' - \sin^2\varphi} + \cos\varphi},$$

the positive directions being as shown in fig. 6. Analogous expressions for the refracted wave are easily deduced from the equations of paragraph 19 if required.

Reflection coefficient, horizontal polarization

35. When the reflection takes place at the surface of the earth, A_3 is identified as the vertically polarized component and B_3 as the horizontally polarized component of the field strength in the reflected wave. It is now convenient to change the notation slightly. The ratios $\frac{A_3}{A_1}$, $\frac{B_3}{B_1}$, may be termed the reflection coefficients for the respective states of polarization and are complex quantities. The reflection of a horizontally polarized wave will be first dealt with, as the results

are somewhat the simpler. The ratio $\frac{B_3}{B_1}$, is a complex number and may be written

$$\frac{B_3}{B_1} = K_h \underline{/\theta_h} = \frac{\sqrt{\varkappa - j\frac{2\sigma}{f} - \sin^2\varphi} - \cos\varphi}{\sqrt{\varkappa - j\frac{2\sigma}{f} - \sin^2\varphi} + \cos\varphi}$$

Then K_h is the arithmetical ratio of the strengths of the reflected and incident fields and θ_h is an angle which must be added to the phase of the incident beam to obtain the phase of the reflected wave. The positive directions must be taken as in fig. 6 when making such additions. The angle θ_h is always negative, lying between zero and $-180°$. If K_h and θ_h are calculated for given values of the constants $\varkappa$, σ, f, and plotted against the angle of incidence φ, the curves obtained always resemble those marked K_h and θ_h in fig. 11a and fig. 11b. respectively.

36. The numeric K_h is very easily calculated for the particular case when $\varphi = 0$, corresponding to vertical incidence, as in the following example.

$$\text{Let } \varkappa = 9, \frac{2\sigma}{f} = 10, \varphi = 0$$

$$\cos\varphi = 1, \ \sin\varphi = 0$$

$$K_h \underline{/\theta_h} = \frac{\sqrt{9 - j\,10} - 1}{\sqrt{9 - j\,10} + 1}$$

$$\text{Let } \sqrt{9 - j\,10} = \nu - j\,\alpha$$

$$9 - j\,10 = \nu^2 - 2j\,\alpha\,\nu - \alpha^2$$

$$\therefore \quad \nu^2 - \alpha^2 = 9$$

$$2\,\alpha\,\nu = 10$$

$$(\nu^2 - \alpha^2)^2 = \nu^4 - 2\,\nu^2\,\alpha^2 + \alpha^4 = 81$$

$$(2\,\alpha\,\nu)^2 = 4\,\nu^2\,\alpha^2 = 100$$

$$\therefore \quad \nu^4 + 2\,\nu^2\,\alpha^2 + \alpha^4 = 181$$

$$\text{and} \quad \nu^2 + \alpha^2 = \sqrt{181} = 13{\cdot}44$$

$$\nu^2 - \alpha^2 = 9$$

$$\therefore \quad 2\,\nu^2 = 22{\cdot}44$$

$$\nu = 3{\cdot}34$$

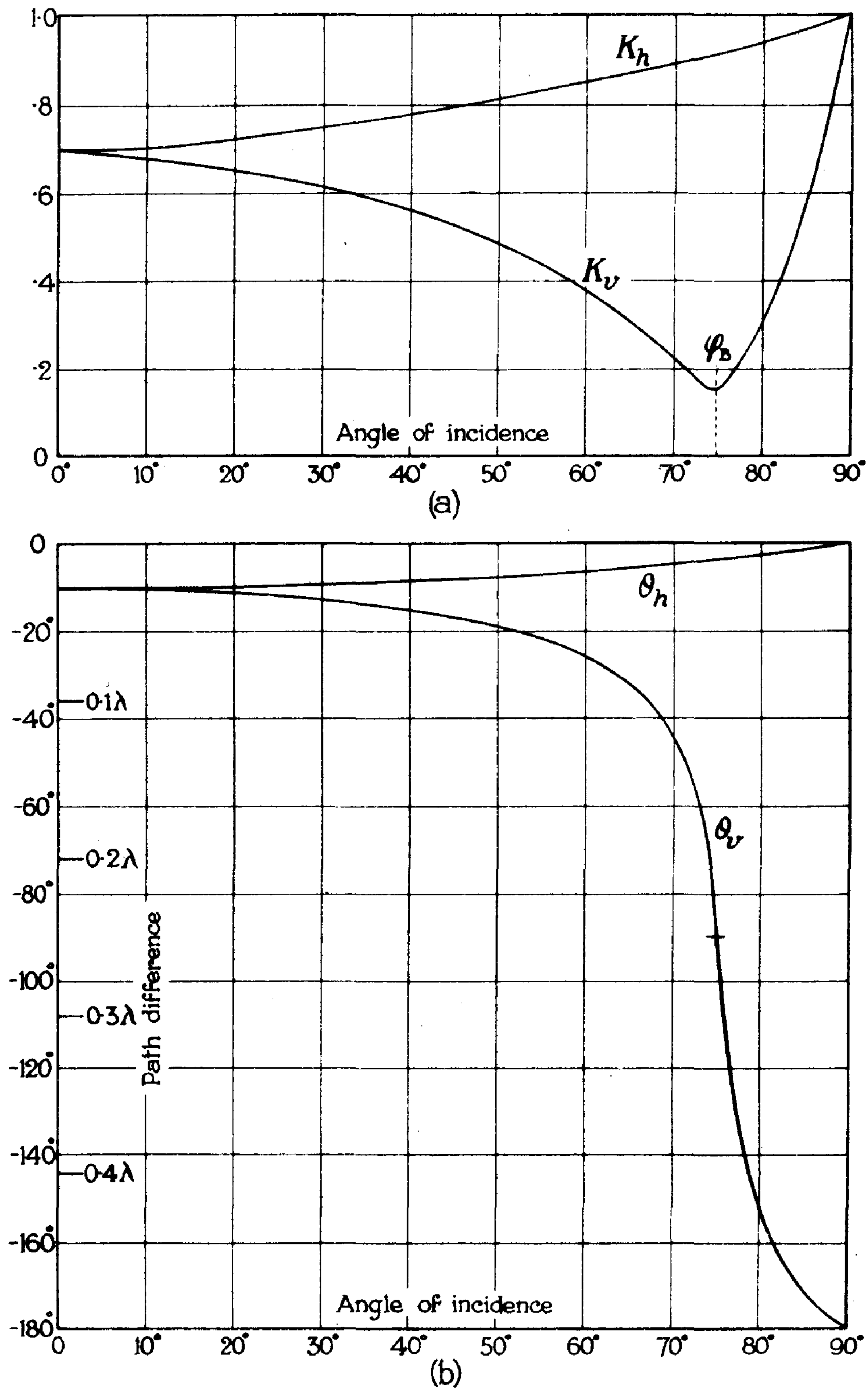
$$\text{and} \quad 2\,\alpha^2 = 4{\cdot}44$$

$$\alpha = 1{\cdot}49$$

$$\text{Therefore} \quad \sqrt{9 - j\,10} = 3{\cdot}34 - j\,1{\cdot}49$$

The accuracy of this result may be checked by squaring $(3{\cdot}34 - j1{\cdot}49)$

$$(3{\cdot}34 - j1{\cdot}49)^2 = 11{\cdot}17 - 2j\,4{\cdot}98 - 2{\cdot}22 = 8{\cdot}95 - j\,9{\cdot}96.$$



REFLECTION COEFFICIENTS

$\left(\kappa = 9, \ \frac{2\sigma}{f} = 10\right)$

FIG. 11
CHAP. XIV

This Page Intentionally Blank

We now have

$$K_h \underline{/\theta_h} = \frac{3\cdot 34 - j\,1\cdot 49 - 1}{3\cdot 34 - j\,1\cdot 49 + 1}$$

$$= \frac{2\cdot 34 - j\,1\cdot 49}{4\cdot 34 - j\,1\cdot 49}$$

$$= \frac{13\cdot 7 - j\,3}{21}$$

$$K_h = 0\cdot 67$$

$$\theta_h = tan^{-1}\left(-\frac{3}{13\cdot 7}\right)$$

$$= -\, tan^{-1}\, 0\cdot 22$$

$$= -\, 12\tfrac{1}{2}^\circ \text{ approximately.}$$

When the wave is incident at an angle of 90°, $\cos\varphi = 0$, and therefore $K_h = 1$, $\theta_h = 0°$. The complete variation of K_h and θ_h with φ, for $\varkappa = 9$, $\frac{2\sigma}{f} = 10$, is shown in fig. 11.

37. When a radio receiver is subject to the action of a downcoming wave, the beam is so wide that the incident and reflected waves cannot be separated as they can be when a parallel beam of light is reflected from a plane surface. The receiver is in general subject to both the incident and reflected beam and it is necessary to investigate the resultant beam when both are present. In fig. 12 let us consider the field at a point P, situated at a height h above the ground. The electric field due to the incident beam along D P sets up a field $\gamma = \hat{\Gamma}\, cos\, \omega t$, at the point P, measured positively above the surface of the paper. The reflected wave arrives at P by the path A B P. Draw H B perpendicular to D P and let H P = d. Then at the point of incidence B the incident wave has a field strength $\hat{\Gamma}\, cos\left(\omega t + \frac{2\pi d_1}{\lambda}\right)$ measured positively above the surface of the paper. The wave reflected at the point B produces at P the field

$$K_h\, \hat{\Gamma}\, cos\left\{\omega t + \frac{2\pi}{\lambda}(d_1 - d_2) + \theta_h\right\}$$

which is positive in the direction below the surface of the paper. Consideration of the figure shows that $d_2 - d_1 = 2h \cos\varphi$, and the electric field at P is therefore

$$\hat{\Gamma}\left\{cos\, \omega t - K_h\, cos\left[\omega t + \theta_h - \frac{2\pi}{\lambda}\left(2h\, cos\, \varphi\right)\right]\right\}$$

38. This expression has an interesting application in the reception on the ground of horizontally polarized waves emitted by an aircraft. If the aircraft is very remote from the receiver, φ will be nearly 90°, and if h is small, $\frac{2\pi}{\lambda}(2h\, cos\, \varphi)$ will also be a very small angle. When φ approaches 90°, K_h is nearly unity and θ_h only a few degrees, hence the electric field strength at the point P is the resultant of two fields of almost equal intensity which are very nearly in antiphase, and is thus very small. If, however, the receiving aerial is at a height h, where

$$\frac{2\pi}{\lambda}(2h\, cos\, \varphi) = \pi, \text{ or } h = \frac{\lambda}{4\, cos\, \varphi},$$

the two fields are of almost equal intensity and arrive at P practically in phase, giving a very strong signal.

Reflection coefficient, vertical polarization

39. Turning now to a consideration of the vertically polarized wave, the phenomena are seen to be more complicated than in the former instance. The ratio of reflected to incident field strength is

$$\frac{A_3}{A_1} = K_v \,\underline{/\theta_v} = \frac{\left(\varkappa - j\frac{2\sigma}{f}\right)\cos\varphi - \sqrt{\varkappa - j\frac{2\sigma}{f} - \sin^2\varphi}}{\left(\varkappa - j\frac{2\sigma}{f}\right)\cos\varphi + \sqrt{\varkappa - j\frac{2\sigma}{f} - \sin^2\varphi}}$$

Here again K_v is the numerical value of the ratio and θ_v is an angle to be added to the phase of the reflected wave ; θ_v is always negative, lying between zero and $-180°$. It is easily seen, by putting $\varphi = 0$, $\cos\varphi = 1$, $\sin\varphi = 0$, that for vertical incidence, $K_v \,\underline{/\theta_v}$ is equal to $K_h \,\underline{/\theta_h}$. As the angle of incidence increases, K_v decreases and reaches a minimum value. Over this range of variation of φ, θ_v approaches the value $-90°$, at first slowly but afterwards more rapidly. The value of φ at which $\theta_v = -\frac{\pi}{2}$ is known as the pseudo-Brewster angle ; instead of complete extinction, such as would occur at the surface of a perfect dielectric, a marked reduction in amplitude of the reflected wave occurs when the wave is incident at the pseudo-Brewster angle. For greater values of φ, the change of phase increases very rapidly and is $-180°$ when $\varphi = 0$. The variation of K_v and θ_v, with φ, is also shown in fig. 11, for $\varkappa = 9$, $\frac{2\sigma}{f} = 10$.

40. The field due to a downcoming wave and its reflection can be worked out in a manner similar to the previous case, but it is not so easy to draw any definite conclusions as to the effect of elevation of the receiving aerial. Reference to fig. 12 shows that if at the point P there is a

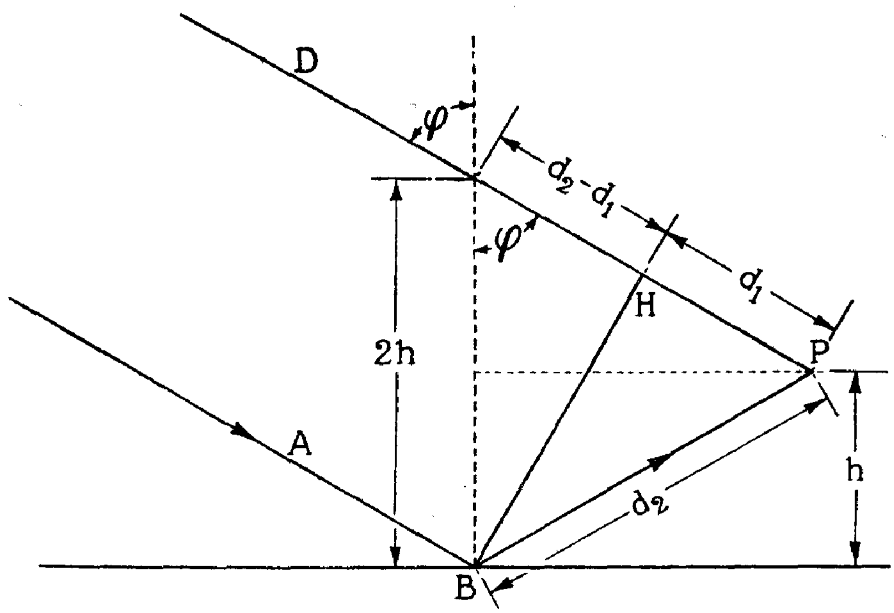


Fig. 12, CHAP. XIV.—Effect of direct and reflected waves upon elevated aerial.

field $\hat{\Gamma}\cos\omega t$ due to the direct action of the downcoming signal, there will be a reflected field of strength

$$K_v\,\hat{\Gamma}\cos\left[\omega t + \theta_v - \frac{2\pi}{\lambda}(2h\cos\varphi)\right]$$

If φ is very nearly 90° and greater than the pseudo-Brewster angle, while h is small, K_v will be nearly unity and θ_v nearly $-180°$, so that the two fields tend to annul each other. Near the ground, therefore, an almost grazing beam and its reflection nearly cancel other each in their effect upon a receiving aerial.

Propagation along ground

41. At first sight, the results obtained in paras. 37, 38 and 40 lead to the conclusion that neither a horizontally nor a vertically polarized wave can be propagated along the earth's surface, for at grazing incidence the direct and reflected waves completely cancel each other. It must however be remembered that the theory is not completely in accordance with fact. Both the incident wave-front and the reflecting surface are assumed to be plane, whereas plane waves are physically unrealizable and the earth's surface is also curved. The theory must therefore not be regarded as an attempt to deny the experimentally established fact that the radiation from a transmitting aerial may be propagated along the earth's surface to some extent, but merely as showing that the mechanism of the ground wave, as this radiation is called, is more complex than it appears. The exact mode of propagation is still debatable. It appears certain that the phenomenon of diffraction is responsible for the propagation to a limited extent, but diffraction alone will not account for the propagation of waves of frequencies higher than about 500 kc/s. According to certain physicists, a vertical aerial on or near the ground gives rise to what is called a surface wave, in addition to the radiation predicted by the simple theory given in Chap. XV. This surface wave must not be confused with the ground wave referred to above, which is assumed to be part of the radiation predicted by the simple theory, but affected by the ground constants in such a manner that it is guided round the protuberance of the earth. The surface wave, where it exists, is strongest at the surface, the field strength both above and below becoming rapidly weaker. Experiments at different frequencies in some cases appear to support and in others to deny the existence of the vertically polarized surface wave. Both theory and experiment, however, agree that a horizontally polarized wave cannot be propagated along a conductive surface to an appreciable distance. Suppose a wave to be initiated in such a manner that it starts to travel over and parallel to the earth's surface, plane polarized at an angle of 45°. The electric field then has components A_1 in the vertical and B_1 in the horizontal plane. Assuming for the moment that A_1 remains vertical, B_1 will be reduced to zero after travelling a comparatively short distance, for it induces radio-frequency currents in the earth, and the energy so dissipated must be supplied by the energy originally possessed by the component B_1.

