Electromagnetic diffraction in optical systems
II. Structure of the image field in an aplanatic system

By B. Richards†

The Computing Machine Laboratory, University of Manchester

and E. Wolf‡

Department of Theoretical Physics, University of Manchester

(*Communicated by D. Gabor, F.R.S.—Received* 19 *February* 1959)

An investigation is made of the structure of the electromagnetic field near the focus of an aplanatic system which images a point source. First the case of a linearly polarized incident field is examined and expressions are derived for the electric and magnetic vectors in the image space. Some general consequences of the formulae are then discussed. In particular the symmetry properties of the field with respect to the focal plane are noted and the state of polarization of the image region is investigated. The distribution of the time-averaged electric and magnetic energy densities and of the energy flow (Poynting vector) in the focal plane is studied in detail, and the results are illustrated by diagrams and in a tabulated form based on data obtained by extensive calculations on an electronic computor. The case of an unpolarized field is also investigated.

The solution is not restricted to systems of low aperture, and the computational results cover, in fact, selected values of the angular semi-aperture α on the image side, in the whole range $0 \leqslant \alpha \leqslant 90°$. The limiting case $\alpha \to 0$ is examined in detail and it is shown that the field is then completely characterized by a single, generally complex, scalar function, which turns out to be identical with that of the classical scalar theory of Airy, Lommel and Struve.

The results have an immediate bearing on the resolving power of image forming systems; they also help our understanding of the significance of the scalar diffraction theory, which is customarily employed, without a proper justification, in the analysis of images in low-aperture systems.

1. Introduction

A knowledge of the structure of an electromagnetic field in the region of focus is of considerable theoretical as well as practical interest. As already indicated in part I of this investigation (Wolf 1959), information about the structure of this complex region is particularly desirable in connexion with the design and the analysis of performance of optical systems of wide angular aperture, both in the field of visible and microwave optics. The construction of large paraboloids for use in radio astronomy has made this problem also of topical interest.

Apart from the usual scalar treatments, the limitations of which were mentioned in part I, the only earlier treatments of this problem appear to be those of Ignatowsky (1919, 1920) and Hopkins (1943, 1945).§ The investigations of Ignatowsky have gone a considerable way towards the solution of the problem. He obtained

† Now at the Computer Section, I.C.I., Ltd, Wilton Works, Middlesbrough, Yorkshire.

‡ Now at the Institute of Optics, University of Rochester, Rochester, N.Y., U.S.A.

§ Since the present investigation was carried out, two related papers have appeared. Burtin (1956) determined the distribution of the electric energy density in the focal plane of a system with angular semi-aperture $\alpha = 45°$ on the image side, when the wave entering the system is linearly polarized. Focke (1957) considered an unpolarized wave and studied the energy density and the energy flow at points in the focal plane in systems of selected angular apertures.

formulae for the electric and magnetic field vectors in the image region of an aplanatic system† and also of a paraboloid of any angular aperture. Unfortunately his deductions from these formulae were chiefly confined to the study of the energy flow across the central bright nucleus of the image; the electric energy density (which is presumably what many detectors, e.g. a photographic plate, record) is not discussed. The researches of Hopkins were mainly concerned with the modification which the Airy pattern undergoes as the angular aperture of the system is increased. While his analysis is not based on the full Maxwell's equations, it does take into account the vectorial nature of the problem and his results are of practical interest in connexion with optical systems the angular semi-apertures of which do not exceed about 40°.

In the present paper a thorough investigation is made of the structure of the electromagnetic field near the focus of an aplanatic system which images a point source at infinity. Formulae are derived for the electric and magnetic vectors in the image space and a number of general properties of the electromagnetic field are deduced, both for polarized and unpolarized incident waves. The formulae are evaluated for a large number of the basic parameters of the problem and the results are presented in the form of diagrams and in tabulated form. Systems of selected angular semi-aperture α on the image side are considered, up to the limiting case $\alpha \to 90°$. The other limiting case, $\alpha \to 0$, is examined in detail and it is found that, in this case, the vector solution is completely characterized by one (generally complex) scalar function, which is found to be identical with that of the classical analyses of Airy (1835), Lommel (1885) and Struve (1886).

The results have an immediate bearing on the resolving power of systems with high angular aperture; they also help our understanding of the significance of the usual scalar methods of optical diffraction theory.

Finally, it may be mentioned that our basic formulae (equations (2·30) and (2·31)) are in agreement with those of Ignatowsky (1919); our deductions, however, go considerably further.

Some preliminary results of this investigation were reported in two previous notes (Richards 1956*a*; Richards & Wolf 1956).

2. Expressions for the field vectors in the image space of an aplanatic system

Consider an optical system of revolution, which images a point source. The imaging will be assumed to be aplanatic, i.e. axially stigmatic and obeying the sine condition. The source will be assumed to be at infinity in the direction of the axis, and to begin with it will be assumed that it gives rise to a linearly polarized monochromatic wave in the entrance pupil of the system. The case of an unpolarized wave will be considered later (§ 5) by averaging over all possible states of polarization. It is assumed that the linear dimensions of the exit pupil are large compared with the wavelength.

† By aplanatic system we mean one which, for a specified axial position of the object point is stigmatic and obeys the Abbe sine condition.

Expressions for the field in the image region of the system may be derived by an application of formulae (2·18) and (2·19) of part I. These formulae express the time-independent parts, **e** and **h** of the electric and magnetic vectors.

$$\mathbf{E}(x,y,z,t) = \mathscr{R}\{\mathbf{e}(x,y,z)\,\mathrm{e}^{-\mathrm{i}\omega t}\}, \quad \mathbf{H}(x,y,z,t) = \mathscr{R}\{\mathbf{h}(x,y,z)\,\mathrm{e}^{-\mathrm{i}\omega t}\}, \tag{2·1}$$

at any point $P(x,y,z)$ in the image space, which is not too close to the exit pupil, in the form

$$\mathbf{e}(x,y,z) = -\frac{\mathrm{i}k}{2\pi}\iint_{\Omega}\frac{\mathbf{a}(s_x,s_y)}{s_z}\,\mathrm{e}^{\mathrm{i}k[\Phi(s_x,s_y)+s_x x+s_y y+s_z z]}\,\mathrm{d}s_x\,\mathrm{d}s_y, \tag{2·2}$$

$$\mathbf{h}(x,y,z) = -\frac{\mathrm{i}k}{2\pi}\iint_{\Omega}\frac{\mathbf{b}(s_x,s_y)}{s_z}\,\mathrm{e}^{\mathrm{i}k[\Phi(s_x,s_y)+s_x x+s_y y+s_z z]}\,\mathrm{d}s_x\,\mathrm{d}s_y, \tag{2·3}$$

where $\mathscr{R}$ denotes the real part. Here **a** and **b** are the 'strength factors' of the unperturbed electromagnetic field $\mathbf{E}^{(i)} = \mathscr{R}\{\mathbf{e}^{(i)}\,\mathrm{e}^{-\mathrm{i}\omega t}\}$, $\mathbf{H}^{(i)} = \mathscr{R}\{\mathbf{h}^{(i)}\,\mathrm{e}^{-\mathrm{i}\omega t}\}$ which is incident on the exit pupil, i.e. **a** and **b** are defined by the relations

$$\mathbf{e}^{(i)} = \frac{\mathbf{a}}{\sqrt{(R_1R_2)}}\,\mathrm{e}^{\mathrm{i}k\mathscr{S}}, \quad \mathbf{h}^{(i)} = \frac{\mathbf{b}}{\sqrt{(R_1R_2)}}\,\mathrm{e}^{\mathrm{i}k\mathscr{S}}, \tag{2·4}$$

where R_1 and R_2 are the principal radii of curvature of the associated wave-front and $\mathscr{S}$ is the eikonal function. Further

$$k = \frac{\omega}{c} = \frac{2\pi}{\lambda} \tag{2·5}$$

is the vacuum wave number, and λ the vacuum wavelength (denoted in part I by k_0 and λ_0 respectively), it being assumed here that the refractive index of the image space is unity. Further, Φ is the aberration function of the system, s_x, s_y, s_z are the components of the unit vector **s** (with its positive direction in the direction of propagation of the light) along a typical ray in the image space, and Ω is the solid angle formed by all the geometrical rays which pass through the exit pupil of the system. According to (2·20) of part I, the strength factors are related by the formula (assuming that in the image space $\epsilon \simeq \mu \simeq 1$)

$$\mathbf{b} = \mathbf{s} \wedge \mathbf{a}. \tag{2·6}$$

In the problem under consideration, the imaging is aplanatic so that the wave-fronts in the image space are spherical, with a common centre at the Gaussian image point. Hence, for all vectors **s** in the solid angle Ω,

$$\Phi(s_x,s_y) = 0. \tag{2·7}$$

The strength factors **a** and **b** may be determined, with an accuracy sufficient for our purposes, by tracing rays through the system up to the region of the exit pupil, and by making use of the laws relating to the variation of the field vectors along each ray. Let AQ_0 be a typical ray entering the system at a height h from the axis and let θ be the angle which the corresponding ray QO in the image space makes with the axis (figure 1). Since the field in the object space is linearly polarized, the field obtained on refraction at the first surface σ_1 is also linearly polarized (the direction

of polarization will, of course, be in general different for different rays); and if the angle of the incidence at σ_1 is small, the angle which the direction of vibrations of the electric (and also of the magnetic) vector makes with the meridional plane of the ray (the plane containing the ray and the axis of the system) will be effectively unchanged by refraction.† Moreover (see Born & Wolf 1959, §3·1·3), in a homogeneous medium the direction of vibration remains unchanged along each ray. By successive repetition of these arguments from surface to surface it follows that the field in the region of the exit pupil is also linearly polarized and that, provided the angle of incidence at each surface of the system is small, the angles which the **e** and **h** vectors in the exit pupil make with the meridional plane of the ray are equal to the corresponding angle in the object space. Further, the magnitude of the field vectors in the region of the exit pupil may be determined from the geometrical law of conservation of energy, taking into account the fact that the system obeys the sine condition.

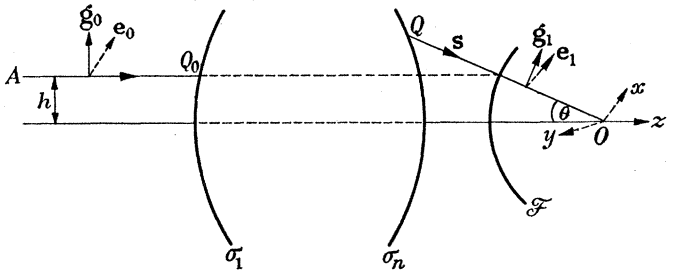

FIGURE 1. The meridional plane of a ray. The axis Ox is in the direction of the electric vector $\mathbf{e}_0$ in the object space.

