X

Interference and diffraction with partially coherent light

10.1 Introduction

So far we have been mainly concerned with monochromatic light produced by a point source. Light from a real physical source is never strictly monochromatic, since even the sharpest spectral line has a finite width. Moreover, a physical source is not a point source, but has a finite extension, consisting of very many elementary radiators (atoms). The disturbance produced by such a source may be expressed, according to Fourier's theorem, as the sum of strictly monochromatic and therefore infinitely long wave trains. The elementary monochromatic theory is essentially concerned with a single component of this Fourier representation.

In a monochromatic wave field the amplitude of the vibrations at any point P is constant, while the phase varies linearly with time. This is no longer the case in a wave field produced by a real source: the amplitude and phase undergo irregular fluctuations, the rapidity of which depends essentially on the effective width $\Delta\nu$ of the spectrum. The complex amplitude remains substantially constant only during a time interval Δt which is small compared to the reciprocal of the effective spectral width $\Delta\nu$; in such a time interval the change of the relative phase of any two Fourier components is much less than 2π and the addition of such components represents a disturbance which in this time interval behaves like a monochromatic wave with the mean frequency; however, this is not true for a longer time interval. The characteristic time $\Delta t = 1/\Delta\nu$ is of the order of the *coherence time* introduced in §7.5.8.

Consider next the light disturbances at two points P_1 and P_2 in a wave field produced by an extended quasi-monochromatic source. For simplicity assume that the wave field is in vacuum and that P_1 and P_2 are many wavelengths away from the source. We may expect that, when P_1 and P_2 are close enough to each other, the fluctuations of the amplitudes at these points, and also the fluctuations of the phases, will not be independent. It is reasonable to suppose that, if P_1 and P_2 are so close to each other that the difference $\Delta\mathcal{S} = SP_1 - SP_2$ between the paths from each source point S is small compared to the mean wavelength $\bar{\lambda}$, then the fluctuations at P_1 and P_2 will effectively be the same; and that some correlation between the fluctuations will exist even for greater separations of P_1 and P_2, provided that for all source points the path difference $\Delta\mathcal{S}$ does not exceed the coherence length $c\Delta t \sim c/\Delta\nu = \bar{\lambda}^2/\Delta\lambda$. We are thus led to the concept of a *region of coherence* around any point P in a wave field.

In order to describe adequately a wave field produced by a finite polychromatic source it is evidently desirable to introduce some measure for the correlation that

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

exists between the vibrations at different points P_1 and P_2 in the field. We must expect such a measure to be closely related to the sharpness of the interference fringes which would result on combining the vibrations from the two points. We must expect sharp fringes when the correlation is high (e.g. when the light at P_1 and P_2 comes from a very small source of a narrow spectral range), and no fringes at all in the absence of correlation (e.g. when P_1 and P_2 each receive light from a different physical source). We described these situations by the terms 'coherent' and 'incoherent' respectively. In general neither of these situations is realized and we may speak of vibrations which are *partially coherent*.

The first investigations which have a close bearing on the subject of partial coherence appear to be due to Verdet* who studied the size of the region of coherence for light from an extended primary source. Later researches by Michelson established the connection between the visibility of interference fringes and the intensity distribution on the surface of an extended primary source† (see §7.3.6), and between the visibility and the energy distribution in a spectral line‡ (see §7.5.8). Michelson's results were actually not interpreted in terms of correlations until much later, but his investigations have contributed to the formulation of the modern theories of partial coherence.§ The first quantitative measure of the correlation of light vibrations was introduced by von Laue,‖ and employed in his researches on the thermodynamics of light beams. Further contributions were made by Berek¶ who used the concept of correlation in investigations relating to image formation in the microscope.

A new stage in the development of the subject began with the publication of a paper by van Cittert,** who determined the joint probability distribution for the light disturbances at any two points on a screen illuminated by an extended primary source. In a later paper†† he also determined the probability distribution for the light disturbances at any one point, but at two different instants of time. Within the accuracy of his calculations the distributions were found to be Gaussian, and he determined the appropriate correlation coefficients.

A different and simpler approach to problems of partial coherence was introduced by Zernike in an important paper‡‡ published in 1938. Zernike defined the 'degree of coherence' of light vibrations in a manner that is directly related to experiment and established a number of valuable results relating to this quantity. While Zernike's degree of coherence is for most practical purposes equivalent to the correlation factor of van Cittert and is also closely related to that of von Laue, his methods seem particularly well suited for the treatment of practical problems of instrumental optics.

* E. Verdet, *Ann. Scientif. l'École Normale Supérieure*, **2** (1865), 291; *Leçons d'Optique Physique* Vol. 1 (Paris, L'Imprimerie Impériale, 1869), p. 106.

† A. A. Michelson, *Phil. Mag.* (5), **30** (1890), 1; **31** (1891), 256; *Astrophys. J.*, **51** (1920), 257.

‡ A. A. Michelson, *Phil. Mag.* (5), **31** (1891), 338; **34** (1892), 280.

§ See F. Zernike, *Proc. Phys. Soc.*, **61** (1948), 158.

‖ M. von Laue, *Ann. d. Physik* (4), **23** (1907), 1, 795.

¶ M. Berek, *Z. Phys.*, **36** (1926), 675, 824; **37** (1926), 387; **40** (1926), 420. Experiments related to Berek's investigations were described by C. Lakeman and J. Th. Groosmuller, *Physica* ('s Gravenhage), **8** (1928), 193, 199, 305.

P. H. van Cittert, *Physica*, **1 (1934), 201.

††P. H. van Cittert, *Physica*, **6** (1939), 1129; see also L. Jánossy, *Nuovo Cimento*, **6** (1957), 111; *ibid.*, **12** (1959), 369.