42. The following account of an experiment actually performed many years ago at the National Physical Laboratory, Teddington, may be of interest. In order to verify or disprove the above conclusion, an attempt was made to produce a wave which travelled over the surface of the earth with horizontal polarization, the transmitting " aerial " being simply a large solenoid of 42 turns, 5 feet in diameter, which was placed with its axis in the vertical plane. This was connected to a condenser having negligible external field, thus constituting a closed, feebly radiating oscillatory circuit. The polarization of the radiated wave at a receiver only a few miles away was found to be vertical. A second experiment for the same purpose produced exactly the same effect. In this case the aerial of a transmitting station was completely removed ; the counterpoise (which measured 600 feet by 150 feet, and was 9 feet above the ground) was cut into two equal portions which were then energized as the aerial and counterpoise respectively of a transmitting aerial system. The natural polarization of such a radiator is horizontal, and it was thought possible that at Slough, a distance of a few miles, the horizontally polarized component of the wave would be at least as strong as the vertically polarized component, but the energy actually received was found to be almost entirely due to a vertically polarized wave.

43. An additional effect of the finite conductivity of the earth is to cause the wave front to be tilted forward in the direction of propagation, because the velocity just within and just above the ground must be the same, i.e., the velocity cannot change abruptly. The velocity at ground level is therefore a few parts in one hundred less than the velocity at a considerable height. Suppose the waves to be polarized in the plane of incidence, and the angle of tilt at the surface of

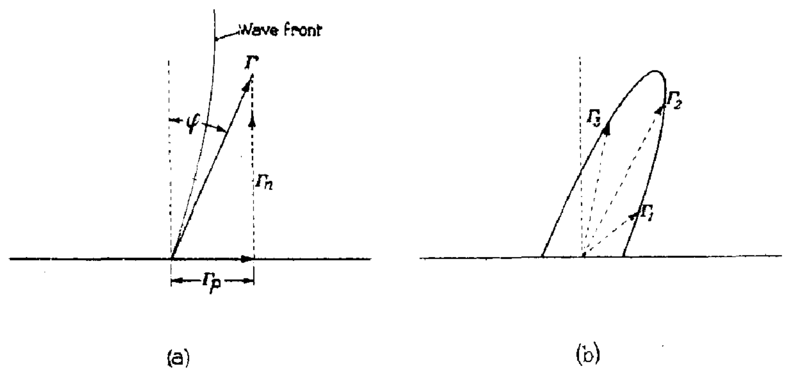


FIG. 13, CHAP. XIV.—Tilting of wave at earth's surface.

the earth (measured from the normal) to be φ (fig. 13a). The electric field vector then consists of two components, (i) a truly vertical component Γ_n and (ii) a horizontal component Γ_p, having its positive direction in the direction of propagation. This component sets up eddy currents in the ground, and as a result of the energy dissipation the horizontal and vertical components are not exactly in phase. The end of a vector representing the instantaneous strength of the electric field will therefore travel round on the elliptical path as shown in fig. 13b. For given values of $\varkappa$ and σ, the tilt angle increases with frequency, so that a high-frequency wave suffers greater absorption than a low-frequency wave.

Attenuation

44. From the foregoing considerations, it will be readily understood that after its inception, the subsequent history of a radio wave largely depends upon the properties of the various media through (and over) which it must travel. These properties affect the direction of the propagation, the strength of the received signal and the polarization of the wave. As the wave travels outward from its point of origin, the strength of the field decreases and the wave is said to be attenuated. The attenuation may be considered as due to the three following causes.

(i) The natural spreading of the wave. This is of the same form, geometrically, as the spreading of an electric field from a point of charge ; if a transmitting aerial were situated in free space, the wave front would be a spherical surface, and the energy density at any point distant x units of length from the radiator would be proportional to $\frac{1}{x^2}$. The strength of the field is proportional to the square root of the energy density and it therefore follows that the field strength, at a distance from the radiator, would be proportional to $\frac{1}{x}$. The reduction of field strength with distance may be referred to as the natural or geometrical attenuation.

(ii) The energy abstracted from the wave due to its passage through any medium possessing finite conductivity. This may be compared with the absorption of light. A piece of window glass may pass only 90 per cent. of the light reaching it, a portion of the remainder being reflected and a portion expended in heating the glass. A second, similar, piece of glass will pass only 90 per cent of that reaching it, so that the two together will pass only 81 per cent. and so on. This absorption follows the law of logarithmic decay, and may be referred to as logarithmic or exponential attenuation.

(iii) Finally, since no radio transmitter is situated in free space, but is either upon or comparatively near to the surface of the earth, attenuation is caused by the passage of the wave over the earth's surface, owing to the finite conductivity of the latter.

Early propagation formulæ

45. Considering the radiation from a vertical aerial of effective height h_e metres situated upon a perfectly conductive earth and carrying a uniform current of I amperes (R.M.S.) at a frequency of f cycles per second, the field strength at a distance of r kilometres, so small that the logarithmic attenuation is negligible, is given by the expression $\Gamma = \frac{4\pi}{10^4} \frac{h_e f I}{r}$ (microvolts per metre). With the frequencies in use in the earliest history of radio-communication, it was found that this formula gave inaccurate results at distances greater than a few miles. It will be noticed that according to the above expression the field strength should increase with frequency, whereas it was found that actually this was not the case, lower frequencies giving a better signalling range for a given value of f. After considerable experimental investigation a factor of the form $\varepsilon^{-Ar\sqrt{f}}$ was added in order to correct for the logarithmic attenuation, and the formula became

$$\Gamma = \frac{4\pi}{10^4} \frac{h_e f I}{r} \varepsilon^{-Ar\sqrt{f}},$$

various values of the constant A being proposed by various workers.

46. Before dismissing these formulæ it may be pointed out that whereas the first factor is derived theoretically on the assumption that the waves are propagated outwards in such a manner that the wave front at any instant is a hemispherical shell, the second factor was deduced empirically. Prior to about 1922 an expression of this type was generally used to forecast the performance of a given aerial; it will be seen that since the factor $\varepsilon^{-Ar\sqrt{f}}$ increases as f decreases, it appears that better signalling ranges should be obtained with low than with high radio frequencies. As a result of this deduction, long-distance communication was performed almost entirely on frequencies below 60 kc/s, some stations of very high power being operated on frequencies as low as 12 kc/s. Frequencies of the order of 1,000 kc/s and above were almost entirely neglected, or in certain instances deliberately adopted to restrict the signalling range for purely local working.

47. In this connection, however, it is unfair to be severely critical of the research workers of this period, for although it was recognized as early as 1912 that the future progress of radio-communication would depend upon the development of C.W. signalling and heterodyne reception, it was not until much later that the valve oscillator was sufficiently developed to be of value for long-distance transmission. Prior to this the C.W. generators available (such as the radio-frequency alternator and Poulsen arc) were either suitable only for low frequency (L/F and VL/F) working, or capable of operation at a much higher efficiency under these conditions. The employment of the lower radio frequencies for signalling purposes was therefore dictated as much by the generators available as by the supposedly superior propagation of these waves. When suitable C.W. oscillators became available, shortly after the end of the war of 1914–1918, systematic research work into the propagation of short waves was undertaken by service, commercial and amateur workers. The great value of the amateur contribution to this research was chiefly due to their world-wide distribution and elaborate though entirely voluntary organization for reporting and verifying the reception of signals transmitted by other amateurs. It was found that the foregoing propagation formulæ were of little value for frequencies higher than about 1,000 kc/s, the obtainable ranges being sometimes very much greater than predicted thereby. The mode of propagation of these waves was therefore seen to be more complex than had hitherto been thought.

48. As long ago as 1882, before the production of electro-magnetic waves by the direct action of electric currents had been achieved, it had been suggested by Balfour Stewart, as a result of the study of terrestrial magnetism, that at a height of about 200 kilometres, instead of being an almost perfect insulant as it is at ground level, the earth's atmosphere should have a conductivity of the same order as that of the earth itself. When, in 1907, Marconi succeeded in signalling between England and Newfoundland, this suggestion was recalled almost simultaneously by Heaviside in England and Kennelly in America. Both these scientists pointed out that if such a conductive region actually existed, wireless waves originating on the earth would be

unable to penetrate it, but would be reflected downwards and so would travel round the protuberance of the earth. Kennelly in particular stated clearly that the effect of such a stratum in the atmosphere would be to decrease the geometrical attenuation ; instead of the field strength varying inversely as the distance r, it would vary inversely as the square root of the distance. Accordingly, it was proposed to predict the probable range of a given transmitter by means of a formula of the form

$$\Gamma = \frac{4\pi\, h_e f I}{10^4\sqrt{r}}\,\varepsilon^{-Ar\sqrt{f}}$$

but this was also found to fail when applied to the propagation of high frequencies.

Nature of the earth's atmosphere

49. Before proceeding further, it is necessary to discuss briefly the nature of the atmosphere. This consists chiefly of nitrogen, water vapour and oxygen, together with traces of many other

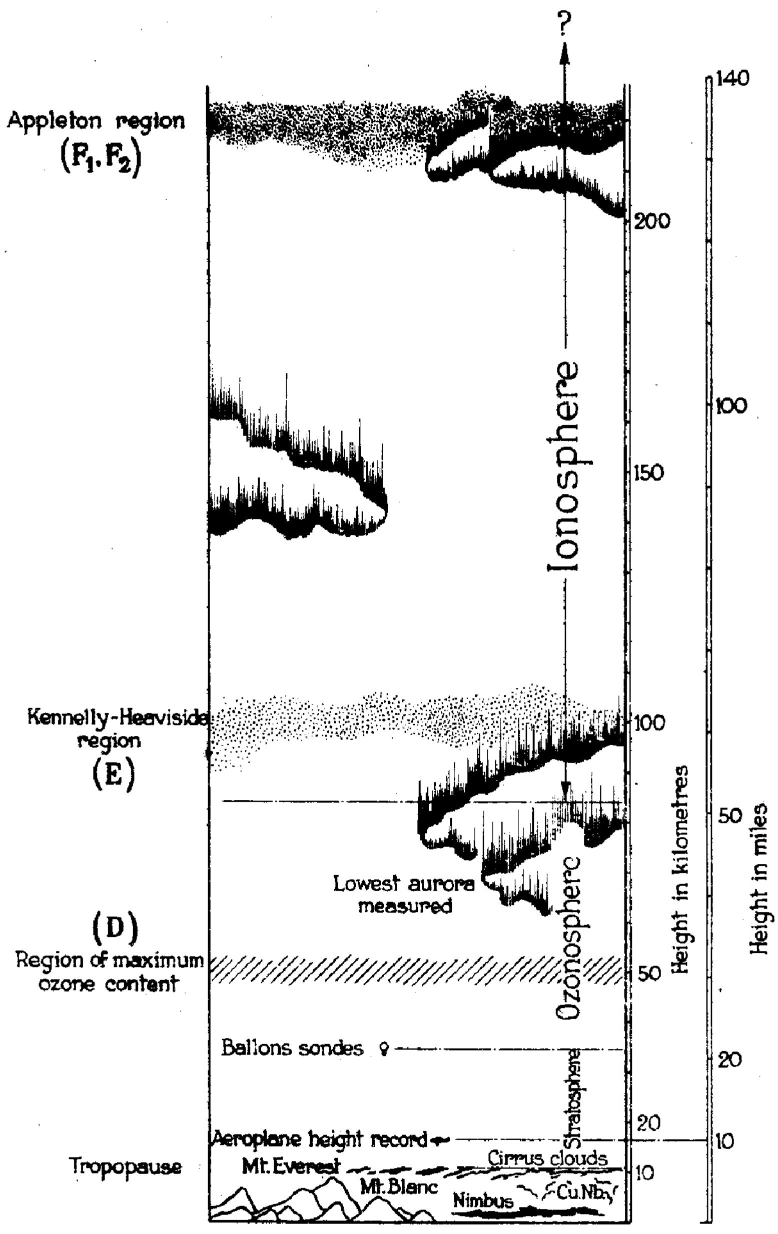


FIG. 14, CHAP. XIV.—The earth's atmosphere.

gases, and is divided into several regions which are shown pictorially in fig. 14. The region immediately above the earth's surface is called the troposphere. Its chief characteristic is the turbulent motion of the air. Winds, cloud formation, rain, etc., are all found in this strata. In the temperate zones the "roof" of the troposphere is about 10 kilometres above ground, and is called the tropopause. In the troposphere the temperature decreases with height, an average of about 6° C. per kilometre. Thus the temperature at the tropopause is about —50° C. over Europe. Over the equator its height is rather more—about 15 kilometres—and its temperature about —75° C.

50. Above the tropopause we have a region of about 20 kilometres deep, called the stratosphere. Human exploration of this region has so far been confined to its lowest levels, but a great deal of knowledge regarding it has been obtained by recording instruments carried by free balloons and by observations of the behaviour of meteors. At about 20 kilometres above ground the pressure is roughly one-twentieth that at the surface and the temperature over the temperate zones about —55° C., while it is lower still over the equator. The upper portion of the stratosphere is also called the ozonosphere, because a comparatively small quantity of ozone is distributed throughout its volume, over a height of from 40 to 60 kilometres. Although small in quantity this ozone absorbs a very large proportion of the ultra-violet radiation from the sun. The temperature of the ozonosphere is raised to some 40° C. as the result of this absorption of solar energy. It is probable that at such heights both oxygen and nitrogen are also present.

The ionosphere

51. The name "stratosphere" was originally given because it was formerly thought that in this region there is little or no wind, and the various gases therefore tend to separate into strata, the heaviest being of course nearest the earth. According to this theory there was little difference between night and day in this region, its temperature being nearly constant and of the order of —50° C. It is now known that the atmosphere at high levels is warmed during the day and cooled at night, so that winds are set up, at any rate up to heights of the order of 100 kilometres. The temperature in the higher portions of the stratosphere may in fact vary between +100° C. and —50° C. at different times of the day and year. The exact agency by which portions of this region become ionized is still in some doubt, but it is almost certain that the most important factor is the emission of ultra-violet rays by the sun. Under certain conditions, it is possible that particles of electrically neutral matter are shot out by the sun, and these may also play a part. At present, however, it is considered that the sun's ultra-violet radiation is sufficiently intense to account for the observed phenomena. The ionized region is usually referred to as the ionosphere. Assuming that the ionizing agent is ultra-violet radiation, it will be understood from fig. 15 that

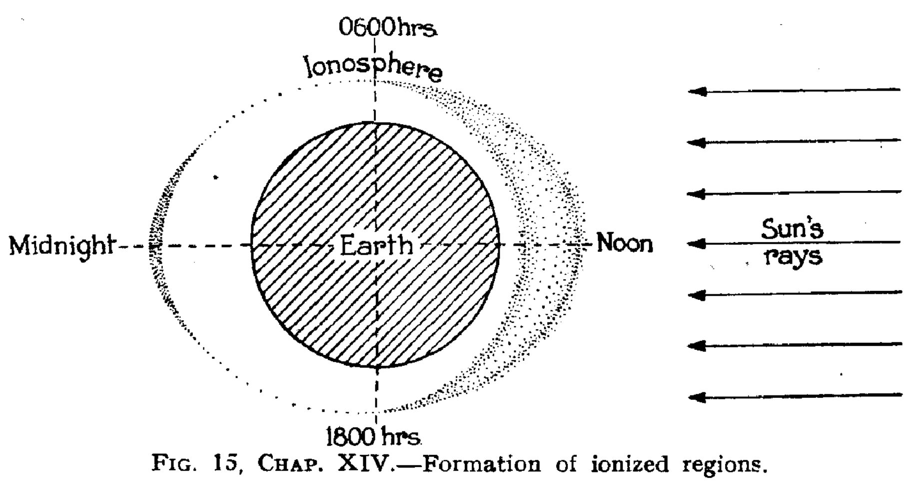


FIG. 15, CHAP. XIV.—Formation of ionized regions.

when any particular portion of the ionosphere is within the shadow of the earth, that portion receives little or no ionizing radiation. The positive and negative ions (or free electrons) into which it is separated during daylight, will therefore tend to re-combine during the dark hours. During daylight over any particular part of the earth then, the ionized region is of greater depth and its lower boundary (which is of course very ill-defined) is nearer the earth than during the night.

The ionized regions E and F

52. Direct evidence of the existence of conducting layers in the upper regions of the atmosphere was first obtained by Appleton and Barnett in 1925. In brief, the later experiments of this series consisted of the transmission of a very short train of waves—called a pulse—which was received at a point near the transmitter. As a rule it was found that on the transmission of a single pulse, three or more pulses were received, the first arriving at the receiver practically instantaneously with the transmission, owing to the proximity of the two stations. The other pulses were received at distinct intervals after the first, and in all cases the interval was consistent with the assumption that the wave had travelled upward until it reached a reflecting layer and had then returned. The first set of measurements, using a frequency of 750 kc/s, indicated that the reflector was situated about 90 to 100 kilometres above the surface of the earth. This phenomenon was expected, and the height at which the reflection occurred was in accordance with theoretical estimates. What the experimenters had achieved, at this stage, was to establish the existence and approximate height of an ionized region.