Let $\mathscr{F}$ be the 'focal sphere', i.e. the sphere with centre at O and with radius f equal to the focal length of the system. Then, according to the sine condition

$$h = f \sin\theta. \tag{2·8}$$

This relation implies that the emergent ray meets the focal sphere at the same height at which the corresponding ray in the object space entered the system (see figure 1).

Consider now all the rays which enter the annulus bounded by circles of radius h and $h+\delta h$. Let δS_0 be the area of the annulus and δS_1 the corresponding area of the focal sphere, and let $\mathbf{e}_0$ and $\mathbf{e}_1$ be the electric amplitude vectors on δS_0 and δS_1 respectively. Then we may write

$$\mathbf{e}_0 = l_0 \hat{\mathbf{e}}_0 \mathrm{e}^{\mathrm{i}k\mathscr{S}_0}, \quad \mathbf{e}_1 = l_1 \hat{\mathbf{e}}_1 \mathrm{e}^{\mathrm{i}k\mathscr{S}_1}, \tag{2·9}$$

where $\hat{\mathbf{e}}_0$ and $\hat{\mathbf{e}}_1$ are unit vectors, in the direction of $\mathbf{e}_0$ and $\mathbf{e}_1$ respectively, l_0 and l_1 are amplitude factors and $\mathscr{S}_0$ and $\mathscr{S}_1$ are the corresponding values of the eikonal. Since the field is linearly polarized, l_0, l_1, $\hat{\mathbf{e}}_0$ and $\hat{\mathbf{e}}_1$ may be taken to be real, provided

† The truth of these statements follows from Fresnel formulae for refraction of a plane wave on a plane interface between two homogeneous media of different refractive indices. Within the accuracy of the present approximation (geometric optics limit, i.e. $k \to \infty$), these formulae also apply to the present case when neither the wave-fronts nor the refracting boundaries are necessarily plane.

that the origin of the phase is suitably chosen. According to the geometrical optics intensity law,

$$l_0^2 \delta S_0 = l_1^2 \delta S_1, \tag{2·10}$$

where we assumed that the refractive index of the object space, like that of the image space, is unity, and that the losses of energy due to reflexion and absorption within the system are negligible. Now from the figure,

$$\delta S_0 = \delta S_1 \cos\theta, \tag{2·11}$$

so that according to (2·10)

$$l_1 = l_0 \cos^{\frac{1}{2}}\theta. \tag{2·12}$$

Now $\mathbf{e}_1$ may be identified with the vector $\mathbf{e}^{(i)}$ of (2·4), with $R_1 = R_2 = f$ and $\mathscr{S} = \mathscr{S}_1$. Hence

$$\mathbf{a} = f l_0 \cos^{\frac{1}{2}}\theta\, \hat{\mathbf{e}}_1. \tag{2·13}$$

To determine $\mathbf{e}_1$ it is convenient to introduce two unit vectors $\mathbf{g}_0$ and $\mathbf{g}_1$ in the meridional plane of the ray, such that $\mathbf{g}_0$ is perpendicular to the ray in the object space and $\mathbf{g}_1$ is perpendicular to the ray in the image space, and both are directed away from the axis (see figure 1). Let Ox, Oy, Oz be Cartesian rectangular co-ordinate axes, with origin at the Gaussian focus, with Ox in the direction of $\mathbf{e}_0$ and with Oz along the axis of the system, pointing away from the plane of the exit pupil into the image space. Finally, let $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ be unit vectors in the direction of the co-ordinate axes.

The electric and the magnetic vectors are orthogonal to the ray (see Born & Wolf 1959, § 3·1). Hence $\mathbf{e}_1$ lies in the plane of $\mathbf{g}_1$ and $\mathbf{g}_1 \wedge \mathbf{s}$, i.e.

$$\hat{\mathbf{e}}_1 = \alpha \mathbf{g}_1 + \beta(\mathbf{g}_1 \wedge \mathbf{s}), \tag{2·14}$$

where α and β are some constants. To determine α and β we make use of a result mentioned earlier, namely that as the light traverses the system, the angle between the electric (and also the magnetic) vector and the meridional plane of the ray remains constant. Moreover, each of these vectors remains on the same side of the meridional plane. These results imply that

$$\left.\begin{aligned} \mathbf{g}_1 . \hat{\mathbf{e}}_1 &= \mathbf{g}_0 . \mathbf{i}, \\ (\mathbf{g}_1 \wedge \mathbf{s}_1) . \hat{\mathbf{e}}_1 &= (\mathbf{g}_0 \wedge \mathbf{k}) . \mathbf{i}. \end{aligned}\right\} \tag{2·15}$$

On substituting from (2·14) into (2·15) we find that

$$\alpha = \mathbf{g}_0 . \mathbf{i}, \quad \beta = (\mathbf{g}_0 \wedge \mathbf{k}) . \mathbf{i} = \mathbf{g}_0 . (\mathbf{k} \wedge \mathbf{i}) = \mathbf{g}_0 . \mathbf{j}, \tag{2·16}$$

and from (2·12), (2·13) and (2·15) it follows that

$$\mathbf{a} = f l_0 \cos^{\frac{1}{2}}\theta [(\mathbf{g}_0 . \mathbf{i})\, \mathbf{g}_1 + (\mathbf{g}_0 . \mathbf{j})\, (\mathbf{g}_1 \wedge \mathbf{s})]. \tag{2·17}$$

It will now be convenient to introduce spherical polar co-ordinates r, θ, ϕ ($r > 0$, $0 \leqslant \theta < \pi$, $0 \leqslant \phi < 2\pi$), with the polar axis $\theta = 0$ in the z-direction, and with the azimuth $\phi = 0$ containing the electric vector in the object space. The components s_x, s_y, s_z of the unit vector $\mathbf{s}$ along a ray in the image space and the co-ordinates (x, y, z) of a point P in the image region may be expressed in the form

$$s_x = \sin\theta \cos\phi, \quad s_y = \sin\theta \sin\phi, \quad s_z = \cos\theta, \tag{2·18a}$$

$$x = r_P \sin\theta_P \cos\phi_P, \quad y = r_P \sin\theta_P \sin\phi_P, \quad z = r_P \cos\theta_P, \tag{2·18b}$$

so that the term in the exponent of the integrals (2·1) and (2·3) becomes

$$s_x x + s_y y + s_z z = r_P \cos\epsilon, \tag{2·19}$$

where

$$\cos\epsilon = \cos\theta\cos\theta_P + \sin\theta\sin\theta_P\cos(\phi - \phi_P). \tag{2·20}$$

Let (θ_0, ϕ_0) and (θ_1, ϕ_1) be the polar angles of $\mathbf{g}_0$ and $\mathbf{g}_1$ respectively. Then evidently (see figure 1)

$$\phi_0 = \phi_1 = \phi - \pi, \quad \theta_0 = \tfrac{1}{2}\pi, \quad \theta_1 = \tfrac{1}{2}\pi - \theta, \tag{2·21}$$

so that

$$\left.\begin{aligned} \mathbf{g}_0 &= -\cos\phi\,\mathbf{i} - \sin\phi\,\mathbf{j}, \\ \mathbf{g}_1 &= -\cos\theta\cos\phi\,\mathbf{i} - \cos\theta\sin\phi\,\mathbf{j} + \sin\theta\,\mathbf{k}. \end{aligned}\right\} \tag{2·22}$$

It follows on substitution from (2·22) and (2·18a) into (2·17) and (2·6) that the Cartesian components of the strength vectors **a** and **b** are

$$\left.\begin{aligned} a_x &= f l_0 \cos^{\frac{1}{2}}\theta[\cos\theta + \sin^2\phi(1-\cos\theta)], \\ a_y &= f l_0 \cos^{\frac{1}{2}}\theta[(\cos\theta - 1)\cos\phi\sin\phi], \\ a_z &= -f l_0 \cos^{\frac{1}{2}}\theta\sin\theta\cos\phi, \end{aligned}\right\} \tag{2·23}$$

$$\left.\begin{aligned} b_x &= f l_0 \cos^{\frac{1}{2}}\theta[(\cos\theta - 1)\cos\phi\sin\phi], \\ b_y &= f l_0 \cos^{\frac{1}{2}}\theta[1 - \sin^2\phi(1-\cos\theta)], \\ b_z &= -f l_0 \cos^{\frac{1}{2}}\theta\sin\theta\sin\phi. \end{aligned}\right\} \tag{2·24}$$

Finally, we also need an expression, in terms of θ and ϕ, for the quantity $\mathrm{d}s_x\mathrm{d}s_y/s_z$, which enters our basic diffraction integrals (2·2) and (2·3). This quantity represents the element $\mathrm{d}\Omega$ of the solid angle and is given by

$$\mathrm{d}s_x\mathrm{d}s_y/s_z = \mathrm{d}\Omega = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi. \tag{2·25}$$

On substituting from (2·23), (2·24), (2·25), (2·11) and (2·7) into (2·2) and (2·3) we obtain the following expressions for the Cartesian components **e** and **h**:

$$\left.\begin{aligned} e_x &= -\frac{\mathrm{i}A}{\pi}\int_0^{\alpha}\int_0^{2\pi} \cos^{\frac{1}{2}}\theta\sin\theta\{\cos\theta + (1-\cos\theta)\sin^2\phi\}\,\mathrm{e}^{\mathrm{i}kr_P\cos\epsilon}\,\mathrm{d}\theta\,\mathrm{d}\phi, \\ e_y &= \frac{\mathrm{i}A}{\pi}\int_0^{\alpha}\int_0^{2\pi} \cos^{\frac{1}{2}}\theta\sin\theta(1-\cos\theta)\cos\phi\sin\phi\,\mathrm{e}^{\mathrm{i}kr_P\cos\epsilon}\,\mathrm{d}\theta\,\mathrm{d}\phi, \\ e_z &= \frac{\mathrm{i}A}{\pi}\int_0^{\alpha}\int_0^{2\pi} \cos^{\frac{1}{2}}\theta\sin^2\theta\cos\phi\,\mathrm{e}^{\mathrm{i}kr_P\cos\epsilon}\,\mathrm{d}\theta\,\mathrm{d}\phi; \end{aligned}\right\} \tag{2·26}$$

$$\left.\begin{aligned} h_x &= \frac{\mathrm{i}A}{\pi}\int_0^{\alpha}\int_0^{2\pi} \cos^{\frac{1}{2}}\theta\sin\theta(1-\cos\theta)\cos\phi\sin\phi\,\mathrm{e}^{\mathrm{i}kr_P\cos\epsilon}\,\mathrm{d}\theta\,\mathrm{d}\phi, \\ h_y &= -\frac{\mathrm{i}A}{\pi}\int_0^{\alpha}\int_0^{2\pi} \cos^{\frac{1}{2}}\theta\sin\theta\{1-(1-\cos\theta)\sin^2\phi\}\,\mathrm{e}^{\mathrm{i}kr_P\cos\epsilon}\,\mathrm{d}\theta\,\mathrm{d}\phi, \\ h_z &= \frac{\mathrm{i}A}{\pi}\int_0^{\alpha}\int_0^{2\pi} \cos^{\frac{1}{2}}\theta\sin^2\theta\sin\phi\,\mathrm{e}^{\mathrm{i}kr_P\cos\epsilon}\,\mathrm{d}\theta\,\mathrm{d}\phi. \end{aligned}\right\} \tag{2·27}$$

Here $\cos\epsilon$ is given by (2·19), α is the angular semi-aperture on the image side, i.e. 2α is the angle which the diameter of the exit pupil subtends at the geometrical focus and A is the constant

$$A = \frac{kfl_0}{2} = \frac{\pi f l_0}{\lambda}. \tag{2·28}$$

The integration with respect to ϕ can immediately be carried out with the help of the following formulae† which are valid for any integral value of n:

$$\left.\begin{aligned}\int_0^{2\pi} \cos n\phi\, \mathrm{e}^{\mathrm{i}\rho\cos(\phi-\gamma)}\,\mathrm{d}\phi &= 2\pi\mathrm{i}^n J_n(\rho)\cos n\gamma,\\ \int_0^{2\pi} \sin n\phi\, \mathrm{e}^{\mathrm{i}\rho\cos(\phi-\gamma)}\,\mathrm{d}\phi &= 2\pi\mathrm{i}^n J_n(\rho)\sin n\gamma.\end{aligned}\right\} \tag{2·29}$$

Here $J_n(\rho)$ is the Bessel function of the first kind and order n. If in (2·26) and (2·27) we use the identities $\cos\phi\sin\phi = \frac{1}{2}\sin 2\phi$, $\sin^2\phi = \frac{1}{2}(1-\cos 2\phi)$ and apply (2·29), we finally obtain the following expressions for the components of the field vectors at a point P in the image region:

$$\left.\begin{aligned}e_x(P) &= -\mathrm{i}A(I_0 + I_2\cos 2\phi_P),\\ e_y(P) &= -\mathrm{i}AI_2\sin 2\phi_P,\\ e_z(P) &= -2AI_1\cos\phi_P,\end{aligned}\right\} \tag{2·30}$$

$$\left.\begin{aligned}h_x(P) &= -\mathrm{i}AI_2\sin 2\phi_P,\\ h_y(P) &= -\mathrm{i}A(I_0 - I_2\cos 2\phi_P),\\ h_z(P) &= -2AI_1\sin\phi_P,\end{aligned}\right\} \tag{2·31}$$

where

$$\left.\begin{aligned}I_0 = I_0(kr_P,\theta_P,\alpha) &= \int_0^\alpha \cos^{\frac{1}{2}}\theta\sin\theta(1+\cos\theta)\,J_0(kr_P\sin\theta\sin\theta_P)\,\mathrm{e}^{\mathrm{i}kr_P\cos\theta\cos\theta_P}\,\mathrm{d}\theta,\\ I_1 = I_1(kr_P,\theta_P,\alpha) &= \int_0^\alpha \cos^{\frac{1}{2}}\theta\sin^2\theta J_1(kr_P\sin\theta\sin\theta_P)\,\mathrm{e}^{\mathrm{i}kr_P\cos\theta\cos\theta_P}\,\mathrm{d}\theta,\\ I_2 = I_2(kr_P,\theta_P,\alpha) &= \int_0^\alpha \cos^{\frac{1}{2}}\theta\sin\theta(1-\cos\theta)\,J_2(kr_P\sin\theta\sin\theta_P)\,\mathrm{e}^{\mathrm{i}kr_P\cos\theta\cos\theta_P}\,\mathrm{d}\theta.\end{aligned}\right\} \tag{2·32}$$

Formulae (2·30) and (2·31) represent the analytic solution of our problem. They express the field at any point P (spherical polar co-ordinates r_P, θ_P, ϕ_P) in terms of the three integrals I_0, I_1 and I_2. We shall now study some consequences of these formulae.

3. The image field

It is convenient at this stage to introduce certain 'optical co-ordinates', which are a natural generalization of the co-ordinates (defined by (3·1*b*) below) employed

† These formulae may be derived as follows. We start from the integral representation of J_n:

$$\int_0^{2\pi} \mathrm{e}^{\mathrm{i}(n\phi+\rho\cos\phi)}\,\mathrm{d}\phi = 2\pi\mathrm{i}^n J_n(\rho)$$

(cf. Watson 1952, p. 20, (5)). We change ϕ into $\phi-\gamma$, multiply both sides of (2·30) by $\mathrm{e}^{\mathrm{i}n\gamma}$ and express the resulting formula as follows:

$$\int_0^{2\pi} \cos n\phi\, \mathrm{e}^{\mathrm{i}\rho\cos(\phi-\gamma)}\,\mathrm{d}\phi + \mathrm{i}\int_0^{2\pi} \sin n\phi\, \mathrm{e}^{\mathrm{i}\rho\cos(\phi-\gamma)}\,\mathrm{d}\phi = 2\pi\mathrm{i}^n J_n(\rho)\,[\cos n\gamma + \mathrm{i}\sin n\gamma].$$

Each side consists of two terms, one of which is an even function of γ and the other an odd function of γ. This is only possible if the even terms are equal to each other and so are the odd terms, and this implies (2·29).

frequently in connexion with diffraction in systems with low angular aperture. We define these optical co-ordinates by the formulae†

$$\left.\begin{aligned} u &= kr_P \cos\theta_P \sin^2\alpha = kz\sin^2\alpha, \\ v &= kr_P \sin\theta_P \sin\alpha \;\; = k\sqrt{(x^2+y^2)}\sin\alpha. \end{aligned}\right\} \tag{3·1a}$$

From now on we shall omit the subscript P in the symbol ϕ_P for the azimuthal angle, and specify the point P of observation by the three parameters u, v and ϕ ($u \gtrless 0$, $v \geqslant 0$, $0 \leqslant \phi < 2\pi$). The geometrical focal plane is given by $u = 0$, the axis by $v = 0$ and the edge of the geometrical shadow ($\sqrt{(x^2+y^2)} = \pm z\tan\alpha$) by $v = \pm u\sec\alpha$.

The integrals (2·32) are now regarded as functions of u and v,

$$\left.\begin{aligned} I_0(u,v) &= \int_0^\alpha \cos^{\frac{1}{2}}\theta \sin\theta(1+\cos\theta) J_0\left(\frac{v\sin\theta}{\sin\alpha}\right) \mathrm{e}^{\mathrm{i}u\cos\theta/\sin^2\alpha}\,\mathrm{d}\theta, \\ I_1(u,v) &= \int_0^\alpha \cos^{\frac{1}{2}}\theta \sin^2\theta J_1\left(\frac{v\sin\theta}{\sin\alpha}\right) \mathrm{e}^{\mathrm{i}u\cos\theta/\sin^2\alpha}\,\mathrm{d}\theta, \\ I_2(u,v) &= \int_0^\alpha \cos^{\frac{1}{2}}\theta \sin\theta(1-\cos\theta) J_2\left(\frac{v\sin\theta}{\sin\alpha}\right) \mathrm{e}^{\mathrm{i}u\cos\theta/\sin^2\alpha}\,\mathrm{d}\theta. \end{aligned}\right\} \tag{3·2}$$

We note that

$$I_n(-u,v) = I_n^*(u,v) \qquad (n = 0, 1, 2), \tag{3·3}$$

where the asterisk denotes the complex conjugate. From (2·30), (2·31) and (3·3) we note the following relations which exist between the components of the field vectors at any two points $P_1(u,v,\phi)$ and $P_2(-u,v,\phi)$, which are symmetrically situated with respect to the focal plane:

$$\left.\begin{aligned} e_x(-u,v,\phi) &= -e_x^*(u,v,\phi), \quad h_x(-u,v,\phi) = -h_x^*(u,v,\phi), \\ e_y(-u,v,\phi) &= -e_y^*(u,v,\phi), \quad h_y(-u,v,\phi) = -h_y^*(u,v,\phi), \\ e_z(-u,v,\phi) &= \;\; e_z^*(u,v,\phi), \quad h_z(-u,v,\phi) = \;\; h_z^*(u,v,\phi). \end{aligned}\right\} \tag{3·4}$$

If $|e_x|$ denotes the amplitude and ψ_x the phase of e_x, the first relation in (3·4) implies that

$$|e_x(-u,v,\phi)| = |e_x(u,v,\phi)|, \tag{3·5a}$$

and‡

$$\psi_x(-u,v,\phi) = -\psi_x(u,v,\phi) + \pi \pmod{2\pi}. \tag{3·5b}$$

Relation (3·5a) shows that for any two points which are symmetrically situated with respect to the focal plane, the amplitudes of the e_x's are the same, and (3·5b) shows that there exists also a simple relation between the phases. The appearance of the additive factor π on the right of (3·5b) is connected with the well-known *phase anomaly* at focus (cf. Born & Wolf 1959, § 8·8, (43)). From (3·4) we see that relations of the form (3·5a) holds also for all the other components. However, relation of the form (3·5b) holds for the phases of e_x, e_y, h_x and h_y only. For the phase of the remaining two components we have, in place of (3·5b), relations of the following form:

$$\psi_z(-u,v,\phi) = -\psi_z(u,v,\phi) \pmod{2\pi}. \tag{3·5c}$$

† When the angular aperture is small ($\alpha \ll 1$), formulae (3·1) reduce to

$$\left.\begin{aligned} u &\sim k(a/R)^2 z, \\ v &\sim k(a/R)\sqrt{(x^2+y^2)}, \end{aligned}\right\} \tag{3·1b}$$

where $\sin\alpha \sim \tan\alpha = a/R$, a being the radius of the exit pupil and R the distance between the exit pupil and the focal plane.

‡ The quantity mod 2π on the right of an equation denotes that the two sides of the equation are indeterminate to the extent of an additive constant $2m\pi$ where m is any integer.

Thus *the components of* **e** *and* **h** *in the direction of the axis of revolution of the system have no phase anomaly.*

From (2·30) and (2·31) it is also seen that

$$\left.\begin{aligned} h_x(u,v,\phi) &= -e_y(u,v,\phi-\tfrac{1}{2}\pi),\\ h_y(u,v,\phi) &= e_x(u,v,\phi-\tfrac{1}{2}\pi),\\ h_z(u,v,\phi) &= e_z(u,v,\phi-\tfrac{1}{2}\pi). \end{aligned}\right\} \tag{3·6}$$

Hence in any fixed plane of observation (u = constant), the **e** and **h** fields are the same but are rotated with respect to each other by 90° around the z-axis; this, of course, might have been expected, since the **e** and **h** fields in the object space have this relationship, and the laws relating to the transmission of these fields through the system are the same.