53. Research on these lines has now proceeded for a number of years and, as a result, we are able to speak with confidence of two regions of maximum ionic density. For brevity these regions are sometimes referred to as "layers," but it must be understood that they possess no definite boundaries. It is convenient to refer to the virtual height of a layer as the height at which signals appear to commence their return journey to earth, or at which the apparent reflection takes place. The critical penetration frequency of any region is defined as the lowest radio-frequency which will penetrate that region at normal (i.e., vertical) incidence. We may think of the two regions as concentric shells of ionized gas surrounding the earth, the inner, which is known as the "E" or Kennelly-Heaviside layer being at a distance of about 100 kilometres above the earth, and the outer, which is known as the "F" or Appleton layer, at an average height of about 230 kilometres. Under certain conditions each of these may be split into two regions of maximum ionic density. We then speak of the E_1 and E_2, F_1 and F_2 layers, E_2 being slightly higher than E_1 and F_2 slightly higher than F_1. In all cases, the virtual heights are referred to.

54. The mechanism by which ionization is set up is thought to be somewhat as follows. Assuming that the principal cause of ionization is the ultra-violet radiation of the sun, we have to consider the absorption of energy which ionization implies. At very great heights there is very little gas present so that although the radiation is intense, very few ions are formed, and very little energy is lost by the radiation. As the latter approaches the earth the gas pressure increases and ionization commences, but since the mere fact of ionization causes loss of energy, the intensity of the radiation falls off rapidly as it approaches the earth. We therefore find the state of affairs to be somewhat as shown by the solid curve of fig. 16, which is drawn on the assumption that the ionizing radiation has a single frequency and that only one kind of gas molecule is present. Actually, of course, this is not the case, and there may be several regions of maximum ionization. In any case the ionizing influence exerted at a lower level, e.g., in the E region, will depend upon the amount of solar radiation absorbed at higher levels.

55. As previously suggested, ions and electrons in a dissociated state have a strong tendency to combine. In the ionosphere, the rate of re-combination depends mainly upon the gas pressure. This is of course much higher at 100 kilometres than at 230 kilometres and consequently the rate of re-combination is much higher in the E than in the F layer. In addition to the E and F regions, it appears probable that portions of the ozonosphere also become ionized, although to a considerably smaller degree than the upper regions. The existence of this region—which has been allotted the letter D—may account in part for the propagation of a ground wave at the lower radio-frequencies, and its virtual disappearance at the higher frequencies. It is probable that the D layer ionization is due in part to the bombardment of gas molecules by cosmic rays.

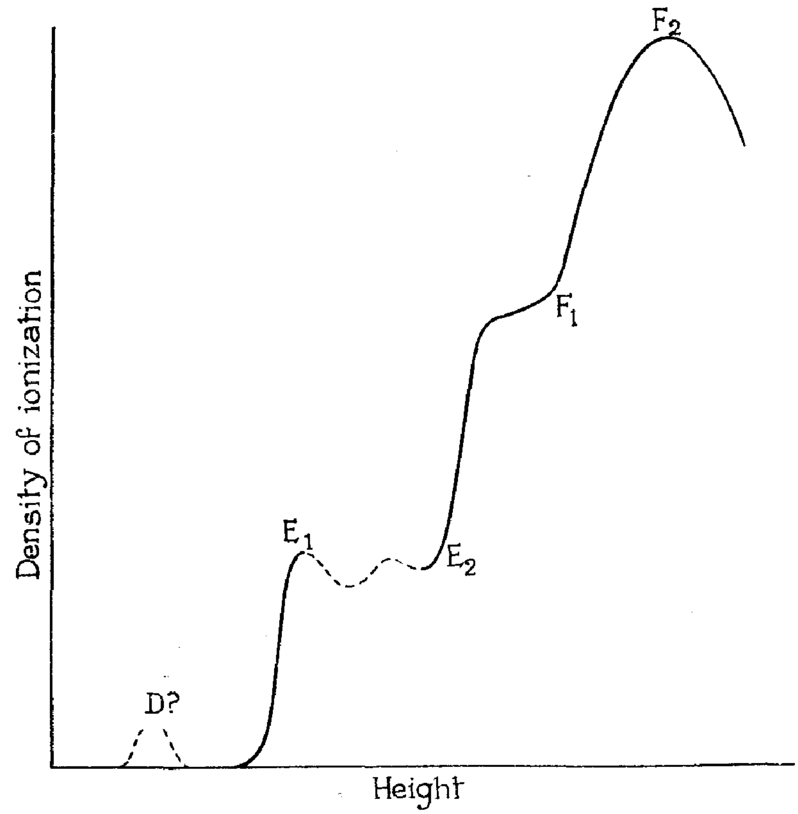


FIG. 16, CHAP. XIV.—Variation of ionization with height.

Seasonal, geographic and solar influences

56. The virtual height of an ionized layer is an important characteristic in that it determines the maximum frequency at which waves will be reflected. For a given transmission distance the angle of incidence at a layer will be greater, the lower the layer. The critical frequency for a given ionization density varies approximately as the secant of the angle of incidence, so that for a given density and signalling distance, a lower layer will reflect waves of a higher frequency. Thus, even though the ionization density of the E layer is considerably lower than in the F layer, the former may on occasion determine the maximum frequency which may be usefully employed. The state of the ionosphere also varies with the latitude ; in general, the density of ionization is greatest in equatorial regions. It also varies seasonally, being greatest in the summer and least in winter. Because the seasons are less clearly defined in equatorial regions than in others, the seasonal variations are of less practical importance than in northern or southern latitudes. In addition to variations associated with day and night conditions, latitude and seasons, progressive changes in the ionosphere occur during the period of the sun-spot cycle—approximately eleven years. The ionization density is highest when the sun-spot activity is a maximum ; it is now increasing and may be expected to reach a maximum in the year 1939. Observations have shown that certain solar disturbances are associated with pronounced variations which tend to repeat at intervals of 27·3 days, this being the period of rotation of the sun. Whether the solar disturbance and the ionization variations are related in the manner of cause and effect, or whether both are due to some unknown cosmic phenomenon, is at present an open question.

Change of polarization

57. Referring to paragraph 33, is will be appreciated that the earth's magnetic field will have a considerable influence upon the passage of a wave through the ionosphere. During its

passage, the wave may be split into two circularly or elliptically polarized components, one having a right-handed and the other a left-handed sense of rotation, looking in the direction of propagation. The former is referred to as the extraordinary and the latter as the ordinary ray. When such splitting occurs, it is necessary to distinguish between the critical frequencies of the two components, and the following notation is used. The symbol f_c is used to denote the critical frequency in general. The critical frequency for the ordinary ray in the F_1 region is denoted by $f^o{}_{F1}$, and for the extraordinary ray $f^x{}_{F1}$. Similarly in the F_2 region the critical frequencies are denoted by $f^o{}_{F2}$ and $f^x{}_{F2}$. The critical frequency for the E region is denoted by f_E, the electrical and magnetic characteristics at this height being such that double refraction (i.e., splitting) occurs only infrequently. In the northern hemisphere the right-handed component, i.e., the extraordinary ray, is absorbed to a greater degree than the left-handed component, and *vice versa* in the southern hemisphere. The greater part of the absorption appears to take place in the lower regions, where the atmospheric pressure is fairly high, the collision frequency large and the ionic density small.

58. Both the absorption and the polarization depend upon the magnitude of $\frac{e\,H}{m\,c}$ where e is the charge of an electron in E.S.U., m its mass in grams and $c = 3 \times 10^{10}$. The strength H of the earth's magnetic field varies from about 0·35 oersted at the equator to about 0·6 oersted at the poles. By making certain simplifying assumptions it is possible to calculate, approximately the frequency f_r at which the electronic oscillations exhibit quasi-resonant properties. Thus if $H = 0{\cdot}4$ oersted

$$f_r = \frac{1}{2\pi}\,\frac{e\,H}{m\,c}$$

$$= \frac{4{\cdot}77 \times 10^{-10} \times 0{\cdot}4}{6{\cdot}28 \times 8{\cdot}8 \times 10^{-28} \times 3 \times 10^{10}}$$

$$= 1{\cdot}15 \times 10^{6} \text{ cycles per second.}$$

At this frequency, a plane-polarized wave having its electric field vector perpendicular to H, and propagated along the axis O Y in fig. 10a, would set the electrons in the ionized region into violent oscillation and numerous collisions would occur between electrons and gas molecules. This would give rise to severe absorption and consequently the frequencies in the quasi-resonant band are but poorly propagated *via* the ionosphere. According to the direction of propagation with respect to the magnetic field, and the latitude, it is found that the frequency band 1 to 2 Mc/s (approximately) suffers almost complete absorption owing to this phenomenon, so that communication by this band must depend chiefly upon the ground wave.

Fading

59. Fading is observed to some extent at all frequencies, but is more pronounced at high frequencies than at low. The term fading is generally applied to variations of signal strength, occurring over only a short interval of time, e.g., a few minutes, but an allied phenomenon is the complete cessation of communication on a given frequency for many hours or even days, after which communication is restored. The latter form of fading often occurs when the transmission path passes near the magnetic poles, and thus is believed to be connected with the mechanism of terrestrial magnetism. The latter phenomenon in turn is associated with the sun-spot activity; it has been observed that when a sun-spot passes over the sun's meridian, as viewed from the earth, a period of fading is generally experienced at a later epoch, which may be from one to three days. Rapid fading, on the other hand, is generally attributed to interference between waves arriving at the receiver by paths of different lengths, and by a variation of the nature of polarization. The effect of interference is similar to that caused by reflection at the surface of the earth (paras. 34–40), but is complicated by the fact that the ionization along the path of the wave through the ionosphere is subject to slow fluctuations, while the respective planes of polarization of the ground and sky waves may also change. If a vertical aerial is in use at a

ground station, the strength of the received signal depends only upon the strength of the vertically polarized component, and if the degree of vertical polarization changes, the signal strength will fluctuate accordingly.

60. About three years ago it was noticed that at certain times, long distance H/F communication would completely fail for periods of 15–30 minutes, sometimes 45 minutes, over the whole of the surface of the earth that was illuminated at the time by the sun. It was found by astronomers that this fading occurred simultaneously with " bright eruptions " of hydrogen gas from the surface of the sun in the neighbourhood of a sun-spot. These bright eruptions produce an intense ultra-violet radiation which is believed to cause a great increase in the intensity of ionization in the E layer. It is probable that the increased attenuation in the E layer is the cause of this kind of fading. During such a period, the transmission on L/F and VL/F is sometimes improved. It was thought at first that this type of fading is subject to a 54-day periodicity, but recent evidence has not confirmed this. It may be expected that H/F fading will occur more frequently and increase in intensity until the sun-spot maximum (1939). It will probably decrease to a minimum about 1944 and so on with the eleven year sun-spot period.

Propagation of low-frequency waves (L/F and VL/F)

61. When the frequency of the signal is below about 300 kc/s, and the range of reception does not exceed about 1,000 miles, most of the energy reaches the receiver by the ground wave, and the season and time of day have little effect. On the other hand, at ranges of from 2,000 to 4,000 miles or so, the energy received is practically confined to that carried by the sky wave, and the strength of signals shows pronounced diurnal and seasonal variations owing to the changes in location and density of the E layer. Between 1,000 and 2,000 miles, both the sky and ground waves are of appreciable strength. Both these are emitted in the same phase, but they do not traverse paths of the same length, and do not, except fortuitously, arrive at the receiver in phase ; the field at the receiving aerial is the vector sum of the two, and if the effective length of either path varies, the signal strength will also vary. When the whole of the transmission path is in daylight, the signal strength of a very distant transmitter is relatively low, but more or less constant, but as the sunset line falls across the transmission path, the field strength generally falls ; when however the whole path is in darkness the intensity usually increases to a high level, which is maintained until sunrise.

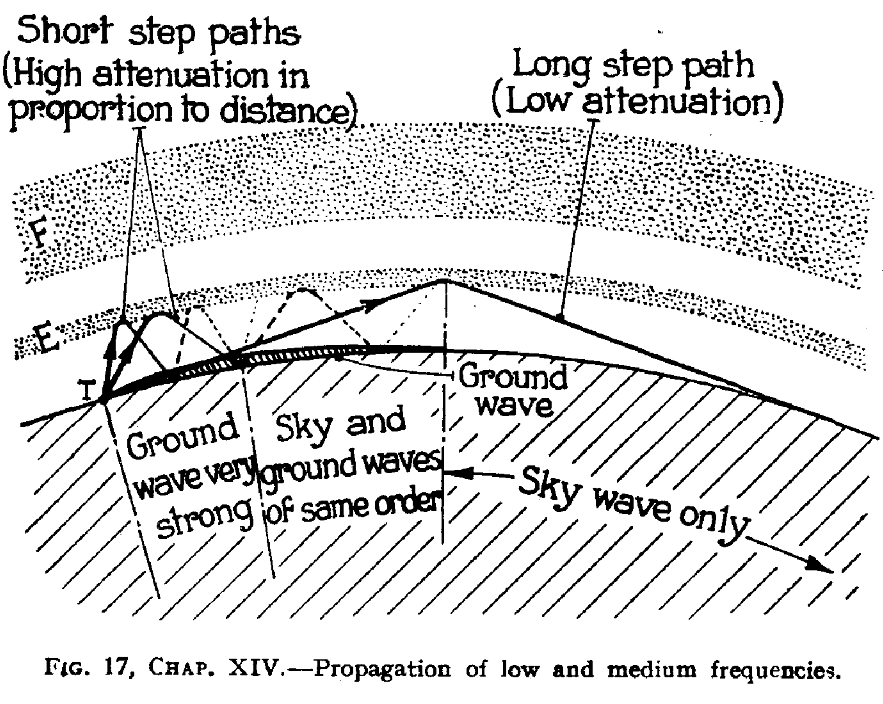


FIG. 17, CHAP. XIV.—Propagation of low and medium frequencies.

62. The propagation of these waves is illustrated in fig. 17 which shows the route taken by several hypothetical rays, according to the vertical angle of radiation at the transmitter. At such low frequencies, the E layer is highly refracting, and the upward-travelling waves are bent downward towards the earth after relatively little penetration, that is, the bending process resembles the reflection of a light wave. The wave, on reaching the earth, is again reflected, and travels upward once more, much energy being lost by absorption at each reflection. The energy which travels in the space between the layer and earth suffers only geometrical attenuation, and consequently, the greater part of the energy received at extreme ranges is that which leaves the transmitter at a low vertical angle and is reflected at the layer only once. It is not, however, possible to design a low-frequency aerial in such a manner as to concentrate the energy in the desired direction, for this would necessitate an aerial of vertical height equal to several half-wavelengths.

63. The diurnal and seasonal variations in signal strength result from variations in the height and density of the E layer; when the layer is low, or the electron density high, the attenuation will be high. As the frequency is increased, these variations become more marked, until, at frequencies of the order of 500 kc/s, long-distance transmission is not possible during daylight and is subject to enormous fluctuation during dark hours. As already observed, low-frequency waves travelling near the surface of the earth are always normally polarized, and the wave front is tilted forward slightly as a result of the absorption of energy by the conductive earth. Fading is rarely observed during daylight but may be troublesome when signalling over extreme ranges at night.