3·1. *Polarization of the image field*

To examine the state of polarization of the image field, we separate the real and imaginary parts of the integrals (3·2) and write

$$I_n(u,v) = I_n^{(r)}(u,v) + \mathrm{i}I_n^{(i)}(u,v) \qquad (n = 0,1,2), \tag{3·7}$$

where $I_n^{(r)}$ and $I_n^{(i)}$ are real. We also write

$$\mathbf{e}(u,v,\phi) = \mathbf{p}(u,v,\phi) + \mathrm{i}\mathbf{q}(u,v,\phi), \tag{3·8}$$

where **p** and **q** are real vectors; they are a pair of conjugate semi-diameters of the polarization ellipse of the electric vector. According to (2·30), (3·7) and (3·8)

$$\left.\begin{aligned} p_x(u,v,\phi) &= A(I_0^{(i)} + I_2^{(i)}\cos 2\phi), & q_x(u,v,\phi) &= -A(I_0^{(r)} + I_2^{(r)}\cos 2\phi),\\ p_y(u,v,\phi) &= AI_2^{(i)}\sin 2\phi, & q_y(u,v,\phi) &= -AI_2^{(r)}\sin 2\phi,\\ p_z(u,v,\phi) &= -2AI_1^{(r)}\cos\phi, & q_z(u,v,\phi) &= -2AI_1^{(i)}\cos\phi. \end{aligned}\right\} \tag{3·9}$$

Since, according to (3·4) the integrals I_0, I_1 and I_2 are all real when $u = 0$ (focal plane), it follows that

$$\left.\begin{aligned} p_x(0,v,\phi) &= 0, & q_x(0,v,\phi) &= -A[I_0(0,v) + I_2(0,v)\cos 2\phi],\\ p_y(0,v,\phi) &= 0, & q_y(0,v,\phi) &= -AI_2(0,v)\sin 2\phi,\\ p_z(0,v,\phi) &= -2AI_1(0,v)\cos\phi, & q_z(0,v,\phi) &= 0. \end{aligned}\right\} \tag{3·10}$$

From (3·10) we see that in the focal plane $\mathbf{p}\,.\,\mathbf{q} = 0$, i.e. the conjugate semi-diameters are at right angles to each other; hence in the focal plane, **p** and **q** are the semi-axes of the polarization ellipse of the electric vector. Moreover, the **p**-axis is perpendicular to the focal plane and the **q**-axis lies in the focal plane. Thus, *the polarization ellipse of the electric vector at any point in the focal plane is at right angles to the focal plane.* The angle $\chi(v,\phi)$ $(-\frac{1}{2}\pi < \chi \leqslant \frac{1}{2}\pi)$ between the plane of the polarization ellipse and the plane $\phi = 0$ (the xz-plane) is given by

$$\tan\chi(v,0) = \frac{q_y(0,v,\phi)}{q_x(0,v,\phi)} = \frac{I_2(0,v)\sin 2\phi}{I_0(0,v) + I_2(0,v)\cos 2\phi}, \tag{3·11}$$

and the two axes of the polarization ellipse are in the ratio

$$\rho(v,\phi) = \frac{|p_z(0,v,\phi)|}{+\sqrt{\{q_x^2(0,v,\phi)+q_y^2(0,v,\phi)\}}} = \frac{2\,|I_1(0,v)\cos\phi|}{+\sqrt{\{I_0^2(0,v)+I_2^2(0,v)+2I_0(0,v)\,I_2(0,v)\cos 2\phi\}}}. \tag{3·12}$$

Along the y-axis ($\phi = \frac{1}{2}\pi$ or $\frac{3}{2}\pi$) only one of the components in (3·10), namely q_x, is different from zero. Hence *at each point of the y-axis* and, in particular, at the focus itself, *the electric field is linearly polarized in the x-direction*, i.e. *in the direction*

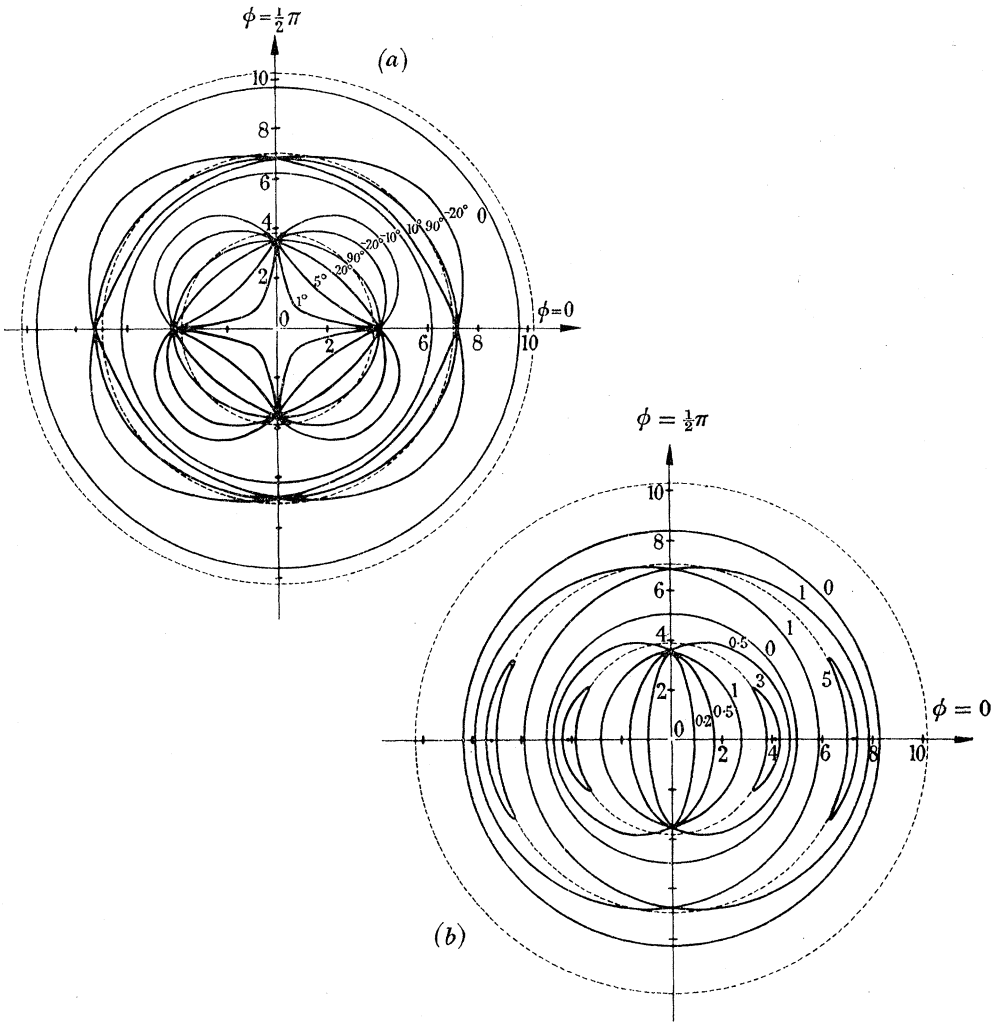

FIGURE 2. Polarization of the electric field in the focal plane of an aplanatic system of angular semi-aperture $\alpha = 60°$ on the image side: (*a*) Contours of $\chi(v,\phi)$; (*b*) Contours of $\rho(v,\phi)$. The field in the object space is linearly polarized with its electric vector in the azimuth $\phi = 0$. The field in the focal plane is, in general, elliptically polarized, with one axis (q) in the focal plane, and the other axis (p) perpendicular to the focal plane. $\chi(v,\phi)$ is the angle between the plane of the polarization ellipse of the electric vector and the meridional plane $\phi = 0$, and $\rho(v,\phi)$ is the ratio of the p and q axes. Radial distances are measured in v-units. The dashed lines indicate the dark rings of the Airy pattern represented by (4·12).

of vibrations of the electric field in the object space. Since **p** also vanishes when $I_1(0, v) = 0$, the electric field is also linearly polarized along circles centred on the focus, whose radii (measured in v-units) are given by the roots of this equation; however, along these circles, the direction of vibrations are not in general in the x-direction.

The behaviour of χ and ρ in the focal plane of an aplanatic system of angular semi-aperture $\alpha = 60°$ is shown in figure 2. The figure was computed from the formulae (3·11), (3·12) and (3·2) by means of the Manchester University Electronic Computer mark I. The details of these and other computations relating to this paper are given in a thesis by one of the authors (Richards 1956*b*).

Figure 2(*a*) shows that in the central nucleus of the image $\chi \sim 0°$, i.e. in this region the electric vector vibrates in planes nearly parallel to the plane which contains the direction of the electric vibrations in the object space and the axis of the system. According to figure 2(*b*) the axial component (e_z) is not negligible in all parts of the central nucleus.

The state of polarization of the field at points on the axis of the system ($v = 0$) is also of interest. Since $I_1(u, 0) = I_2(u, 0) = 0$ for all values of u, it follows from (2·30) that on the axis $e_y = e_z = 0$. Hence *the electric vector at each point on the axis of revolution in the image space is linearly polarized, and its direction is the same as the direction of the electric vector in the object space.*

3·2. *The energy density*

Let

$$\left.\begin{aligned} \langle w_e(u,v,\phi)\rangle &= \frac{1}{8\pi}\langle \mathbf{E}^2\rangle = \frac{1}{16\pi}(\mathbf{e}.\mathbf{e}^*),\\ \langle w_m(u,v,\phi)\rangle &= \frac{1}{8\pi}\langle \mathbf{H}^2\rangle = \frac{1}{16\pi}(\mathbf{h}.\mathbf{h}^*),\\ \langle w(u,v,\phi)\rangle &= \langle w_e(u,v,\phi)\rangle + \langle w_m(u,v,\phi)\rangle = \frac{1}{16\pi}(\mathbf{e}.\mathbf{e}^* + \mathbf{h}.\mathbf{h}^*), \end{aligned}\right\} \tag{3·13}$$

denote the electric energy density, the magnetic energy density and the total energy density, each averaged over a time interval which is large compared to the basic period $T = 2\pi/\omega$. From (2·30) and (3·13), we have

$$\left.\begin{aligned} \langle w_e(u,v,\phi)\rangle &= \frac{A^2}{16\pi}\{|I_0|^2 + 4\,|I_1|^2\cos^2\phi + |I_2|^2 + 2\cos 2\phi\,\mathscr{R}(I_0 I_2^*)\},\\ \langle w_m(u,v,\phi)\rangle &= \frac{A^2}{16\pi}\{|I_0|^2 + 4\,|I_1|^2\sin^2\phi + |I_2|^2 - 2\cos 2\phi\,\mathscr{R}(I_0 I_2^*)\},\\ \langle w(u,v,\phi)\rangle &= \frac{A^2}{8\pi}\{|I_0|^2 + 2\,|I_1|^2 + |I_2|^2\}, \end{aligned}\right\} \tag{3·14}$$

where $\mathscr{R}$ denotes the real part.

We note that because of (3·3) (as, of course, is also evident from (3·4))

$$\langle w_e(-u,v,\phi)\rangle = \langle w_e(u,v,\phi)\rangle, \tag{3·15}$$

with similar relations involving $\langle w_m\rangle$ and $\langle w\rangle$. Hence *the distributions of the time-averaged electric energy density, magnetic energy density and total energy density are symmetrical about the focal plane* $u = 0$.

Further, we see that

$$\langle w_m(u, v, \phi)\rangle = \langle w_e(u, v, \phi - \tfrac{1}{2}\pi)\rangle. \tag{3·16}$$

Thus *the distribution of the time-averaged magnetic energy density is identical with the distribution of the time-averaged electric energy density, but the distributions are rotated with respect to each other by* 90° *about the axis of the system.* We also note that $\langle w\rangle$ is independent of ϕ, so that *the loci of constant time-averaged total energy density are surfaces of revolution about the axis of revolution of the system.*

Since, for all u, $I_1(u, 0) = I_2(u, 0) = 0$, we see from (3·14) that along *the axis of revolution* in the image space†

$$\langle w_e\rangle = \langle w_m\rangle = \tfrac{1}{2}\langle w\rangle = \frac{A^2}{16\pi}\,|I_0(u, 0)|^2. \tag{3·17}$$

When $u = v = 0$, we have from (3·2)

$$\begin{aligned} I_0(0, 0) &= \int_0^\alpha \cos^{\frac{1}{2}}\theta \sin\theta(1 + \cos\theta)\,\mathrm{d}\theta \\ &= \tfrac{16}{15}\{1 - \tfrac{5}{8}(\cos^{\frac{3}{2}}\alpha)(1 + \tfrac{3}{5}\cos\alpha)\}, \end{aligned} \tag{3·18}$$

so that *at the focus* itself

$$\langle w_e\rangle = \langle w_m\rangle = \tfrac{1}{2}\langle w\rangle = \frac{A^2}{15\pi}\{1 - \tfrac{5}{8}(\cos^{\frac{3}{2}}\alpha)(1 + \tfrac{3}{5}\cos\alpha)\}. \tag{3·19}$$

In figure 3, contours, computed from (3·14), of the time-averaged electric energy density in the focal planes of aplanatic systems of selected angular semi-apertures α are shown. The contours of the time-averaged magnetic energy density are, according to (3·16), identical with those for the time-averaged electric energy, but are rotated by 90° about the normal to the plane of the figure.

The first diagram in figure 3 represents the limiting case $\alpha \to 0$. We shall see later (§4) that in this case our solution reduces to that obtained on the basis of the usual scalar diffraction theory. Thus the first diagram is the ordinary *Airy diffraction pattern*: the contours are circles, the first zero contour being given by $v = 1{\cdot}22\pi = 3{\cdot}83$. As α increases, the pattern is seen to lose its rotational symmetry; the contours in the neighbourhood of the focus are then approximately elliptical, with their major axes in the direction of the electric vector in the object space ($\phi = 0$). Further away from the centre the contours are of a more complicated form and closed loops appear around certain points in the azimuths $\phi = 0, \frac{1}{2}\pi, \pi, \frac{3}{2}\pi$. For large values of α, the ellipticity of the contours becomes quite pronounced and the first minimum in the azimuth $\phi = 0$ is well outside the first zero of the Airy pattern, while in the azimuth $\phi = \frac{1}{2}\pi$ it is well inside. Hence, *if the wave in the object space is linearly polarized and detectors of electric energy* (e.g. *a photographic plate*) *are used, our solution predicts an increase in the resolving power in wide aperture systems for measurements in the azimuth at right angles to that of the electric vector of the incident wave.* This conclusion is in agreement with a prediction of Hopkins (1943, 1945) and appears to be supported by early experiments recorded by Carpenter (1901).

† Curves which illustrate the behaviour of these quantities along the axis of revolution are given in Richards (1956*a*, p. 358). This paper also contains curves which give the ratio (as function of α) between the values at focus computed from (3·19) and those computed by the application of the Huygens–Kirchhoff scalar diffraction theory.

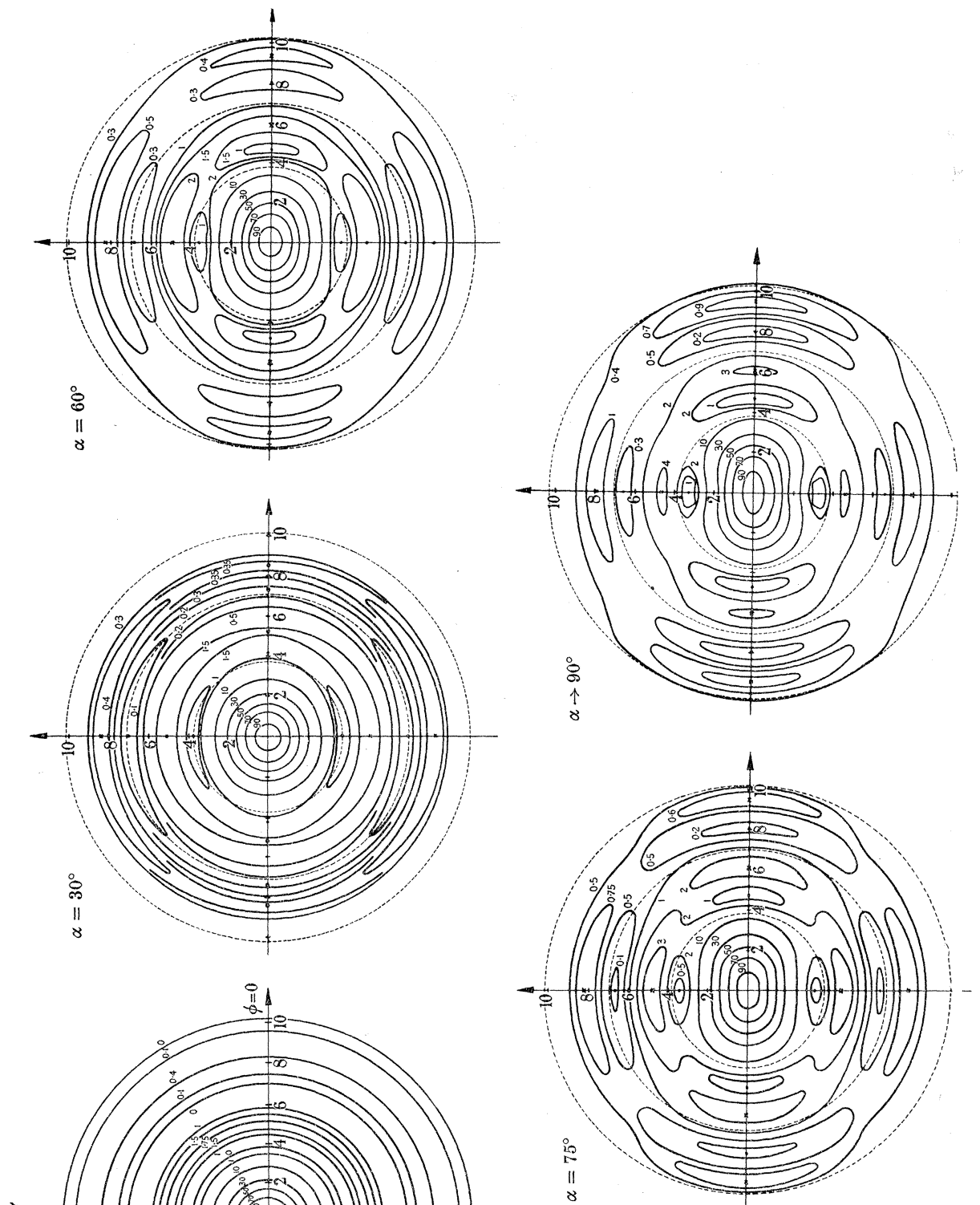

FIGURE 3

For legend see foot of p. 371.

The behaviour of the time-averaged electric energy density along the azimuths $\phi = 0$ and $\phi = \frac{1}{2}\pi$ in the focal plane is shown in detail in figure 4. It is seen that along the azimuth $\phi = 0$, the minima are not true zeros; the minima are, however, zeros along the azimuth $\phi = \frac{1}{2}\pi$. The behaviour of the time averaged total energy density, which as already noted is independent of the azimuth, is shown in figure 5*a*.

3·3. *The energy flow (Poynting vector)*

The time-averaged Poynting vector is given by

$$\langle \mathbf{S} \rangle = \frac{c}{4\pi} \langle \mathbf{E} \wedge \mathbf{H} \rangle = \frac{c}{8\pi} \mathscr{R}(\mathbf{e} \wedge \mathbf{h}^*). \tag{3·20}$$

On substituting from (2·30) and (2·31) into (3·20) we obtain the following expressions for the components of $\langle \mathbf{S} \rangle$:

$$\left.\begin{aligned} \langle S_x \rangle &= \frac{cA^2}{4\pi} \cos\phi \mathscr{I}\{I_1(I_2^* - I_0^*)\}, \\ \langle S_y \rangle &= \frac{cA^2}{4\pi} \sin\phi \mathscr{I}\{I_1(I_2^* - I_0^*)\}, \\ \langle S_z \rangle &= \frac{cA^2}{8\pi} \{|I_0|^2 - |I_2|^2\}, \end{aligned}\right\} \tag{3·21}$$

where $\mathscr{I}$ denotes the imaginary part.

We see that the magnitude of the time-averaged Poynting vector is given by

$$|\langle \mathbf{S} \rangle| = \frac{cA^2}{8\pi} \{(|I_0|^2 - |I_2|^2) + 4(\mathscr{I}[I_1(I_2^* - I_0^*)])^2\}^{\frac{1}{2}}. \tag{3·22}$$

Since this expression is independent of the azimuth ϕ, *the loci of constant* $|\langle \mathbf{S} \rangle|$ *are surfaces of revolution about the axis of the system.*

Along the axis of revolution in the image space ($v = 0$) (3·22) reduces to

$$|\langle \mathbf{S} \rangle| = \frac{cA^2}{8\pi} |I_0(u, 0)|^2; \tag{3·23}$$

on comparison with (3·17) we see that the relation $|\langle \mathbf{S} \rangle| = c\langle w \rangle$ holds at all points on the axis.

From (3·21) we readily deduce that *at each point in the image space the Poynting vector lies in the meridional plane which contains that point*, and that it makes an angle γ with the positive z-axis, given by

$$\left.\begin{aligned} \sin\gamma &= \frac{2\,|\mathscr{I}\{I_1(I_2^* - I_0^*)\}|}{[(|I_0|^2 - |I_2|^2)^2 + 4\{\mathscr{I}[I_1(I_2^* - I_0^*)]\}^2]^{\frac{1}{2}}}, \\ \cos\gamma &= \frac{|I_0|^2 - |I_2|^2}{[(|I_0|^2 - |I_2|^2)^2 + 4\{\mathscr{I}[I_1(I_2^* - I_0^*)]\}^2]^{\frac{1}{2}}}. \end{aligned}\right\} \tag{3·24}$$

FIGURE 3. Contours of the time-averaged electric energy density $\langle w_e \rangle$ in the focal plane of an aplanatic system of angular semi-aperture α on the image side.