The 600–1,500 kc/s band

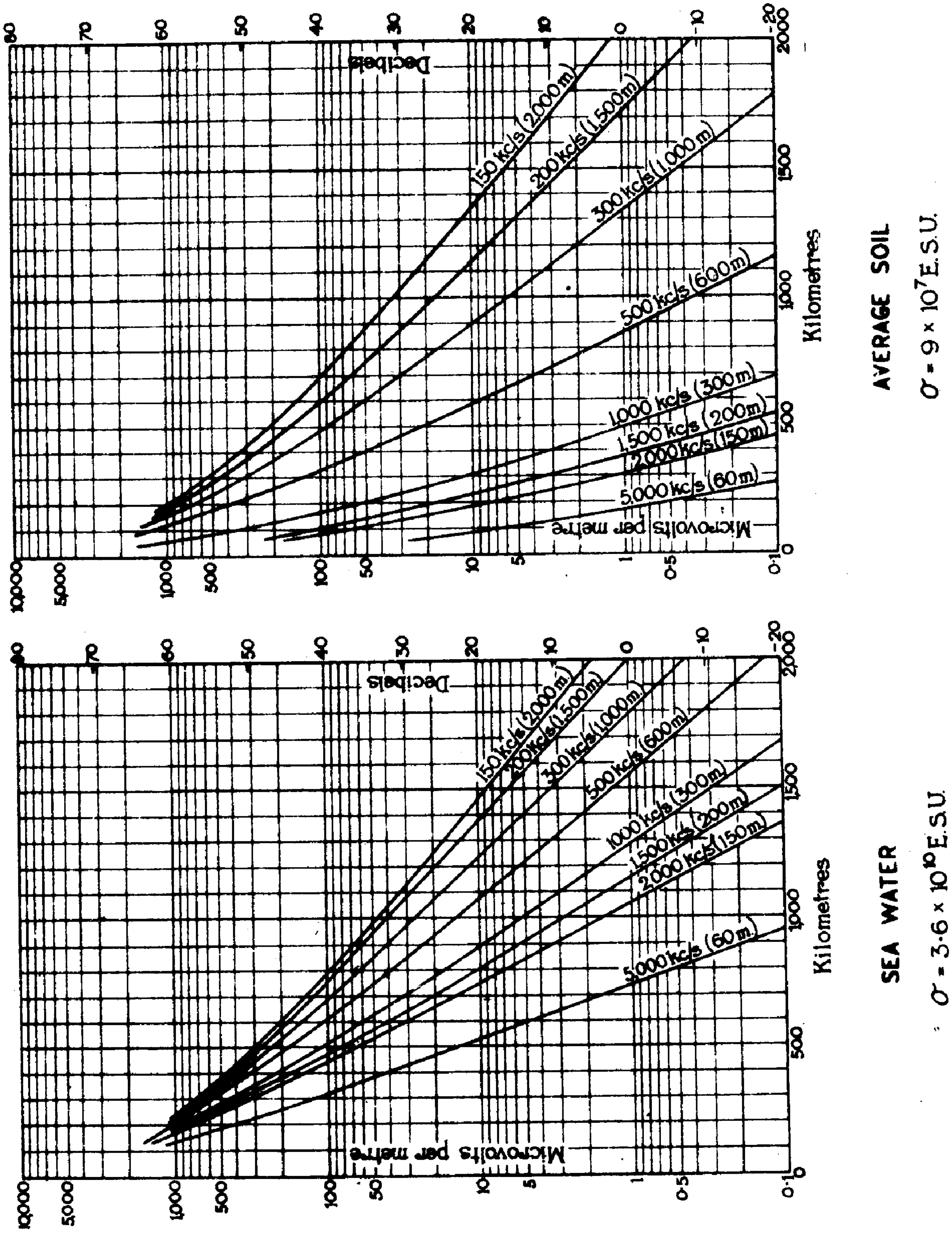
64. These frequencies are chiefly used for the broadcasting of entertainment programmes, and their propagation has been studied by a very large number of investigators. The chief point to be observed is that with this type of transmission, what is required is a comparatively loud, and very constant, signal over a region within a few hundred miles of the transmitter. The region in which reasonably high-quality telephony can be received is that in which the ground wave is powerful enough to give a high degree of discrimination against all interference. The field strength required in industrial districts is much higher than in rural ones, owing to the higher noise level: Experience indicates that for satisfactory service a field strength of from 5 to 30 millivolts per metre is necessary in towns, whereas a field of only ·1 millivolt per metre may suffice in the country.

Field-strength curves (ground wave)

65. Fig. 18a shows the variation with distance in the intensity of the ground wave, when the transmission is wholly over sea water. The left-hand scale gives the field strength for a radiated power of one kilowatt; if P kilowatts are radiated the field-strength scale must be multiplied by $\sqrt{P}$. The right-hand scale gives the field intensity in decibels with reference to an arbitrary level of 1 micro-volt per metre. If P kilowatts are radiated the quantity $10 \log_{10} P$ decibels must be added to this scale. For example, at a frequency of 1,000 kc/s the field strength at a distance of 900 kilometres is 10 micro-volts per metre, which is 20 db. above the reference level. A radiated power of 4 kilowatts gives a field strength of 20 micro-volts per metre, which is $(20 + 10 \log_{10} 4) = 26$ db. above the reference level. Fig. 18b is similar to fig. 18a except that transmission is assumed to take place over soil of average conductivity. The field strengths actually obtained may be greater or less than those given by the curves, owing to the possible presence of the sky wave as explained above.

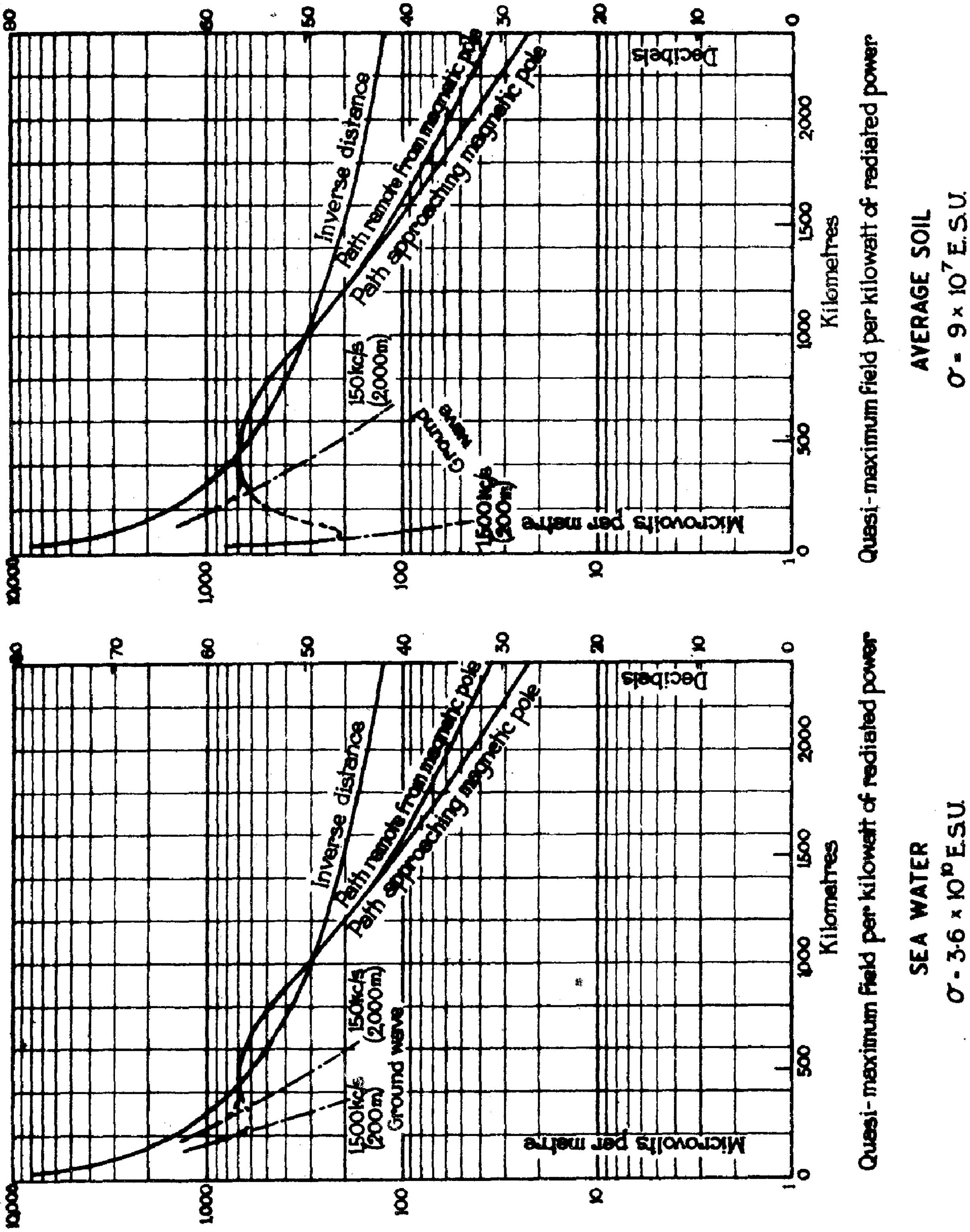
Field-strength curves (night propagation)

66. When the whole of the transmission path is in darkness, the curves given in fig. 19 may be used to estimate the probable intensity at the receiver, for all frequencies between 150 and 1,500 kc/s. The two sets of curves correspond with those of fig. 18 (for sea water and average soil respectively). The quantity plotted is referred to as the quasi-maximum field strength (per kilowatt radiated). This is defined as the strength which is exceeded by the actual field strength for only five per cent. of the time. The average field strength is, however,



DAYLIGHT PROPAGATION (GROUND WAVE)

FIG. 18
CHAP. XIV



NIGHT PROPAGATION

FIG. 19
CHAP. XIV

only about 35 per cent. of the quasi-maximum value. It will be seen that in each graph there are three curves. One, the inverse distance curve, gives the field which would be obtained on a flat earth with geometrical attenuation only, and is inserted merely for comparison. If the transmission is substantially east to west or *vice versa*, in other than equatorial latitudes, the great circle path between transmitter and receiver may approach the magnetic pole, and the absorption will be greater than if the path is remote from the pole. The latter condition corresponds to transmission in the north-south or south-north directions. The ground-wave intensity is approximately shown by the chain-dotted curves. Where the intensities of ground and sky waves are of the same order, the conditions are very variable and the curves are shown in broken line.

Propagation of high-frequency waves (1,500 to 30,000 kc/s.)

67. At frequencies above 1,500 kc/s. the ground wave is rapidly attenuated ; long-distance communication is dependent entirely upon the sky wave, and thus upon the frequency, the height of the reflecting layer, and the electron density in it. The angle at which the wave enters the layer is also of primary importance. The most striking features involved are illustrated in fig. 20,

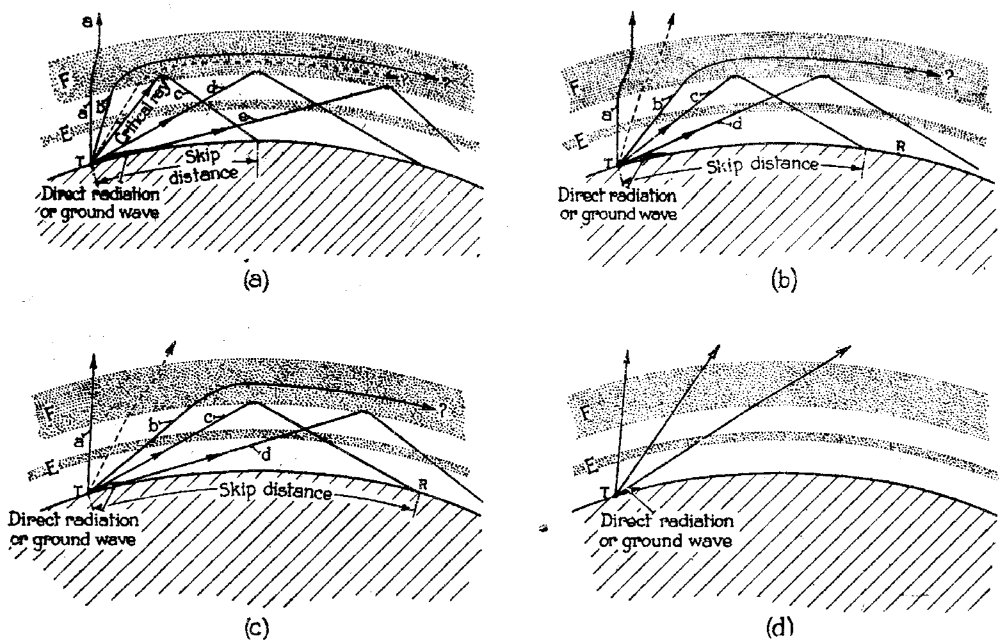


FIG. 20, CHAP. XIV.—Propagation of high frequencies.

which shows the paths of rays of different frequencies emitted at various angles to the vertical. These rays are shown as penetrating the E and reaching the F region. In fig. 20a, the frequency is comparatively low, i.e., in the neighbourhood of 1,500 kc/s. The ray a which is incident almost perpendicularly upon the layer, is subjected to little bending, and passing the region of maximum ionization, escapes past the upper boundary of the layer. Ray b reaches the layer at such an angle that it is just bent parallel to the earth when it reaches the region of maximum density. After travelling parallel to the earth for some distance, it may reach a portion of the layer of greater density at the same height. The ray is then bent downwards and will return to earth. The ray c which leaves the transmitter at a still smaller angle to the ground does not penetrate so far as the region of maximum density but is bent downward much earlier, returning to earth comparatively near to the transmitter. No rays between a and b will return to earth, and in the region between the transmitter and the point at which c returns, only the ground wave will be effective in giving any signal. The ground wave, however, is heavily attenuated, and as a result there is a considerable region in which no signal is received at all ; the distance thus " skipped over " by the sky wave is called the skip distance.

68. In fig. 20b, the frequency is assumed to be somewhat higher, and the refracting property of the layer at this frequency is less than in the previous instance. Ray C, which just reaches the layer and is immediately bent downwards, reaches the ground at a greater distance from the transmitter, i.e., the skip distance is greater. In fig. 20c, a further increase of frequency results in the passage through the layer of almost all the high-angle radiation and ground absorption of almost all the low-angle radiation. The only energy reaching the distant receiver is that reaching the layer at an angle only just above the horizontal. Finally in fig. 20d, the frequency has been increased to such an extent that the F layer is practically non-refracting. No energy is returned to earth by the layer and no long-distance communication is possible. The highest frequency suitable for daylight long-range communication is about 25 Mc/s.

69. It is now seen that for any given frequency there must be a certain critical angle ; radiation at an angle greater than the critical, passes through the layer and is lost. The critical angle becomes smaller as the frequency increases, because the refracting property is inversely proportional to the frequency. The effect of the electron density is the reverse ; the higher the density, the smaller is the critical angle. The skip distance depends upon the critical angle and upon the height of the layer ; it increases as the frequency is raised.

Skip distance/frequency curves

70. It is impossible to state any simple relation between frequency and skip distance. As, however, this question is of first importance in H/F signalling, continuous endeavours have been made to correlate theory and the observed facts, and to produce graphical information which will be of service to those responsible for frequency allocation for different forms of communication. It must be emphasized that such graphs cannot predict with complete certainty the frequency which will give reliable communication between any two points, under any specified conditions of time, date and latitude. Besides the ionospheric variations of a more or less predictable nature already referred to, there are day-to-day changes which appear to be very erratic. Thus a graph drawn to illustrate skip distance/frequency relations over a period may give a rough average picture of these relations for the periods in question, but there will inevitably be certain days when the actual relation differs considerably from those indicated by the graph. Subject to these limitations, the graphs shown in figs. 21 to 33, may be used. These cover the total range of latitude from 65° north to 65° south. The actual zone covered by each graph and the seasons for which it is applicable are shown in a " key page " preceding the graphs. Each graph is made up of seven distance curves marked $d = 0$, $d = 200$ etc., up to $d = 1200$. The distance d is in miles. The ordinate of each graph is marked in frequency (Mc/s) and the abscissa represents local time. It will be observed that five graphs are required for all latitudes other than the zone 15° north to 15° south, which is covered by a single graph for all months of the year. The material on which the graphs for latitudes north and south of 55° are based was very scanty, and these are considered to be less reliable than the remainder.

71. If a short-wave transmitter is communicating with a receiver d miles away, there will be one particular frequency at which waves sent out by the transmitter, and reflected by the ionosphere, will just graze the distant receiver. This is the skip frequency for the distance d miles. With reference to fig. 20c, if the transmitter T is sending on a frequency of f Mc/s and R is the point at which the descending skywave first reaches the earth, then f Mc/s is the skip frequency for the distance T R. If the transmitter T uses a frequency much below f Mc/s, the point R should be well within the zone of the downcoming skywave. Conversely, if the transmitter uses a frequency higher than f Mc/s, the downcoming skywave will strike the earth beyond the point R and no signals will be heard at that point. Local time, at any place, is the time given by a clock set to 1200 hours at the instant when the sun crosses the meridian up to that place. The local time, at any place, may be determined by means of a map, bearing in mind that the change of time for each degree of longitude is 4 minutes ; an example of this determination is given below.

Example :—What is the local time at Alexandria, Egypt, when it is 1500 hours B.S.T. in London ?
(1500 hours B.S.T. is 1400 hours G.M.T.)

Reference to a map shows that Alexandria is 30° East of Greenwich and therefore the sun crosses the meridian through Alexandria 30 × 4 minutes or two hours before it crosses the meridian through Greenwich. 1400 hours G.M.T. at London therefore corresponds to 1600 hours local time at Alexandria, i.e., when it is 1500 hours B.S.T. in London the local time in Alexandria is 1500 hours.

72. When using the skip distance graph to find a suitable frequency for communication between two points such as T and R (fig. 20) the local time to be used is strictly the local time at a point midway between T and R, and not the local time at either T or R. In practice it is seldom necessary to make any correction for the difference in time between the ends and the middle of the signalling path, for even in the worst case—with T and R on the 65th parallel of latitude and 1,200 miles apart—the time difference does not exceed 1·36 hours.

73. The graphs have been so planned that it is possible to use them to estimate approximately (for distances up to 1,200 miles) the skip distance for any given frequency, or alternatively the skip frequency for a given distance. It is probable that the graphs will be used mostly as an aid in finding the most suitable frequency for communication between two known points under specified conditions of time and date, but attention must again be drawn to the fact that it is not possible to predict with certainty the most suitable frequency for reliable communication. If the signalling is to take place between two stations whose geographical position is known, the skip frequency for the distance of separating the stations under specified conditions of time and place may be estimated approximately from the appropriate graph. The most suitable frequency for signalling will, however, be below the real skip frequency, for when a transmitter T is sending to a distant receiver at R (fig. 20), the receiver must not be just on the edge of the skip zone, but must lie within the zone of the downcoming skywave, as in fig. 20b. At first sight it might appear that any frequency well below the skip frequency would be suitable for the desired communication. This is not necessarily so, for the attenuation increases as the frequency decreases, so that if too low a frequency is employed, communication may fail because of the excessive attenuation.