The field in the object space is linearly polarized with its electric vector in the azimuth $\phi = 0$. The crosses, dots and circles indicate maxima, minima and saddle points respectively. Radial distances are measured in v-units. The values are normalized to 100 at the focus. The first figure, which represents the limiting case $\alpha \to 0$, is identical with the classical intensity pattern of Airy, characterized by (4·12). The dashed circles in the other figures represent the dark rings in the Airy pattern.

Using the relations (3·3) we see from (3·22) that

$$|\langle \mathbf{S}(-u, v, \phi)\rangle| = |\langle \mathbf{S}(u, v, \phi)\rangle|, \tag{3·25}$$

i.e. *the magnitude of the time-averaged Poynting vector is symmetric with respect to the focal plane.* Further, from (3·24), we have

$$\sin\{\gamma(-u, v, \phi)\} = -\sin\{\gamma(u, v, \phi)\}, \quad \cos\{\gamma(-u, v, \phi)\} = \cos\{\gamma(u, v, \phi)\},$$

so that, for $u \neq 0$,

$$\gamma(-u, v, \phi) = 2\pi - \gamma(u, v, \phi). \tag{3·26}$$

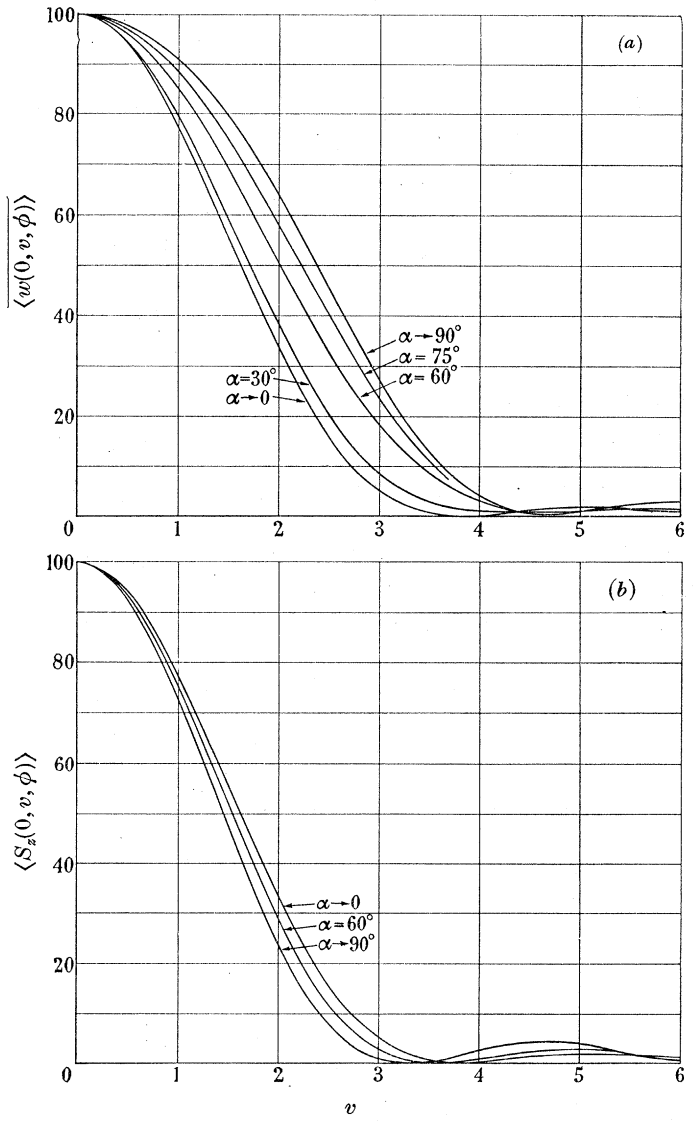

FIGURE 4. The variation of the time-averaged electric energy $\langle w_e\rangle$ along the two principal meridional sections of the focal plane in an aplanatic system of selected angular semi-aperture α in the image space: (*a*) Meridional section $\phi = 0$; (*b*) meridional section $\phi = \frac{1}{2}\pi$.

The field in the object space is linearly polarized with its electric vector in the azimuth $\phi = 0$. The values are normalized to 100 at the focus. The curves representing the limiting case $\alpha \to 0$ are identical with the classical Airy intensity curve given by (4·12).

When $u = 0$, we have from (3·21), since the integrals $I_n(0, v)$ $(n = 0, 1, 2)$ are all real,

$$\langle S_x \rangle = \langle S_y \rangle = 0, \tag{3·27}$$

i.e. *at any point in the focal plane, the time-averaged energy flow*, as represented by the Poynting vector, *is perpendicular to the focal plane.* Moreover,

$$\left.\begin{aligned} &\langle S_z \rangle > 0 \quad \text{if} \quad |I_0(0, v)| > |I_2(0, v)|, \\ \text{and} \quad &\langle S_z \rangle < 0 \quad \text{if} \quad |I_0(0, v)| < |I_2(0, v)|. \end{aligned}\right\} \tag{3·28}$$

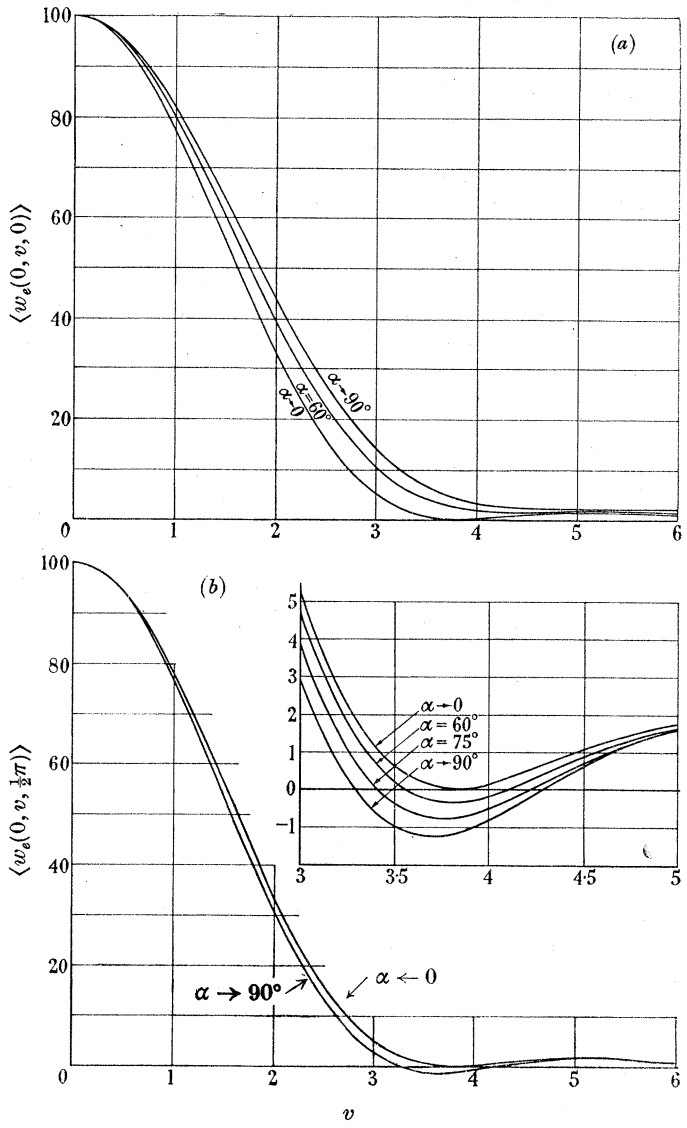

FIGURE 5. The variation of the time-averaged total energy density (*a*) and of the only non-vanishing component of the time-averaged energy flow (*b*) along any meridional section ϕ = const. of the focal plane of an aplanatic system of angular semi-aperture α on the image side. The curves represent the case of polarized as well as unpolarized wave. The curves in (*a*) also represent the variation of the time-averaged electric and magnetic energy densities of an unpolarized wave. The values are normalized to 100 at focus.

The second line implies that at points in the focal plane where $|I_2(0,v)| > |I_0(0,v)|$, the energy flow is directed back towards the object space.†

The computed curves, which show the behaviour of $\langle S_z \rangle$ in the focal plane of aplanatic systems of selected angular semi-apertures α are shown in figure 5. It is of interest to note that these (normalized) curves do not substantially depend on α. In particular, in the region of the Rayleigh limit, where the curves do not drop below 0·8 of their respective maxima, the differences (for any particular v-value in this region) are less than 1 %, in the full range ($0 \leqslant \alpha \leqslant 90°$) of the angular semi-aperture.

4. Approximate form of the solution for systems with a small angular aperture

The structure of images in systems with small angular aperture is usually investigated on the basis of a scalar diffraction theory. The intensity distribution in the focal plane of an aberration-free system of revolution which images a point source was first determined in this way by Airy (1835). His analysis was extended by Lommel (1885) and Struve (1886), who determined the distribution in the whole three-dimensional neighbourhood of the focus. It is of interest to examine the form which our solution takes when the angular semi-aperture α is small, and to compare it with these early solutions.

When α is small enough and u and v are not large compared with unity, we may replace the trigonometric factors in the amplitudes of the integrands in I_0, I_1 and I_2 in (3·2) by the leading term in their series expansions. In the exponents of the integrands we retain both the first and the second term in the expansion of $\cos\theta$ since, when u is not small compared to unity, the contribution of the second term is evidently also significant. The integrals then become

$$\left.\begin{aligned} I_0(u,v) &= 2\,\mathrm{e}^{\mathrm{i}u/\alpha^2}\int_0^\alpha \theta J_0\left(\frac{v\theta}{\alpha}\right)\mathrm{e}^{-\mathrm{i}u\theta^2/2\alpha^2}\,\mathrm{d}\theta,\\ I_1(u,v) &= \mathrm{e}^{\mathrm{i}u/\alpha^2}\int_0^\alpha \theta^2 J_1\left(\frac{v\theta}{\alpha}\right)\mathrm{e}^{-\mathrm{i}u\theta^2/2\alpha^2}\,\mathrm{d}\theta,\\ I_2(u,v) &= \tfrac{1}{2}\mathrm{e}^{\mathrm{i}u/\alpha^2}\int_0^\alpha \theta^3 J_2\left(\frac{v\theta}{\alpha}\right)\mathrm{e}^{-\mathrm{i}u\theta^2/2\alpha^2}\,\mathrm{d}\theta. \end{aligned}\right\} \tag{4·1}$$

For small x, $J_n(x) \sim x^n$ and we see that I_1 and I_2 are of lower order in α than I_0, so that these integrals may be neglected in comparison with I_0. Hence (2·30) and (2·31) reduce to

$$e_x = h_y = -\mathrm{i}AI_0, \tag{4·2}$$

$$e_y = e_z = h_x = h_z = 0. \tag{4·3}$$

Thus *in a system with a small angular aperture, the field in the image region is effectively linearly polarized; and the directions of the two field vectors in the image region are the same as their directions in the object space.* We see that in this case the image field is completely specified by one (complex) scalar wave (e_x or h_y), which is represented by the integral I_0 alone.

† This interesting fact was already deduced by Ignatowsky (1919).

To evaluate I_0 we introduce a new variable $\rho = \theta/\alpha$ and obtain

$$I_0(u,v) = 2\alpha^2 \mathrm{e}^{\mathrm{i}u/\alpha^2} \int_0^1 \rho J_0(v\rho)\, \mathrm{e}^{-\frac{1}{2}\mathrm{i}u\rho^2}\, \mathrm{d}\rho. \tag{4·4}$$

The integral on the right may be evaluated in terms of two of the Lommel functions

$$U_n(u,v) = \sum_{s=0}^{\infty} (-1)^s \left(\frac{u}{v}\right)^{n+2s} J_{n+2s}(v), \tag{4·5}$$

introduced by Lommel in the analysis already referred to. In fact†

$$\int_0^1 \rho J_0(v\rho)\, \mathrm{e}^{\frac{1}{2}\mathrm{i}u\rho^2}\, \mathrm{d}\rho = \frac{1}{u}\, \mathrm{e}^{-\frac{1}{2}\mathrm{i}u} [U_1(u,v) + \mathrm{i}U_2(u,v)]. \tag{4·6}$$

From (4·2), (4·4) and (4·6) we obtain the following expressions for the only two non-vanishing field components in the image region:

$$e_x = h_y = -2\mathrm{i}A\alpha^2 \frac{\exp\left\{\mathrm{i}u\left(\frac{1}{\alpha^2} - \frac{1}{2}\right)\right\}}{u} [U_1(u,v) + \mathrm{i}U_2(u,v)]. \tag{4·7}$$

The expressions (3·14) for the energy densities now become

$$\begin{aligned} \langle w_e \rangle = \langle w_m \rangle = \tfrac{1}{2}\langle w \rangle &= \frac{A^2}{16\pi} |I_0|^2 \\ &= \frac{A^2\alpha^4}{4\pi} \frac{1}{u^2} [U_1^2(u,v) + U_2^2(u,v)]. \end{aligned} \tag{4·8}$$

Formulae (3·21) for the components of the Poynting vector become

$$\begin{aligned} \langle S_x \rangle &= \langle S_y \rangle = 0, \\ \langle S_z \rangle &= \frac{cA^2}{8\pi} |I_0|^2 = \frac{cA^2\alpha^4}{2\pi} \frac{1}{u^2} [U_1^2(u,v) + U_2^2(u,v)]. \end{aligned} \tag{4·9}$$

Thus, *when α is small, the energy flow,* as represented by the time-averaged Poynting vector, *is in the direction of the positive z-axis.* On comparison of (4·8) and (4·9) we also see that the relation

$$|\langle \mathbf{S} \rangle| = c\langle w \rangle \tag{4·10}$$

then holds everywhere in the image region.

The expressions on the right of (4·8) and (4·9) are proportional to the classical solution of Lommel and Struve for the 'intensity' in the image region of an aberration-free system.

† See Lommel (1885) or Born & Wolf (1959, § 8·8).

The U_n-series converge for all u and v values, but are convenient for computations only when $|u/v| < 1$. When $|u/v| > 1$, equation (4·4) may be evaluated in terms of two of the Lommel's V_n-functions defined by

$$V_n(u,v) = \sum_{s=0}^{\infty} (-1)^s (v/u)^{n+2s} J_{n+2s}(v).$$

The U and the V functions are related by the formula [cf. Watson 1952, p. 537]

$$U_{n+1}(u,v) - (-1)^{n-1} V_{n-1}(u,v) = \sin\{\tfrac{1}{2}(u + v^2/u - n\pi)\}.$$

When $u = 0$ (focal plane) we have from (4·7), since

$$\lim_{u\to 0}\frac{U_1(u,v)}{u} = \frac{J_1(v)}{v}, \qquad \lim_{u\to 0}\frac{U_2(u,v)}{u} = 0,$$

$$e_x = h_y = -\mathrm{i}A\alpha^2\left(\frac{2J_1(v)}{v}\right), \tag{4·11}$$

and (4·8) and (4·9) become

$$\langle w_e\rangle = \langle w_m\rangle = \tfrac{1}{2}\langle w\rangle = \frac{1}{2c}|\langle \mathbf{S}\rangle| = \frac{A^2\alpha^4}{16\pi}\left(\frac{2J_1(v)}{v}\right)^2. \tag{4·12}$$

The expression on the right of (4·12) is proportional to the classical solution of Airy for the 'intensity' in the image plane of an aberration-free system.

5. Images formed by an unpolarized wave

So far we have assumed that the wave entering the system is linearly polarized. We shall now briefly consider the case when the wave is unpolarized.

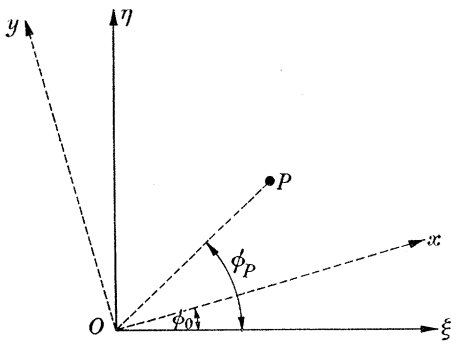

Figure 6. Change of reference system. Section at right angles to the axis of revolution.

An unpolarized incident wave may become partially polarized on refraction (or on reflexion) at the successive surfaces of the system. However, if the angles of incidence at each surface are small, as is assumed, the degree of polarization introduced in this way will also be small and we may, therefore, consider the wave emerging from the system to be effectively unpolarized. If further we assume that the wave is quasi-monochromatic, with mean angular frequency ω, the expressions for the time-averaged energy densities and energy flow may be obtained by averaging the corresponding expressions relating to the polarized wave over all possible states of polarization.

To carry out this averaging it is convenient to choose a new set of co-ordinate axes $O\xi$, $O\eta$ at right angles to each other and to the axis of revolution. Let ϕ_0 be the angle between $O\xi$ and the direction Ox of the electric vibrations of the incident wave, and let ϕ_P be the angle between $O\xi$ and the azimuth which contains the point P of observation (see figure 6). Then the first formula in (3·14) for the time-averaged electric energy density at a point P of observation becomes, since $\phi = \phi_P - \phi_0$,

$$\langle w_e\rangle = (A^2/16\pi)\{|I_0|^2 + 4|I_1|^2\cos^2(\phi_P-\phi_0) + |I_2|^2 + 2\mathscr{R}(I_0 I_2^*)\cos[2(\phi_P-\phi_0)]\}. \tag{5·1}$$

If we denote by vertical bars an average over all possible values of ϕ_0 $(0 \leqslant \phi_0 < 2\pi)$ it follows that when the incident wave is unpolarized the time-averaged electric energy density at P is†

$$\overline{\langle w_e \rangle} = \frac{1}{2\pi}\int_0^{2\pi} \langle w_e \rangle \, \mathrm{d}\phi = \frac{A^2}{16\pi}\{|I_0|^2 + 2\,|I_1|^2 + |I_2|^2\}. \tag{5·2}$$

We note that $\overline{\langle w_e \rangle}$ is independent of ϕ_P so that the loci of the constant averaged electric energy are surfaces of revolution about the axis of the system, as was to be expected from symmetry. In a strictly similar manner we obtain expressions for $\overline{\langle w_m \rangle}$ and $\overline{\langle w \rangle}$, and we have in all

$$\overline{\langle w_e \rangle} = \overline{\langle w_m \rangle} = \tfrac{1}{2}\overline{\langle w \rangle} = (A^2/16\pi)\{|I_0|^2 + 2\,|I_1|^2 + |I_2|^2\}. \tag{5·3}$$

Comparison with the last expression in (3·14) shows that each term in (5·3) is also equal to one-half of the total energy density of the polarized wave. The distribution of this quantity in the focal plane has already been given in figure 5.

Next consider the energy flow. It is again convenient to transform first the expressions for the polarized case so that the co-ordinates which appear in them are referred to the fixed axes $O\xi$, $O\eta$, independent of the direction of polarization of the incident wave. Let $\langle S_\xi \rangle$ and $\langle S_\eta \rangle$ be the components of the time-averaged Poynting vector in the direction $O\xi$, $O\eta$. Then

$$\left.\begin{aligned} \langle S_\xi \rangle &= \langle S_x \rangle \cos\phi_0 - \langle S_y \rangle \sin\phi_0, \\ \langle S_\eta \rangle &= \langle S_x \rangle \sin\phi_0 + \langle S_y \rangle \cos\phi_0. \end{aligned}\right\} \tag{5·4}$$

On substituting for $\langle S_x \rangle$ and $\langle S_y \rangle$ from (3·21), where again ϕ is replaced by $\phi_P - \phi_0$ we find that $\langle S_\xi \rangle$ and $\langle S_\eta \rangle$ are given by the same formulae as $\langle S_x \rangle$ and $\langle S_y \rangle$ but with ϕ_P written in place of ϕ. Since these formulae are independent of ϕ_0, they apply not only to the polarized wave but also to the unpolarized one, i.e.

$$\overline{\langle S_\xi \rangle} = \langle S_\xi \rangle = (cA^2/4\pi)\,\mathscr{I}\{I_1(I_2^* - I_0^*)\}\cos\phi_P, \tag{5·5a}$$

$$\overline{\langle S_\eta \rangle} = \langle S_\eta \rangle = (cA^2/4\pi)\,\mathscr{I}\{I_1(I_2^* - I_0^*)\}\sin\phi_P. \tag{5·5b}$$

The remaining component, in the direction of the axis of revolution is, as before, given by the last formulae in (3·21),

$$\overline{\langle S_z \rangle} = \langle S_z \rangle = (cA^2/8\pi)\{|I_0|^2 - |I_2|^2\}. \tag{5·5c}$$

It immediately follows that the results expressed by (3·22), (3·23) and (3·24) hold also for the unpolarized wave.