74. Where short-wave communication must be established along a certain route and there is local information based on practical experience, concerning the behaviour of short waves along that route, or other routes similar to it, every endeavour should be made to check deduction from the graph by information derived from practical experience. When no local information is available, deduction from the graph should be a valuable guidance to the order of frequency likely to be most useful.

75. In addition to their use as an aid in determining the most suitable frequency for signalling under certain specified conditions, the graphs indicate how frequency conditions along a signalling path are likely to vary with the time of day, e.g., if communication between two points 600 miles apart is satisfactory from 0600 hours to 1000 hours on a certain frequency and then begins to fail, the general trend of the $d = 600$ curves, on the appropriate graph, will indicate whether an increase or decrease of frequency is more likely to enable communication to be satisfactorily maintained after 1000 hours.

Numerical examples

76. Examples illustrating the use of the skip distance frequency graphs are given below. In the earlier examples, no mention is made of any correction to allow for variation of ionospheric conditions with the solar cycle ; the method of making such correction is explained later.

Example 1

In fig. 20c, let T represent a short-wave transmitter situated in a place on latitude 20° S., and the distance T R be 400 miles. What will be the skip frequency for the distance T R at 0300 hours, 1000 hours and 1950 hours on a day in November ? What frequency would you recommend for communication between T and a receiver at R from 0930 to 1100 hours ?

CHAPTER XIV.—PARA. 76
(Example 2)

(*a*) The skip frequency for any given distance *d* and time is found by reading off the height in Mc/s of the point where the vertical line corresponding to the time cuts the appropriate *d* curve for the distance in question. The graph to be used for this problem is No. 6, and the vertical line through 0300 hours cuts the $d = 400$ curve at a height of 5·75 Mc/s. This is the skip frequency at 0300 hours; by applying the same method to the other two times, it will be found that at 1000 hours the skip frequency is 13·5 Mc/s, and at 1950 hours the skip frequency is 15·25 Mc/s.

(*b*) Reading from the $d = 400$ curve, it will be seen that the skip frequency between 0930 and 1100 hours varies between about 13·2 and 13·5 Mc/s. For reliable communication it is essential to use a frequency below the actual skip frequency; for best results the distant receiver at R must be well within the zone of the downcoming waves, and not just on the edge of the skip zone. On the other hand, the frequency chosen must not be too low; otherwise, if the waves passing from T towards R through the ionosphere are being reflected from the F region, they may be so heavily attenuated that they never reach R at all, or they may reach R but be too weak to give readable signals in the receiver.

(*c*) It is quite impossible to lay down any rule for the relation between skip frequency and best signalling frequency, but, if no local information is available as a check on results deduced from the skip distance graphs, it would be advisable to try as a signalling frequency, a frequency about 20 per cent. below the deduced skip frequency. Applying this to the case now under consideration, where the skip frequency is 13·2 to 13·5 Mc/s, a frequency of about 10·6 Mc/s (13·2 Mc/s less 20 per cent.) would be recommended as likely to give communication between 0930 and 1100 hours.

Note.—If the frequency 20 per cent. below the deduced skip frequency gave communication, but with signals weaker and less reliable than might reasonably be expected, the indication might be either that the distant receiver was too near the edge of the skip zone, or that excessive attenuation was occurring. Matters might then be improved by an increase or a decrease of frequency, and only experiment could decide which of these two courses would be the better to adopt. If the frequency 20 per cent. below the deduced skip frequency gave no communication at all, the probable indication would be that the deduced skip frequency was too high, and that trials could then be made with frequencies 30 or even 40 per cent. below the deduced skip frequency.

Example 2

A short-wave transmitter situated on a latitude 32° N. sends on a frequency of 16 Mc/s (18·75 m.) at noon on a certain day in January. What skip distance would you expect with this frequency? Estimate the probable skip distances for frequencies of (*a*) 12 Mc/s (25 m.), (*b*) 9 Mc/s (33·33 m.) and (*c*) 6 Mc/s (50 m.), time and date as for the frequency of 16 Mc/s.

Graph No. 7 must be used for this problem. If a point is plotted whose X ordinate is 1200 hours and whose Y ordinate is 16 Mc/s, it will be found to lie on the curve marked $d = 800$. This means that the skip distance for the frequency of 16 Mc/s, is 800 miles at noon (1200 hours).

(*a*) The point whose co-ordinates are 1200 hours and 12 Mc/s lies about midway between the curves $d = 400$ and $d = 600$ from which it follows that the skip distance for 12 Mc/s at noon is about 500 miles.

(*b*) The point whose co-ordinates are 1200 hours and 9 Mc/s lies between the $d = 200$ and $d = 400$ curves and is nearer the $d = 200$ curve. By marking in the point and estimating its position relative to both curves, it will be seen that the skip distance for 9 Mc/s at noon is of the order of 250 miles.

(*c*) If the point whose co-ordinates are 1200 hours and 6 Mc/s is plotted, it will be found to lie below the $d = 0$ curve. This means that at noon the frequency of 6 Mc/s is below the critical frequency of the ionosphere, so that all 6 Mc/s waves sent out by the transmitter, including those travelling vertically upwards, are reflected back to earth. At noon, therefore, the skip distance for the frequency of 6 Mc/s is zero.

Example 3

A short-wave transmitter situated on latitude 20° N. sends on a frequency of 5 Mc/s (60 m.) throughout the whole of a 24 hours' period in February. How will skip distance vary with time throughout the 24 hours ?

If a ruler is laid across graph No. 3, parallel to the time axis, and at a height of 5 Mc/s, it will be seen that the horizontal line for 5 Mc/s starts on the $d = 0$ curve, cuts twice across the $d = 200$ curve, then cuts across the $d = 0$ curve and remains below this curve thereafter. Translating these results into words, there is no skip distance at midnight, but from midnight to about 0315 hours skip distance increases steadily until it reaches a maximum value of about 220 miles at 0315 hours. It then starts to decrease and by 0645 hours has fallen to zero. From 0645 hours to midnight skip distance remains at zero, i.e., there is no skip effect during these hours.

Example 4

A comparatively low power short-wave transmitter in Palestine is working in January to a receiver situated 700 miles east of it. Communication is opened on a frequency of 4·5 Mc/s at midnight and remains satisfactory until about 0400 hours, when the distant receiver reports that signals are falling in strength and becoming unreadable. To what can this falling-off of signals be attributed, and would an increase or decrease of wavelength be advisable if it be required to continue communication until 0700 hours ?

From a reference to a map and the time of the year (January) it is obvious that graph No. 7 must be used for this problem. By sketching in a "d" curve for 700 miles intermediate between the curves $d = 600$ and $d = 800$, it can be seen that the skip frequency for 700 miles is about 6·5 Mc/s at midnight, after which it increases steadily. At 0400 hours it has reached nearly 9 Mc/s and is rising fairly steeply. Since communication is satisfactory up to 0400 hours on the frequency of 4·5 Mc/s, the indication is that the falling-off of signals which starts at about 0400 hours is not the result of using too high a frequency; i.e., the downcoming waves are not coming down to earth beyond the receiver. It may, therefore, be assumed that the falling-off of signals is due to the fact that the frequency of 4·5 Mc/s is being increasingly attenuated in its passage through the ionosphere, so that in order to maintain communication from 0400 to 0700 hours, it would be advisable to change to some higher frequency nearer to the actual skip frequencies of 9 to 12·5 Mc/s.

Example 5

Communication is required between a short-wave station at Bombay (Lat. 19° N.) and another similar station at Calcutta, which is approximately due east of Bombay. Both stations have an aerial power of about 0·5 kilowatts. What frequencies might be tried as likely to give good results from 0900 to 1200 hours, and from 2100 to 2400 hours (Bombay local time), during the months of May, June, and July ?

Calcutta is approximately 1,100 miles from Bombay, from which it differs in longitude by about 18°. There is, therefore, a time difference of 72 minutes, and the local times at the mid-point of the signalling path corresponding to the Bombay times specified above will be 0936 to 1236 hours and 2136 to 0036 hours.

(*a*) By using graph No. 6 and interpolating between the $d = 1{,}000$ and $d = 1{,}200$ curves, the skip frequencies for different times within the required first signalling period are found to be approximately as follows :—

0936 hours—20·5 Mc/s.
1036 hours—21·0 Mc/s.
1136 hours—20·75 Mc/s.
1236 hours—20·0 Mc/s.

It is obvious that some frequency below 20 Mc/s is indicated as the signalling frequency, and, if no local information were available as a check on the figures above, a frequency of 16 Mc/s (20 Mc/s less 20 per cent.) might be tried for communication during the period in question (but also *see* Note to Example 1).

(*b*) For the second signalling period the skip frequencies deduced from the graph are as follows :—

2136 hours—22·5 Mc/s.
2236 hours—20·0 Mc/s.
2336 hours—17·0 Mc/s.
0036 hours—13·75 Mc/s.

(*c*) The rapidly falling skip frequency makes it unlikely that, with the small power available, one frequency could be used successfully for the whole period under consideration. Communication might probably be started on a frequency of about 16 Mc/s (20 Mc/s less 20 per cent.), and if this were found to fail, very possibly between 2230 and 2330 hours, a change could then be made to a frequency in the region of 11 Mc/s, i.e., a frequency about 20 per cent. below 13·75 Mc/s (also *see* Note to Example 1).

Corrections for sun-spot cycle

77. It was stated in an earlier paragraph that ionization density changes progressively throughout the period of a solar cycle. As a solar cycle proceeds from a condition of minimum sun-spot activity to a condition of maximum sun-spot activity, ionization density increases, and therefore there is a progressive decrease in the value of skip distance for any given frequency. The skip distance for a given frequency rises to its highest value when the sun-spot activity is a minimum and falls to its lowest value when the sun-spot activity is a maximum. Conversely, the skip frequency for any given distance has its highest value when sun-spot activity is a maximum, and its lowest value when sun-spot activity is a minimum.

Percentage Additions to Frequency to Allow for Solar-Cycle Effects

Year	Latitude					
	10°	20°	30°	40°	50°	60°
1930	3	6	9	12	15	18
1931	2	4	6	8	10	12
1932	1	2	3	4	5	6
1933	0	0	0	0	0	0
1934	0	0	0	0	0	0
1935	1	2	3	4	5	6
1936	2	4	6	8	10	12
1937	3	6	9	12	15	18
1938	4	8	12	16	20	24
1939	5	10	15	20	25	30
1940	4	8	12	16	20	25
1941	3	6	9	12	15	18
1942	2	4	6	8	10	12
1943	1	2	3	4	5	6
1944	0	0	0	0	0	0
1945	0	0	0	0	0	0
1946	1	2	3	4	5	6
1947	2	4	6	8	10	12
1948	3	6	9	12	15	18
1949	4	8	12	16	20	24

The skip distance graphs have been plotted for a time of minimum sun-spot activity, which last occurred about 1933/1934, but by using the above table it is possible to make approximate corrections for the progressive ionospheric changes which have occurred since that time. Notice that the percentage additions to frequency which must be made to allow for the solar cycle vary with latitude; the higher the latitude, the greater is the change of ionospheric conditions resulting from changing solar conditions. Two examples illustrating the use of the correction table are given below.

Example 6

In 1933/34 signalling along a path 1,000 miles long and lying between latitudes 25° and 35° N. was carried out between 1000 and 1300 hours every day in February by using a frequency of 13 Mc/s. What frequency should be employed for similar work in 1937?

On reference to the table, it will be seen that for latitudes 30° a 9 per cent. addition to the 1933/34 frequency should be made to allow for the difference between ionospheric conditions in 1937 and 1933/34, i.e., the 1937 frequency should be 13 Mc/s + 1·17 Mc/s = 14·17 Mc/s.

Example 7

Correct the answers to (*a*), (*b*) and (*c*) of Example 2, the corrected answers to give skip distances for the year 1937 for the frequencies of 12, 9 and 6 Mc/s.

(*a*) A frequency of 12 Mc/s in 1937 is equivalent to a frequency of f Mc/s in 1933/34, where

$$f + \cdot 09f = 12$$

$$\text{i.e., } f = \frac{12}{1 \cdot 09} = 11 \text{ Mc/s.}$$

From graph No. 7 the skip distance at noon for this frequency is approximately 400 miles.

Applying the reasoning above to cases (*b*) and (*c*) of Example 2, the 1933/34 frequencies corresponding to 1937 frequencies of 9 and 6 Mc/s are 8·25 Mc/s and 5·5 Mc/s, and these give skip distances at noon of just under 200 miles and zero respectively. Hence, for 1937, the skip distances at noon for the three frequencies are :—

(*a*) 12 Mc/s—400 miles (approx.).

(*b*) 9 Mc/s—200 miles (approx.).

(*c*) 6 Mc/s—zero.

This Page Intentionally Blank

SKIP-DISTANCE/FREQUENCY GRAPHS

(i) The graphs are numbered consecutively from 1 to 26, each diagram, figs. 21 to 33 inclusive, containing two graphs.

(ii) The appropriate graph number for any latitude and month is given in the following table :—

Hemisphere. Northern	Hemisphere. Southern	Lat. 15° N. to 15° S.	Lat. 15° to 25° N. or S.	Lat. 25° to 35° N. or S.	Lat. 35° to 45° N. or S.	Lat. 45° to 55° N. or S.	Lat. 55° to 65° N. or S.
Jan., Nov., Dec.	May, June, July ..	1	2	7	12	17	22
Feb., Oct. ..	April, Aug. ..	1	3	8	13	18	23
March, Sept. ..	March, Sept. ..	1	4	9	14	19	24
April, Aug. ..	Feb., Oct. ..	1	5	10	15	20	25
May, June, July	Jan., Nov., Dec. ..	1	6	11	16	21	26

(iii) The graphs are calculated for the minimum period of a sun-spot cycle ; corrections must be applied for other periods.

(iv) The range "*d*" is in miles.