When the angular semi-aperture α is sufficiently small $(\alpha \to 0)$ one may, as in § 4, neglect I_1 and I_2 in comparison with I_0. The above formulae then become, when (4·4) and (4·6) are used:

$$\overline{\langle w_e \rangle} = \overline{\langle w_m \rangle} = \tfrac{1}{2}\overline{\langle w \rangle} = \frac{A^2\alpha^4}{4\pi}\frac{1}{u^2}[U_1^2(u,v) + U_2^2(u,v)], \tag{5·6}$$

$$\overline{\langle S_\xi \rangle} = \overline{\langle S_\eta \rangle} = 0, \tag{5·7a}$$

$$\overline{\langle S_z \rangle} = \frac{cA^2\alpha^4}{2\pi}\frac{1}{u^2}[U_1^2(u,v) + U_2^2(u,v)]. \tag{5·7b}$$

† The factor l_0 entering the constant $A = \pi f l_0/\lambda$ may now be related to the averaged electric energy density $\overline{\langle w_e \rangle}_i$ of the incident wave in the object space by the formula

$$\overline{\langle w_e \rangle}_i = (1/16\pi)\,(\mathbf{e}_0 \cdot \mathbf{e}_0^*) = (1/16\pi)\,l_0^2.$$

Here U_1 and U_2 are again two of the functions defined by (4·5). For points of observation in the focal plane ($u = 0$) these formulae reduce to

$$\overline{\langle w_e \rangle} = \overline{\langle w_m \rangle} = \tfrac{1}{2}\overline{\langle w \rangle} = \frac{1}{2c}\,|\langle \mathbf{S} \rangle| = \frac{A^2\alpha^4}{16\pi}\left(\frac{2J_1(v)}{v}\right)^2. \tag{5·8}$$

As already noted in connexion with the corresponding formulae (4·8), (4·9) and (4·12) the quantities on the right-hand sides of (5·6) and (5·7*b*) are proportional to the classical solution of Lommel and Struve for the 'intensity' in the image region

TABLE 1. COMPARISON OF DATA RELATING TO THE STRUCTURE OF THE ELECTROMAGNETIC FIELD IN THE FOCAL PLANE ($u = 0$) OF APLANATIC SYSTEMS FOR SELECTED VALUES OF ANGULAR SEMI-APERTURE α ON THE IMAGE SIDE

$\langle w_e(u, v, \phi) \rangle$ and $\langle w_m(u, v, \phi) \rangle$ represent the time-averaged electric energy density and magnetic energy density respectively. $\langle S_z(u, v, \phi) \rangle$ represents the only non-vanishing component (in the direction of the axis of the system) of the time-averaged energy flow. All the quantities are normalized to 100 at the focus $u = v = 0$.

Quantities with a vertical bar refer to an unpolarized wave, the others to a linearly polarized wave in the object space, with its electric vector in the azimuth $\phi = 0$. The third row in each division of the table also represents one-half of the time-averaged total energy density $\langle w \rangle$, i.e.

$$\overline{\langle w_e \rangle} = \overline{\langle w_m \rangle} = \tfrac{1}{2}\overline{\langle w \rangle} = \tfrac{1}{2}\langle w \rangle.$$

The entries in the column $\alpha \to 0$ are identical with values given by the intensity formulae (4·12) of Airy.

value	quantity	v: $\alpha \to 0$	$\alpha = 30°$	$\alpha = 60°$	$\alpha = 75°$	$\alpha \to 90°$
80	$\langle w_e(0,v,0)\rangle = \langle w_m(0,v,\frac{1}{2}\pi)\rangle$	0·93	0·98	1·16	1·33	1·48
	$\langle w_e(0,v,\frac{1}{2}\pi)\rangle = \langle w_m(0,v,0)\rangle$	0·93	0·93	0·89	0·86	0·84
	$\overline{\langle w_e(0,v,\phi)\rangle} = \overline{\langle w_m(0,v,\phi)\rangle}$	0·93	0·95	1·00	1·03	1·05
	$\langle S_z(0,v,\phi)\rangle = \overline{\langle S_z(0,v,\phi)\rangle}$	0·93	0·93	0·93	0·93	0·91
60	$\langle w_e(0,v,0)\rangle = \langle w_m(0,v,\frac{1}{2}\pi)\rangle$	1·40	1·47	1·73	1·95	2·10
	$\langle w_e(0,v,\frac{1}{2}\pi)\rangle = \langle w_m(0,v,0)\rangle$	1·40	1·38	1·32	1·30	1·26
	$\overline{\langle w_e(0,v,\phi)\rangle} = \overline{\langle w_m(0,v,\phi)\rangle}$	1·40	1·42	1·50	1·55	1·58
	$\langle S_z(0,v,\phi)\rangle = \overline{\langle S_z(0,v,\phi)\rangle}$	1·40	1·40	1·39	1·38	1·36
40	$\langle w_e(0,v,0)\rangle = \langle w_m(0,v,\frac{1}{2}\pi)\rangle$	1·83	1·94	2·27	2·50	2·63
	$\langle w_e(0,v,\frac{1}{2}\pi)\rangle = \langle w_m(0,v,0)\rangle$	1·83	1·82	1·76	1·70	1·64
	$\overline{\langle w_e(0,v,\phi)\rangle} = \overline{\langle w_m(0,v,\phi)\rangle}$	1·83	1·87	1·99	2·05	2·10
	$\langle S_z(0,v,\phi)\rangle = \overline{\langle S_z(0,v,\phi)\rangle}$	1·83	1·83	1·82	1·82	1·79
20	$\langle w_e(0,v,0)\rangle = \langle w_m(0,v,\frac{1}{2}\pi)\rangle$	2·36	2·50	2·93	3·14	3·23
	$\langle w_e(0,v,\frac{1}{2}\pi)\rangle = \langle w_m(0,v,0)\rangle$	2·36	2·33	2·22	2·16	2·10
	$\overline{\langle w_e(0,v,\phi)\rangle} = \overline{\langle w_m(0,v,\phi)\rangle}$	2·36	2·41	2·56	2·66	2·73
	$\langle S_z(0,v,\phi)\rangle = \overline{\langle S_z(0,v,\phi)\rangle}$	2·36	2·36	2·34	2·33	2·27
1st min.	$\langle w_e(0,v,0)\rangle = \langle w_m(0,v,\frac{1}{2}\pi)\rangle$	3·83	4·15	4·75	4·75	4·70
	$\langle w_e(0,v,\frac{1}{2}\pi)\rangle = \langle w_m(0,v,0)\rangle$	3·83	3·77	3·56	3·40	3·29
	$\overline{\langle w_e(0,v,\phi)\rangle} = \overline{\langle w_m(0,v,\phi)\rangle}$	3·83	3·95	—	—	—
	$\langle S_z(0,v,\phi)\rangle = \overline{\langle S_z(0,v,\phi)\rangle}$	3·83	3·80	3·80	3·76	3·70
2nd min.	$\langle w_e(0,v,0)\rangle = \langle w_m(0,v,\frac{1}{2}\pi)\rangle$	7·02	7·03	8·10	8·10	8·00
	$\langle w_e(0,v,\frac{1}{2}\pi)\rangle = \langle w_m(0,v,0)\rangle$	7·02	6·95	6·85	6·70	6·55
	$\overline{\langle w_e(0,v,\phi)\rangle} = \overline{\langle w_m(0,v,\phi)\rangle}$	7·02	—	—	—	—
	$\langle S_z(0,v,\phi)\rangle = \overline{\langle S_z(0,v,\phi)\rangle}$	7·02	7·02	7·02	7·02	6·95

of an aberration-free system, and the expression on the right-hand side of (5·8) is proportional to the classical solution of Airy for the 'intensity' in the image plane.

Let us denote by suffix zero expressions (such as (5·6)) which refer to a low aperture system ($\alpha \to 0$). Then from (4·8), (4·9), (5·3), (5·6) and (5·7) the following relations are seen to hold:

$$\langle S_z \rangle_0 = \overline{\langle S_z \rangle_0} = c\overline{\langle w \rangle_0} \leqslant c\overline{\langle w \rangle} = c\langle w \rangle, \tag{5·9}$$

$$\overline{\langle w_e \rangle_0} = \overline{\langle w_m \rangle_0} = \tfrac{1}{2}\overline{\langle w \rangle_0} \leqslant \tfrac{1}{2}\overline{\langle w \rangle} = \overline{\langle w_m \rangle} = \overline{\langle w_e \rangle}. \tag{5·10}$$

In practice, detectors of electric energy are usually used. The relation $\langle w_e \rangle_0 \leqslant \overline{\langle w_e \rangle}$ implies that the pattern recorded in any particular receiving plane (w = const.) will then be broader in a system with a wide angular aperture than in one with a low angular aperture. For patterns in the focal plane, this effect is seen in figure 5*a* which also shows that $\overline{\langle w_e \rangle}$ has no exact zeros, nor pronounced subsidiary maxima.

Finally, some of the main data which relate to the structure of the image in the focal plane of an aplanatic system are summarized in table 1. Data relating to both polarized and unpolarized incident waves are given.

The very extensive calculations on which the diagrams and the table in this paper are based were carried out on the Manchester University Electronic Computer mark I. We are indebted to Mr R. A. Brooker for helpful advice on computational techniques. We are also obliged to Miss B. Wood for help with construction of the contour diagrams shown in figures 2 and 3.

One of us (B. R.) wishes to acknowledge the award of a grant from the National Research and Development Corporation; the other (E. W.) is indebted to Manchester University for the award of an Imperial Chemistry Industries Research Fellowship during the tenure of which the main part of this work was carried out. The investigation was completed when he was a guest at the Institute of Mathematical Sciences (Division of Electromagnetic Research), New York University and was partially supported by the U.S. Air Force Cambridge Research Center, Air Research and Development Command, under contract No. AF 19 (604) 5238.

References

Airy, G. B. 1835 *Trans. Camb. Phil. Soc.* **5**, 283.
Born, M. & Wolf, E. 1959 *Principles of optics.* London: Pergamon Press.
Burtin, R. 1956 *Optica Acta*, **3**, 104.
Carpenter, W. B. 1901 *The microscope and its revelations*, 8th ed., p. 381. London: Churchill.
Focke, J. 1957 *Optica Acta*, **4**, 124.
Hopkins, H. H. 1943 *Proc. Phys. Soc.* **55**, 116.
Hopkins, H. H. 1945 *Nature, Lond.* **155**, 275.
Ignatowsky, V. S. 1919 *Trans. Opt. Inst. Petrograd*, vol. I, paper IV.
Ignatowsky, V. S. 1920 *Trans. Opt. Inst. Petrograd*, vol. I, paper V.
Lommel, E. 1885 *Abh. bayer. Akad. Wiss.* **15**, 233.
Richards, B. 1956*a* Contribution to *Symposium on astronomical optics and related subjects* (ed. Z. Kopal), p. 352. Amsterdam: North Holland Publishing Company.
Richards, B. 1956*b* Ph.D. Thesis, University of Manchester.
Richards, B. & Wolf, E. 1956 *Proc. Phys. Soc.* B, **69**, 854.
Struve, H. 1886 *Mém. Acad. St. Petersbourgh* (7), **34**, 1.
Watson, G. N. 1952 *Treatise on the theory of Bessel functions*, 2nd ed. Cambridge Univ. Press.
Wolf, E. 1959 *Proc. Roy. Soc.* A, **253**, 349.