This Page Intentionally Blank

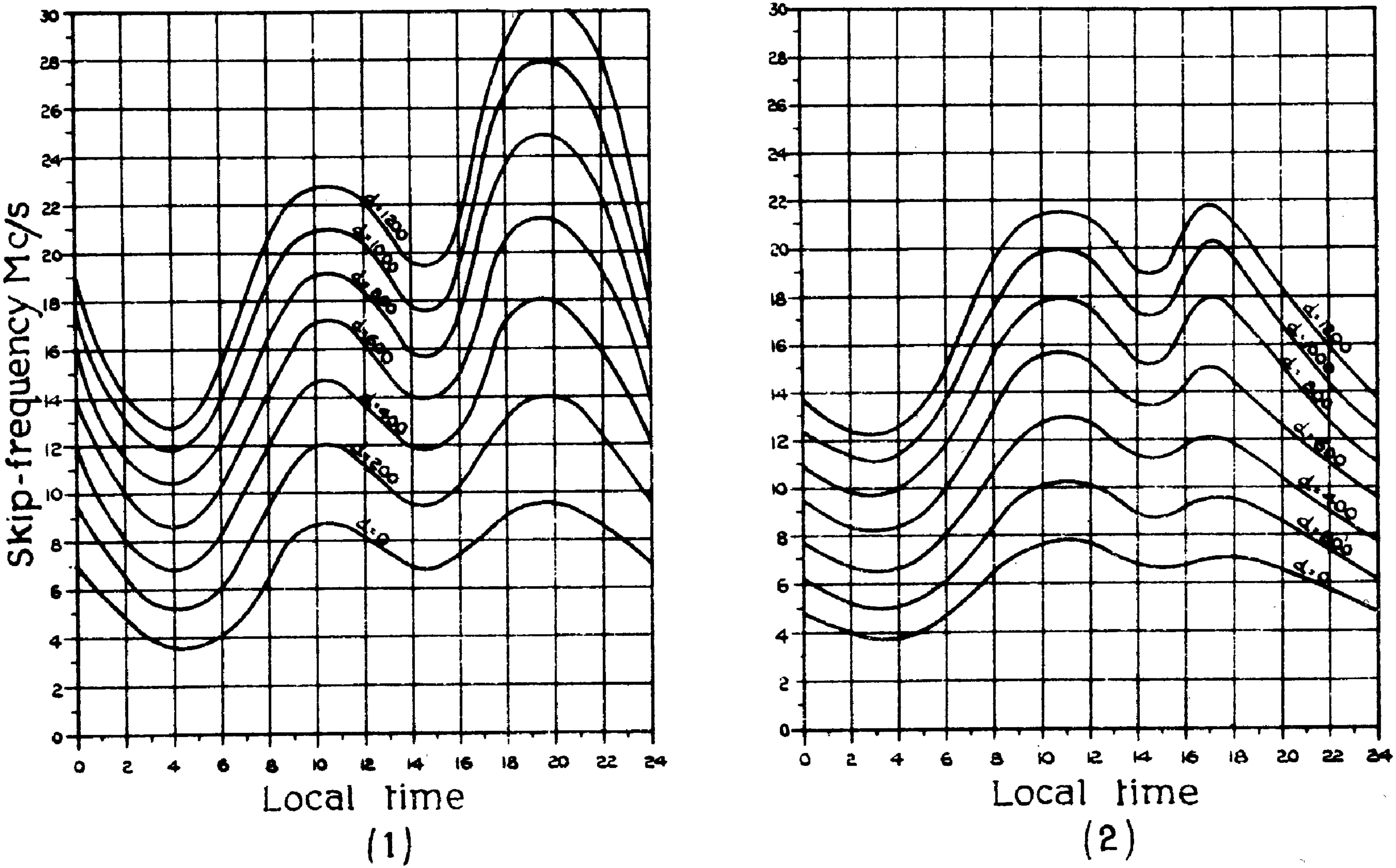
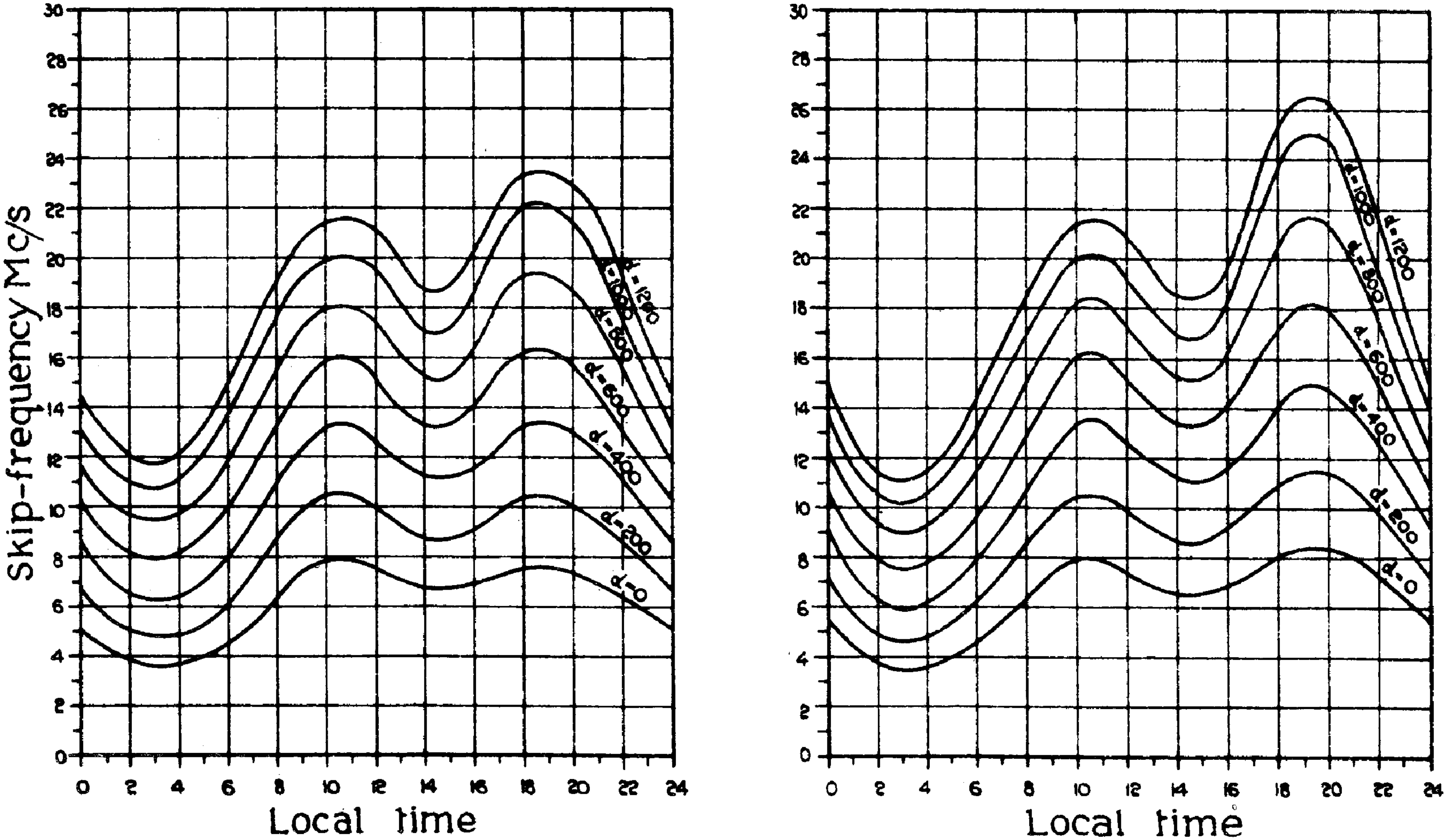
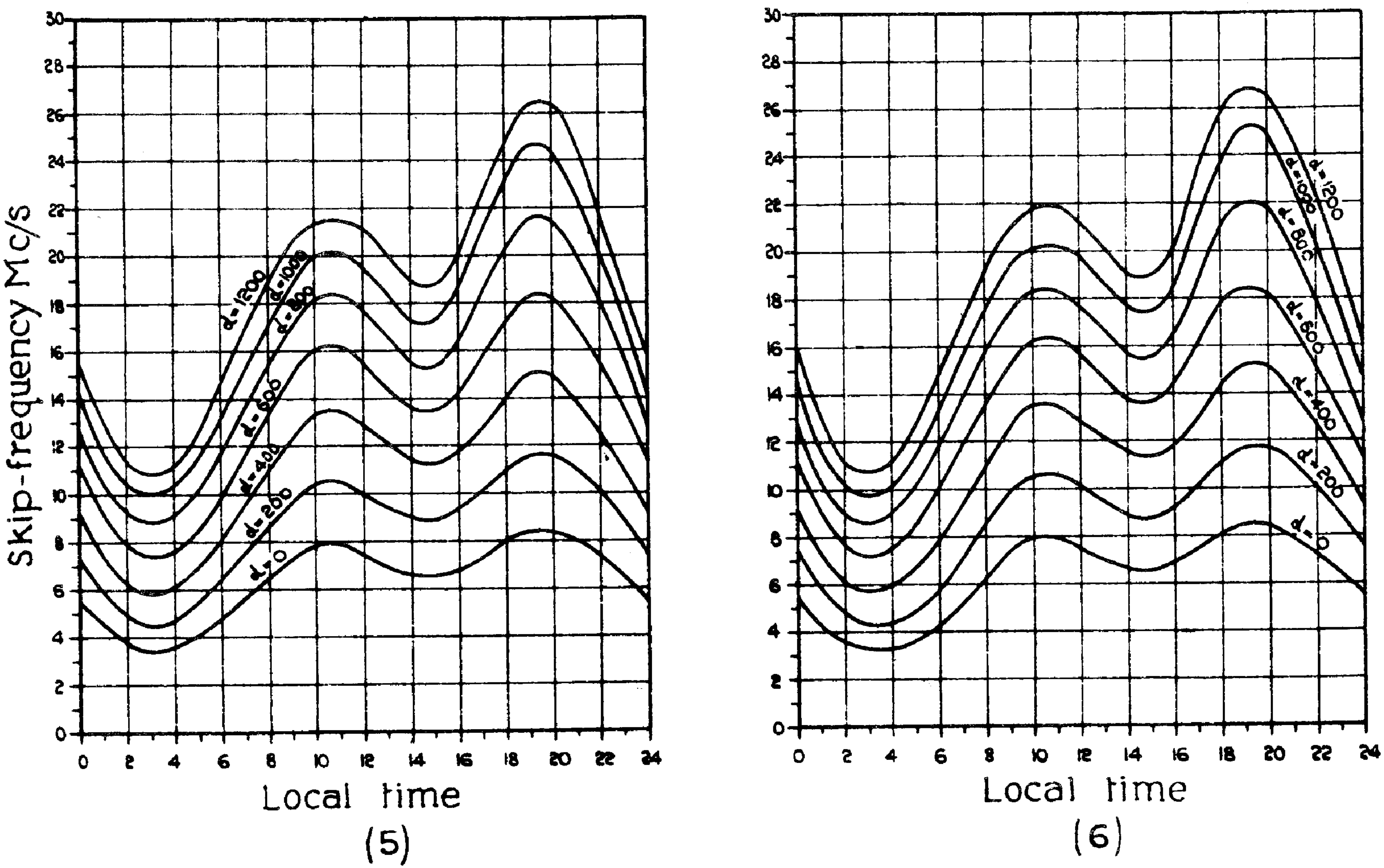


FIG. 21
CHAP. XIV



(3) (4)

FIG. 22
CHAP. XIV



(5) (6)

FIG.23
CHAP. XIV

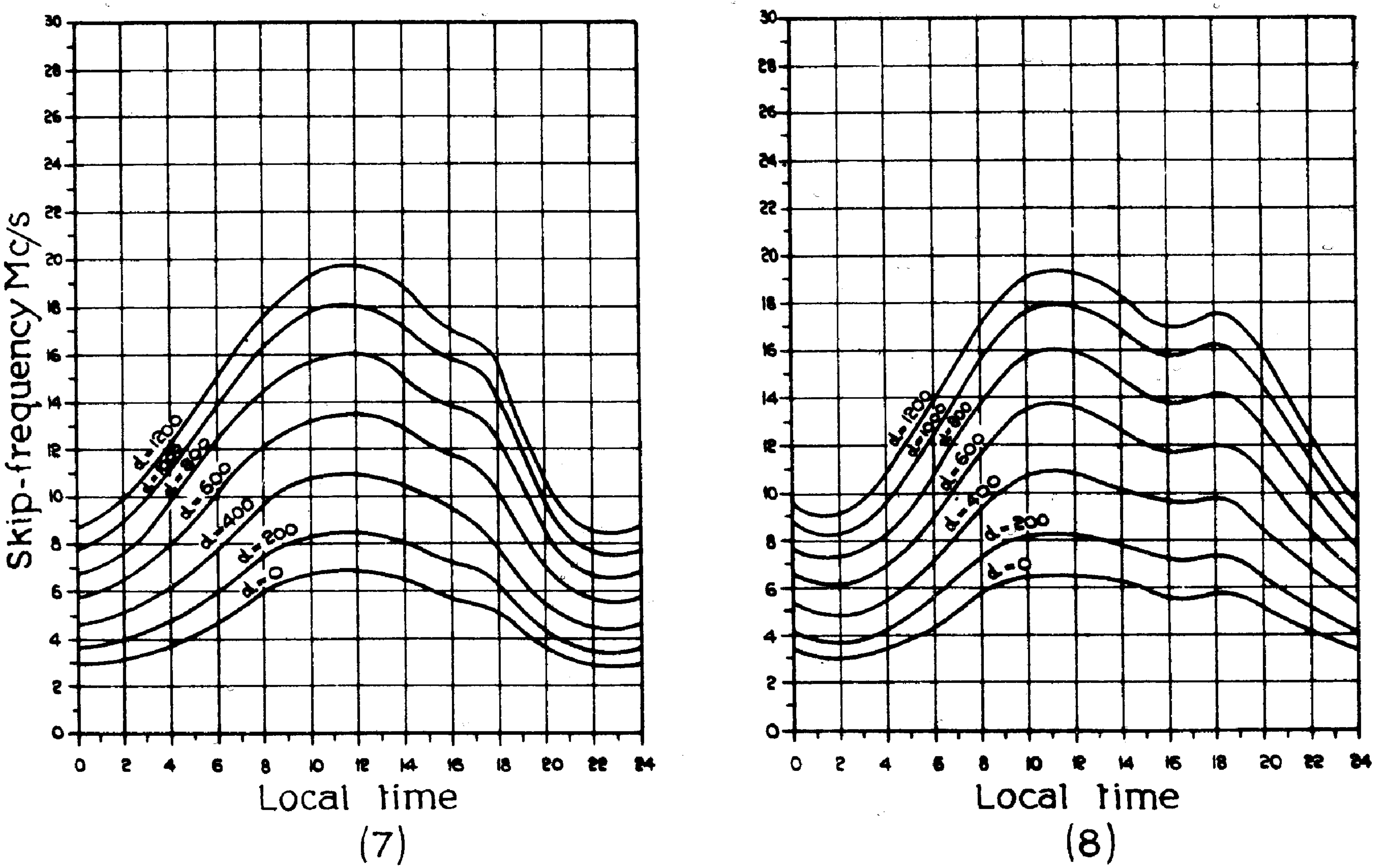


FIG. 24
CHAP. XIV

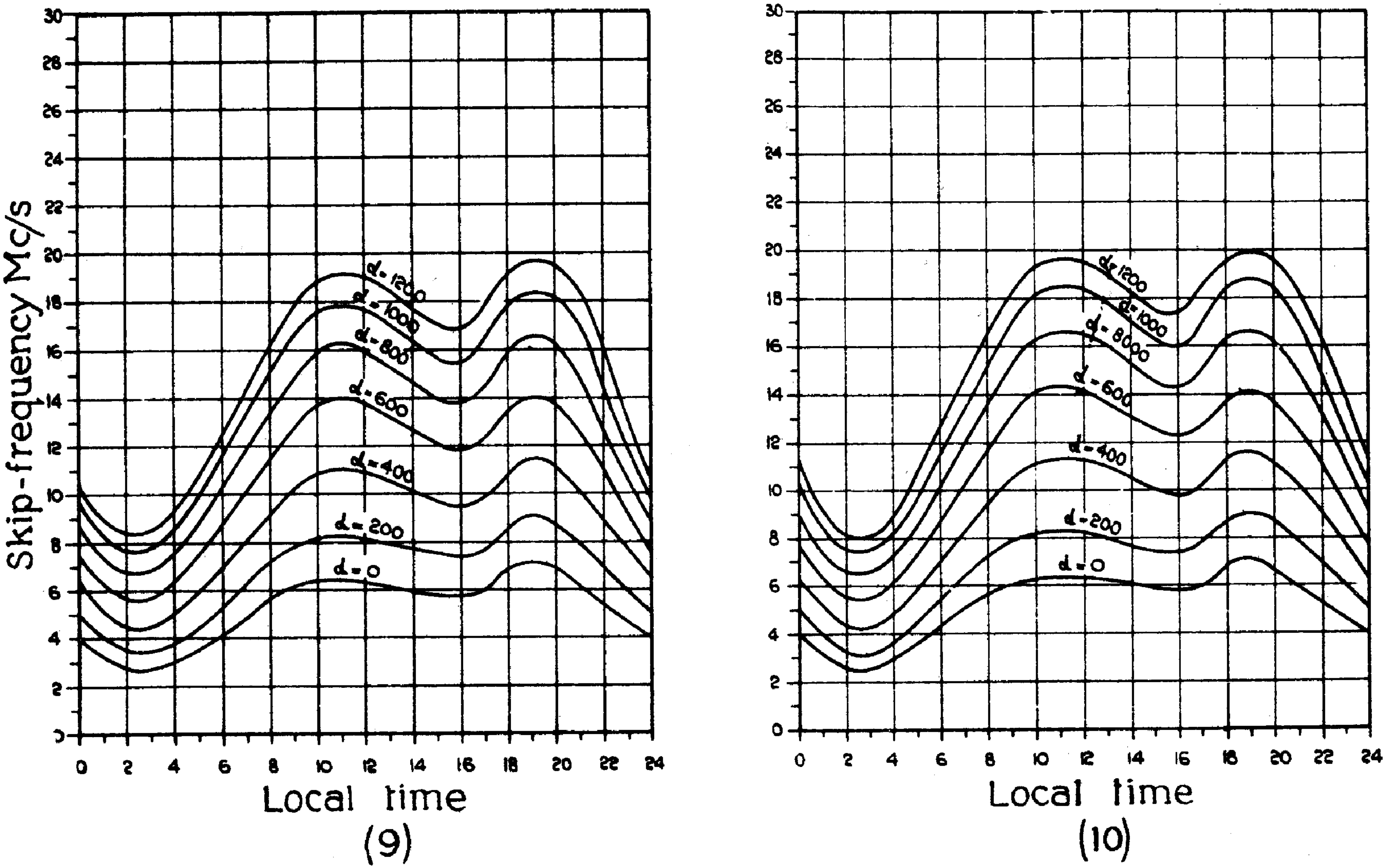


FIG. 25
CHAP. XIV

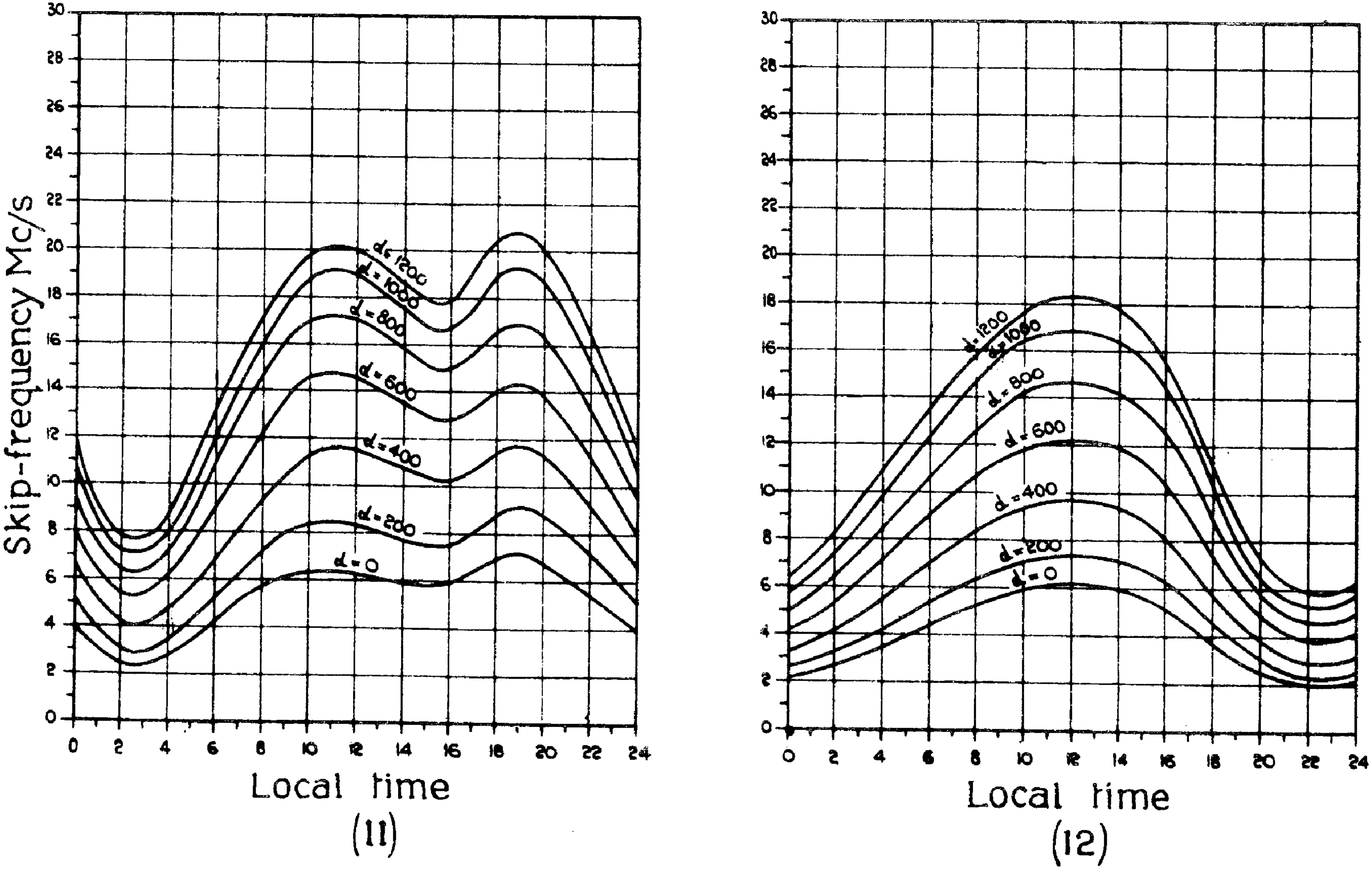


FIG. 26
CHAP. XIV

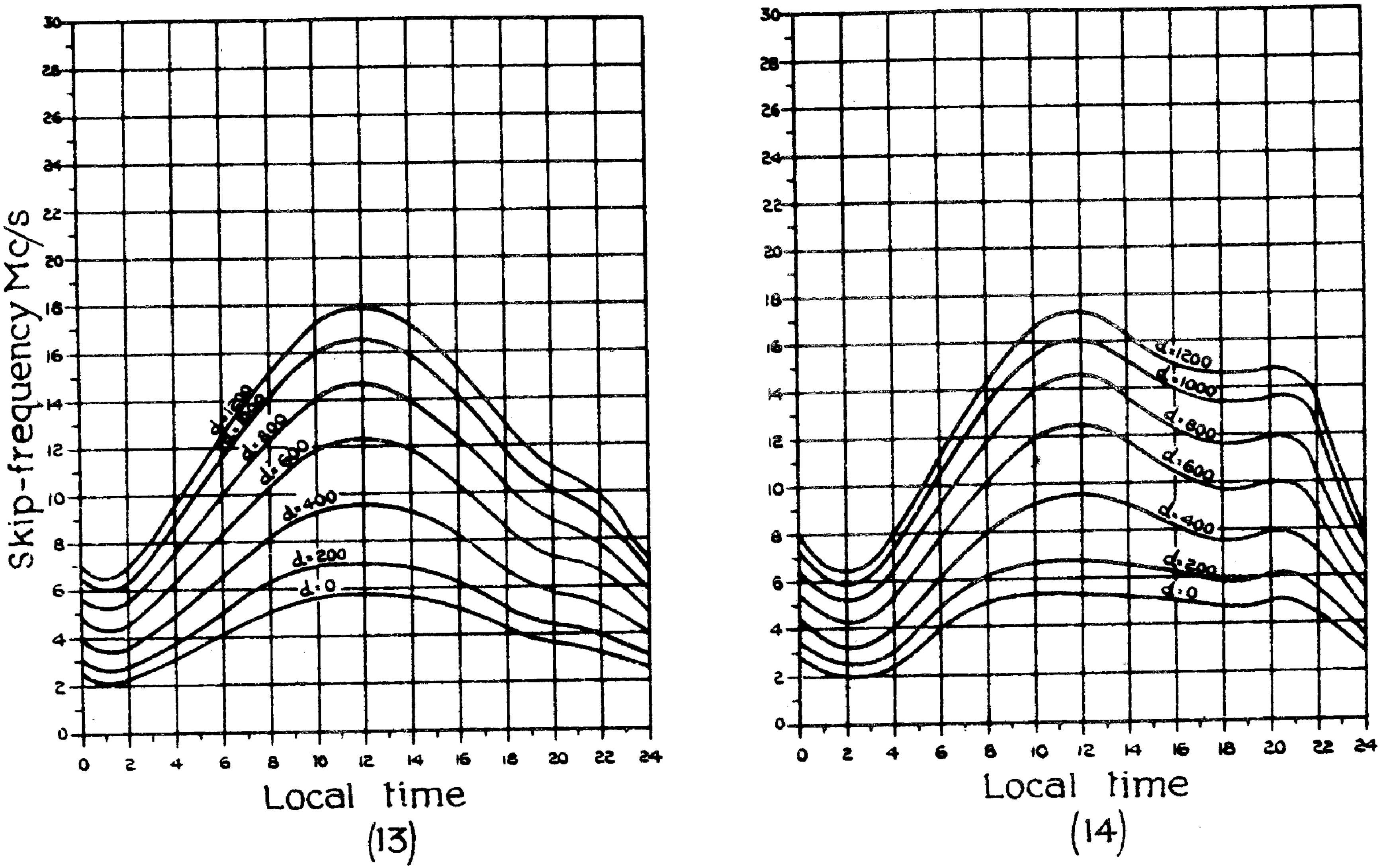
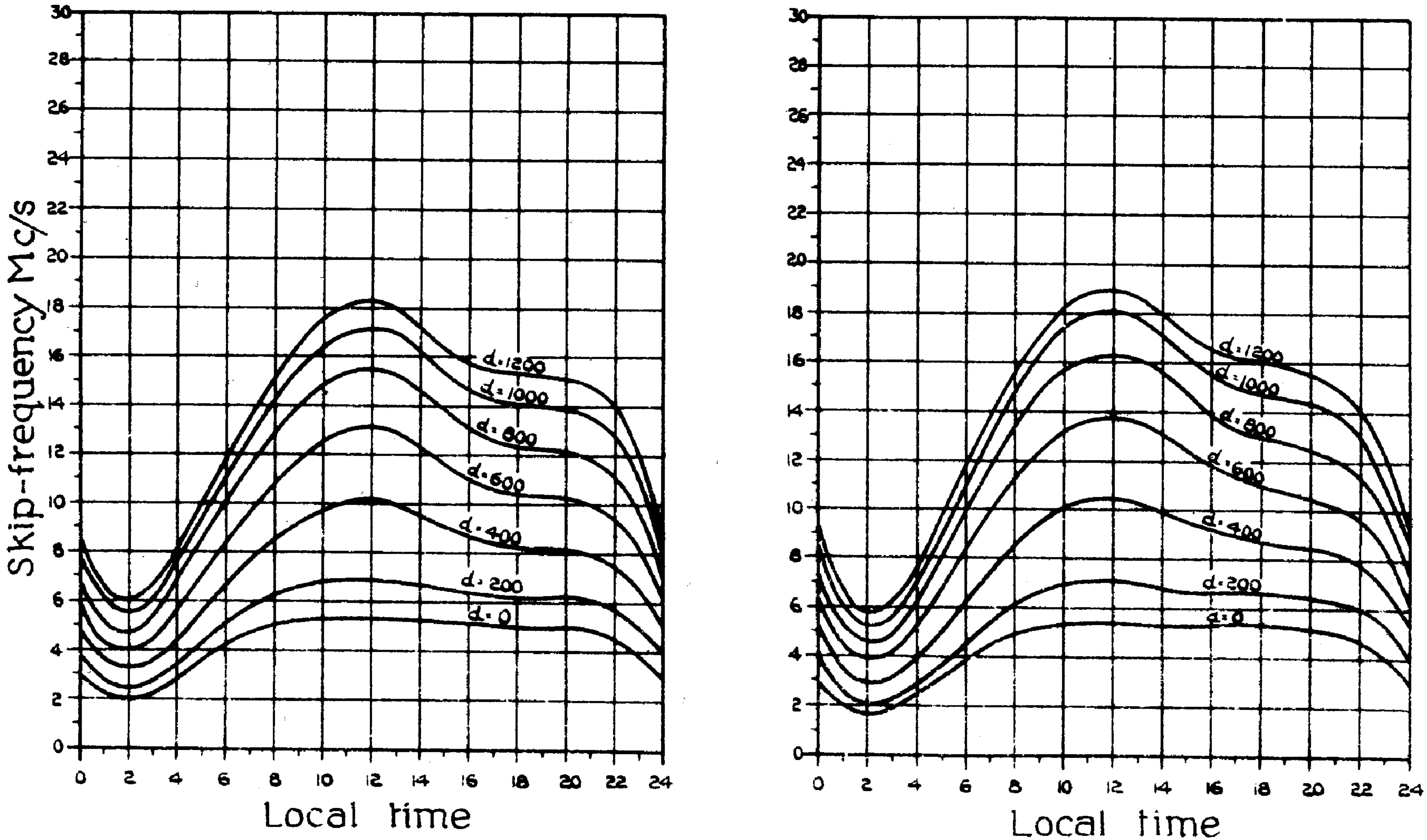


FIG. 27
CHAP XIV

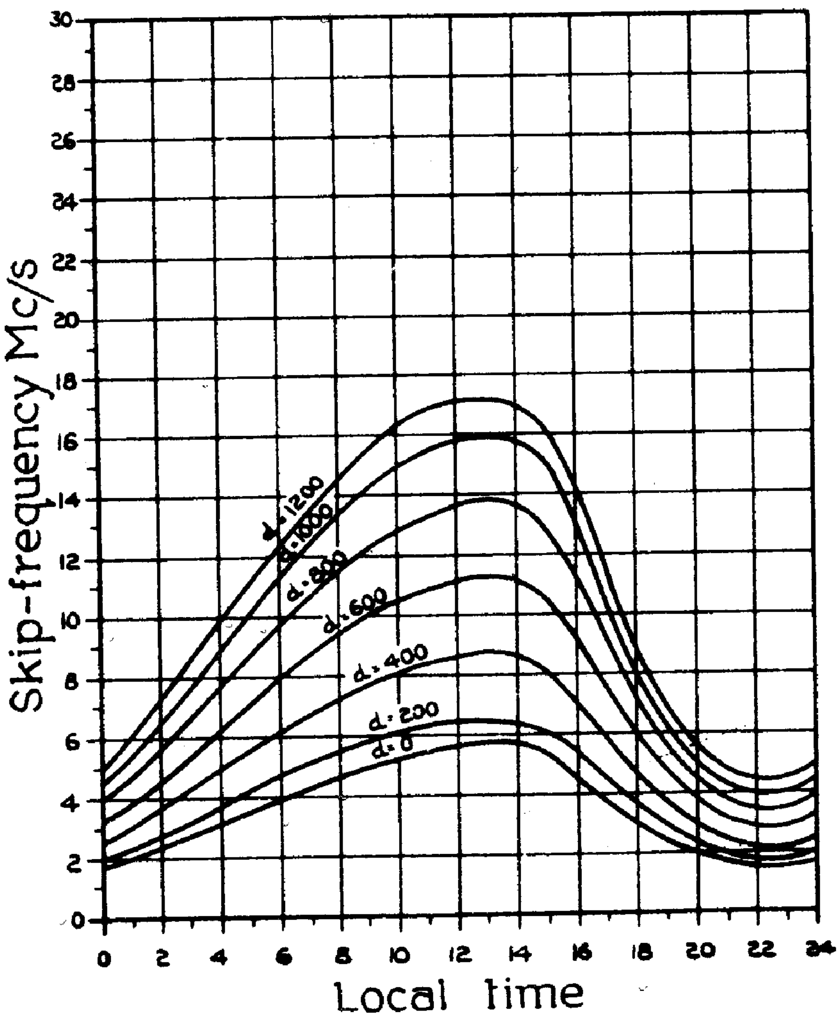


(15)

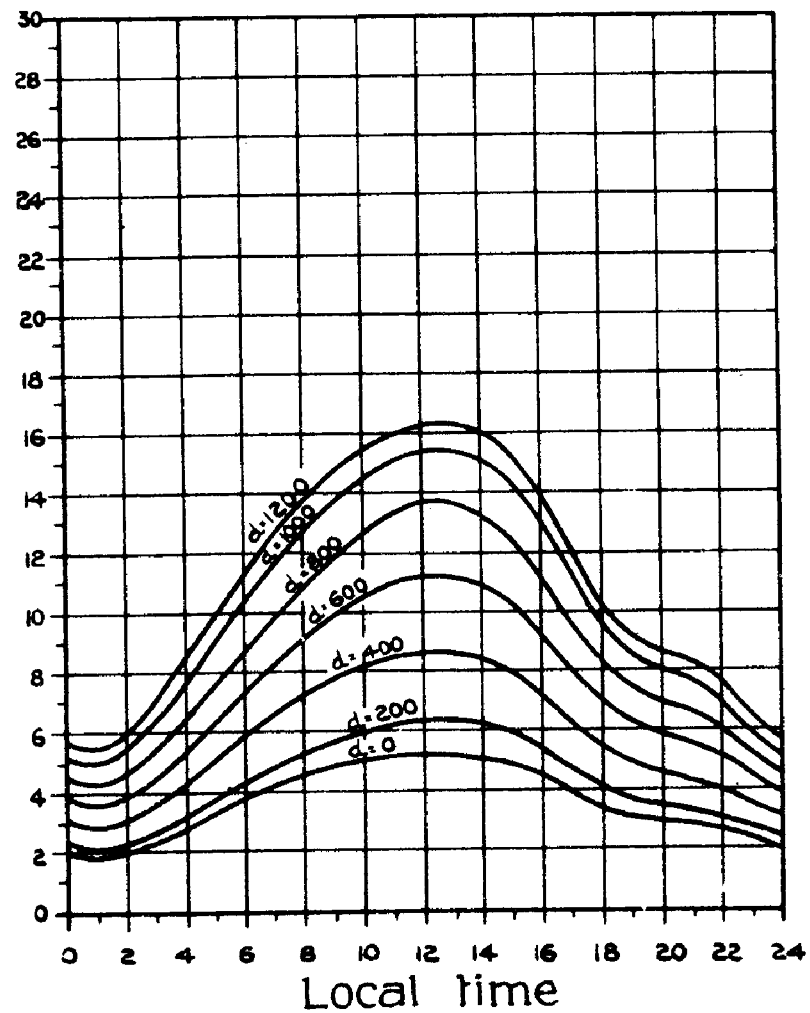
(16)

FIG. 28
CHAP. XIV

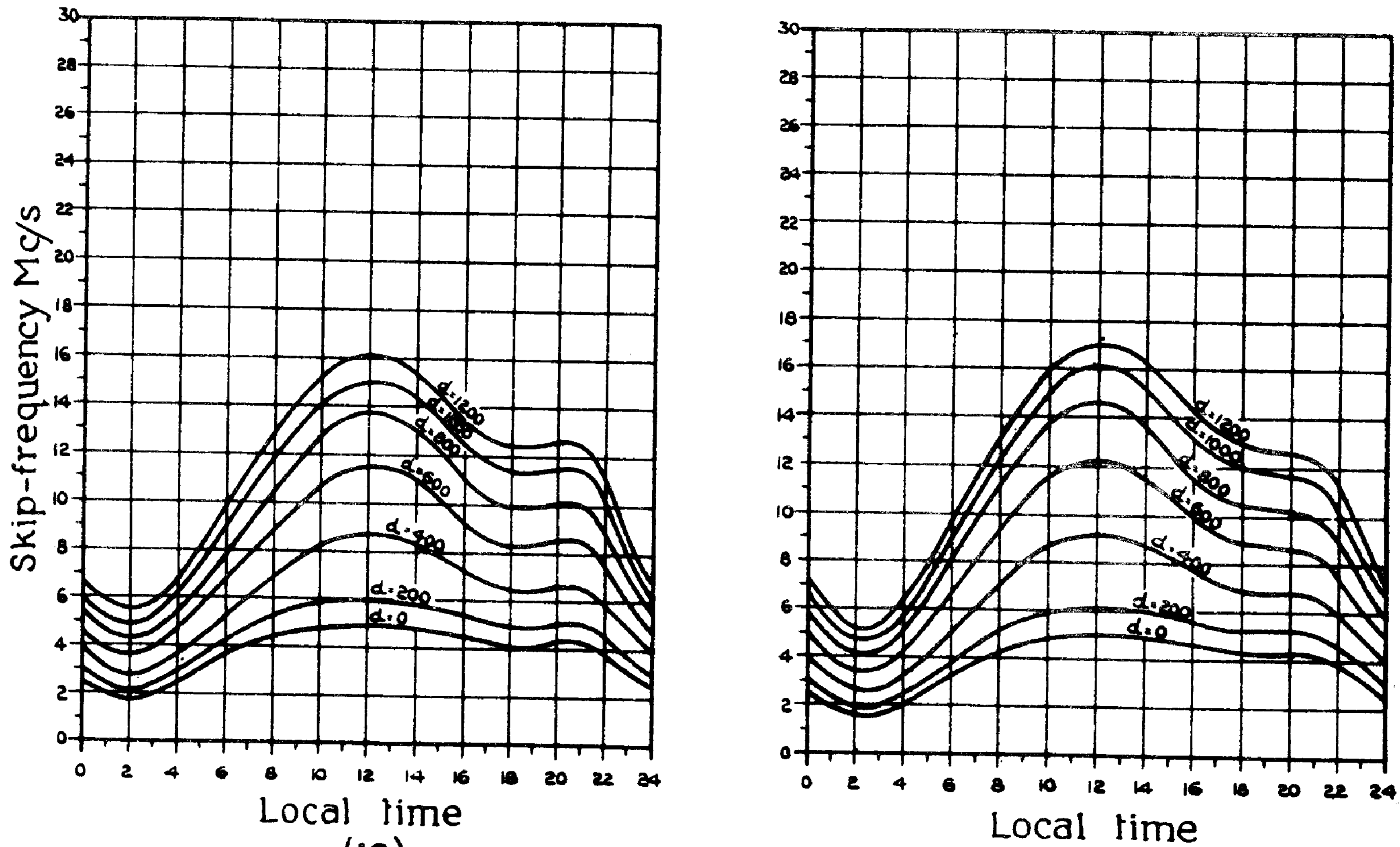
FIG. 29
CHAP. XIV



(17)

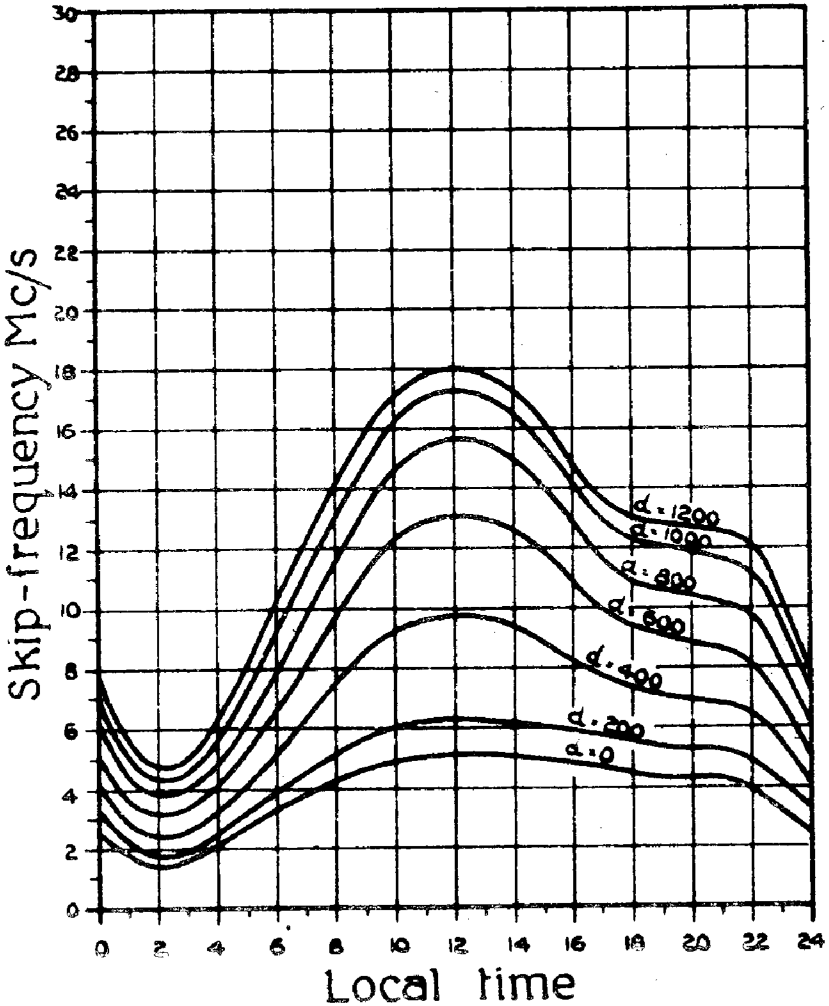


(18)

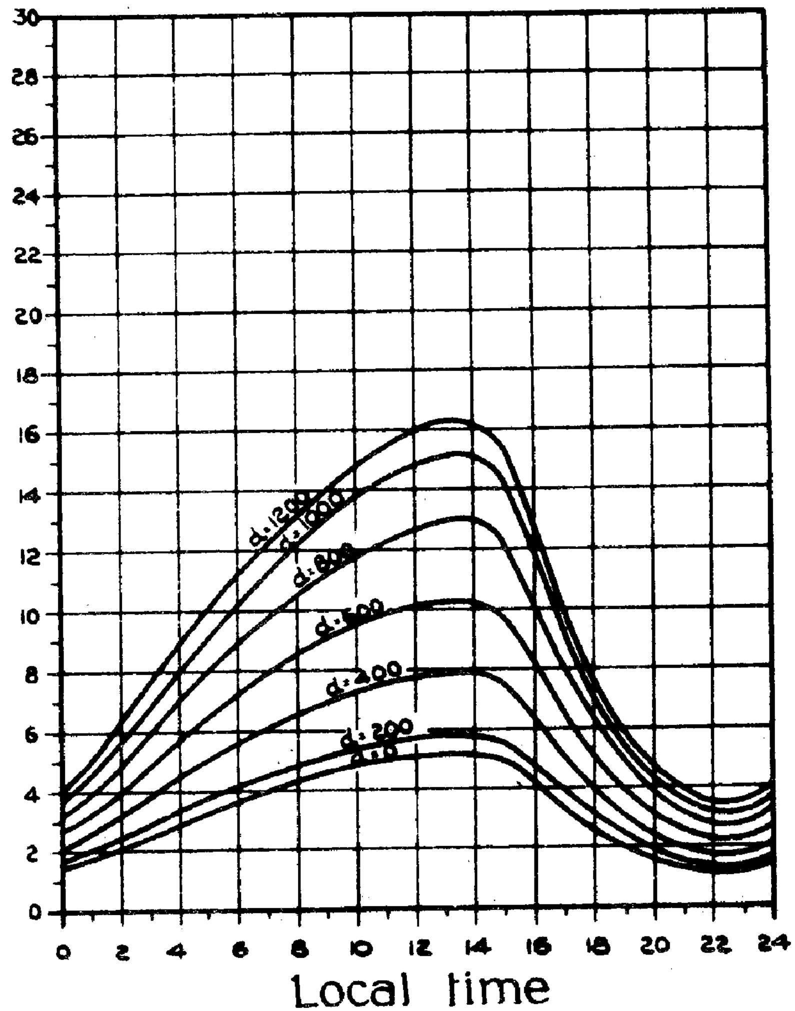


(19) (20)

FIG. 30
CHAP. XIV

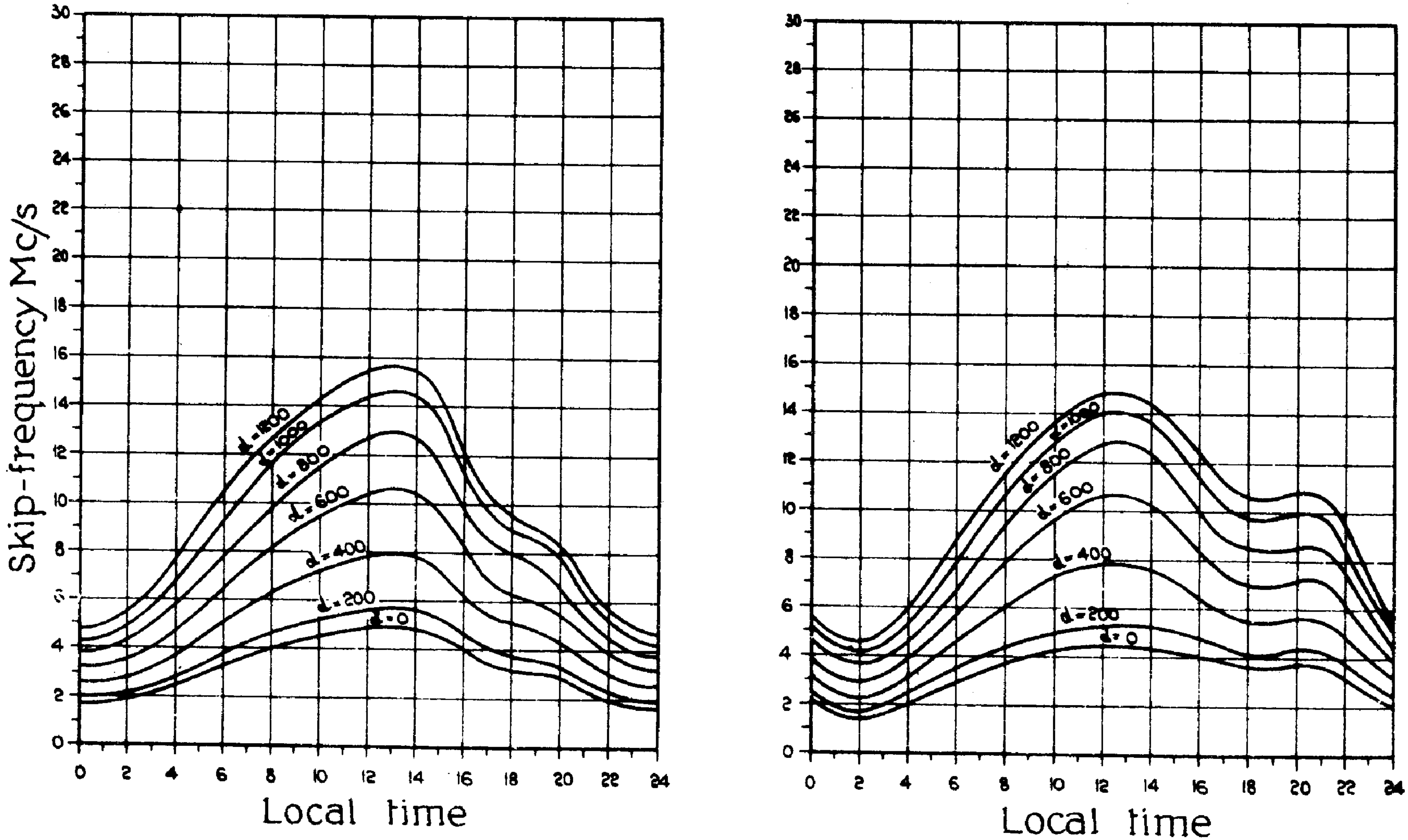


(21)



(22)

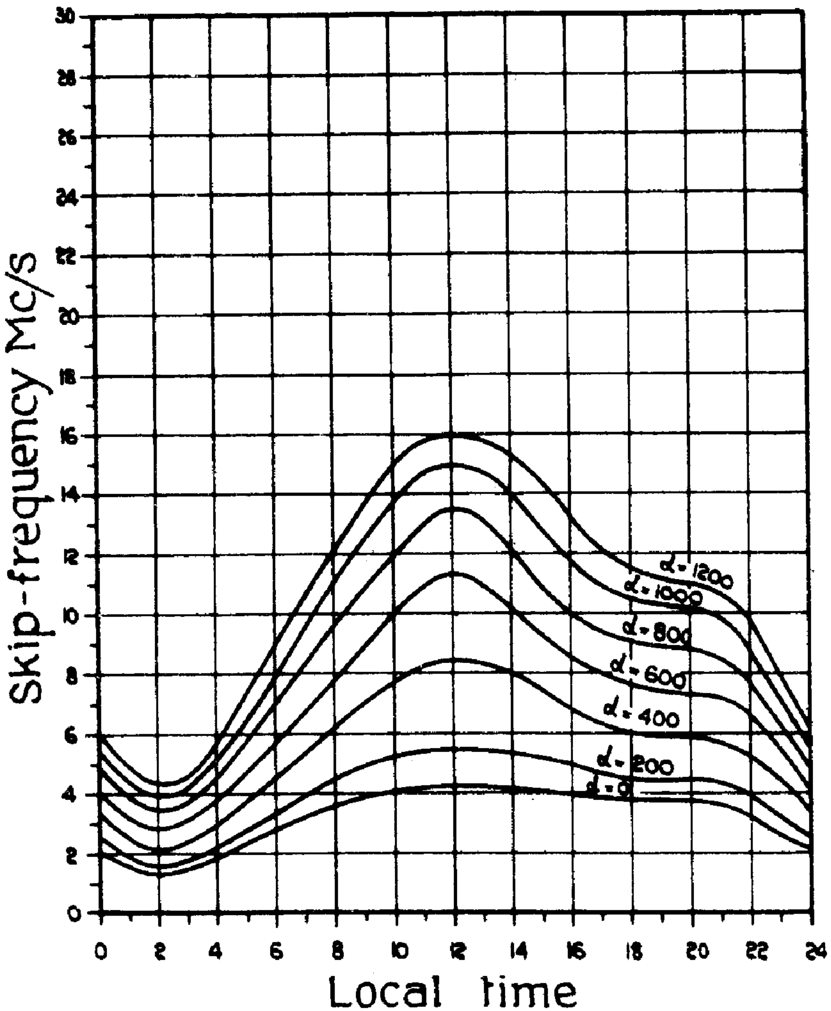
FIG. 31
CHAP. XIV



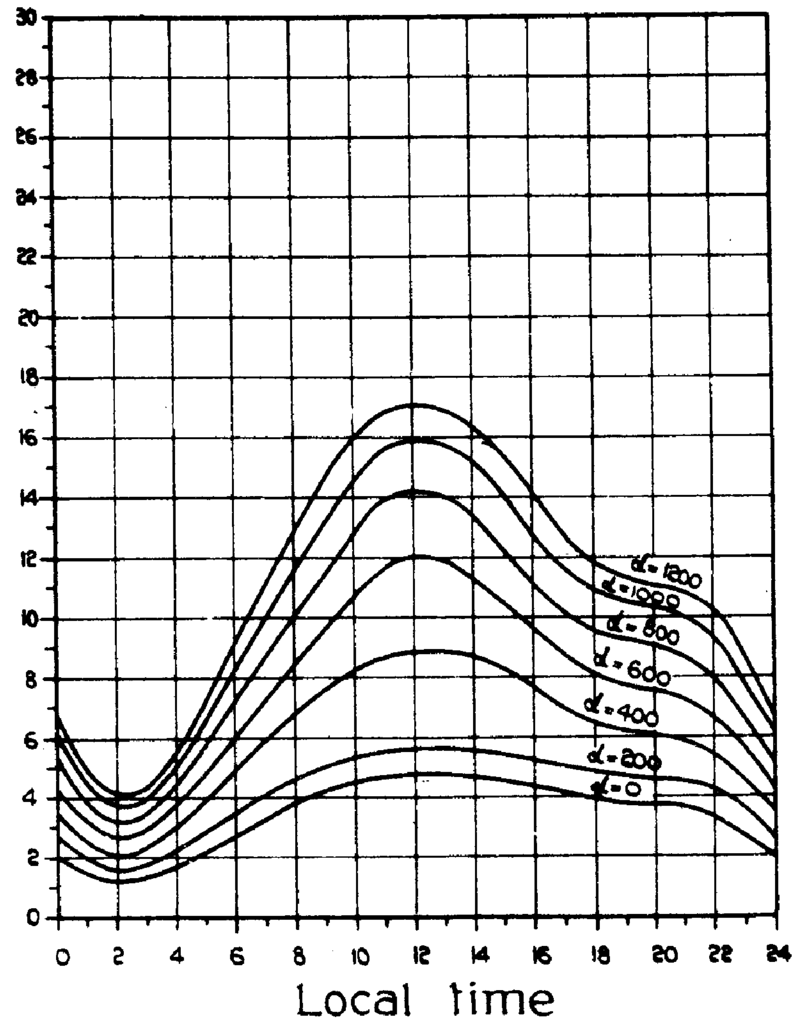
(23)

(24)

FIG. 32
CHAP XIV



(25)



(26)

FIG. 33
CHAP. XIV

This Page Intentionally Blank