‡‡F. Zernike, *Physica*, **5** (1938), 785.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

The methods were simplified still further and applied to the study of image formation and resolving power by Hopkins.*

The investigations mentioned so far bridged the gap between two extreme cases, namely, that of complete coherence and complete incoherence, but the results obtained were still somewhat restricted in that they mainly applied to quasi-monochromatic light and to situations where the path differences between the interfering beams are sufficiently small. To deal with more complex situations and to formulate the theory on a rigorous basis, a further generalization was necessary. This was carried out by Wolf,† and independently by Blanc-Lapierre and Dumontet‡ and involved the introduction of more general correlation functions. These correlation functions were found to obey two wave equations; a result which implies that not only the optical disturbance but also the correlation between disturbances is propagated in the form of waves. In the light of this result many of the theorems established previously obtained a relatively simple interpretation.

The correlation functions mentioned so far characterize the correlation between the light vibrations at *two* space-time points. Correlation functions which represent correlations at two points, at any particular frequency component, were introduced much later§. Such 'second-order' correlation functions are entirely adequate for the analysis of the usual optical experiments involving interference and diffraction of light from steady sources.‖ For the analysis of more sophisticated experiments, higher-order correlation functions — i.e. correlation functions involving the field variables in total power higher than the second — may be needed.¶

* H. H. Hopkins, *Proc. Roy. Soc.*, A, **208** (1951), 263; *ibid*. A, **217** (1953), 408.

Related investigations by D. Gabor and H. Gamo utilize the concept of partial coherence in the study of optical transmission from the standpoint of information theory. (D. Gabor, *Proc. Symp. Astr. Optics*, ed. Z. Kopal (Amsterdam, North Holland Publishing Company, 1956), p. 17; *Proc. Third Symposium on Information Theory*, ed. C. Cherry (London, Butterworths Scientific Publications, 1956), p. 26; H. Gamo, *J. Appl. Phys. Japan*, **25** (1956), 431; *Progress in Optics*, Vol. 3, ed. E. Wolf [Amsterdam, North Holland Publishing Company and New York, J. Wiley and Sons, 1964), p. 187.]

† E. Wolf, *Proc. Roy. Soc.*, A, **230** (1955), 246. See also *ibid*., A, **225** (1954), 96; *Nuovo Cimento*, **12** (1954), 884.

‡ A. Blanc-Lapierre and P. Dumontet, *Rev. d'Optique*, **34** (1955), 1.

§ E. Wolf, *Opt. Commun.* **38** (1981), 3; *J. Opt. Soc. Amer.*, **72** (1982), 343; *ibid A* **3** (1986), 76; G. S. Agarwal and E. Wolf, *J. Mod. Opt.*, **40** (1993), 1489.

‖ More precisely from sources which give rise to a stationary field in the sense of the definition given on p. 561 below. For the analysis of nonstationary fields similar correlation functions may be employed, but they must be defined in terms of ensemble averages rather than time averages. For a stationary field the two types of averaging will usually give the same result.

¶ When the light originates in a thermal source, such as incandescent matter or a gas discharge, one may assume that the joint probability distribution of the field at n space-time points is, to a good approximation, Gaussian. As is well known (see, for example, J. J. Freeman, *Principles of Noise* (New York, J. Wiley & Sons, Inc., 1958), p. 245–247), such distributions are completely specified by second-order correlation functions, which in turn implies that all the higher-order correlation functions associated with thermal light may be expressed in terms of the correlation functions of the second order. This, however, is not so for light from nonthermal sources, such as a laser.

Some higher-order coherence effects and the appropriate correlation functions are discussed in L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge, Cambridge University Press, 1995).

Analogous quantum mechanical correlation functions were introduced by R. J. Glauber, in *Quantum Electronics*, 3*rd Congress*, Vol. 1, eds. N. Blombergen and P. Grivet (New York, Columbia University Press; Paris, Dunod, 1964), p. 111 and *Phys. Rev.*, **130** (1963), 2529. The relation between the classical and quantum treatments is discussed by E. C. G. Sudarshan, *Phys. Rev. Lett.*, **10** (1963), 277 and by J. R. Klauder and E. C. G. Sudarshan, *Fundamentals of Quantum Optics* (New York, W. A. Benjamin, Inc., 1968). *See also* L. Mandel and E. Wolf, *Rev. Mod. Phys.*, **37** (1965), 231; and L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge, Cambridge University Press, 1995).

Footnote continued on page 557.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

An attractive feature of the theory of partial coherence is the fact that it operates with quantities (namely with correlation functions and with time averaged intensities) that may, in principle, be determined from experiment. This is in contrast with the elementary optical wave theory, where the basic quantity is not measurable because of the very great rapidity of optical vibrations. In the present chapter we shall study the properties of partially coherent wave fields and we shall illustrate the results by a number of examples of practical interest. We shall only be concerned with the case of light, but our analysis has a close bearing on other fields; in particular, similar considerations apply in connection with correlation techniques for measurements of radio stars* and for the exploration of the ionosphere by radio waves.†

The mathematical techniques employed in connection with partial coherence are also very suitable for the analysis of partial polarization. Here one is concerned with phenomena which can be interpreted in terms of correlation between orthogonal components of the electromagnetic field vectors. Early investigations in this direction are due to G. G. Stokes.‡ Modern treatments which employ the concepts of correlation functions and correlation matrices are chiefly due to Wiener,§ Perrin,‖ Wolf¶ and Pancharatnam.** This topic will be discussed in §10.9 of this chapter.

10.2 A complex representation of real polychromatic fields

In discussing monochromatic wave fields we have found it useful to regard each real wave function as the real part of an associated complex wave function. In the present chapter we shall be concerned with polychromatic (i.e. nonmonochromatic) fields. It will again be useful to employ a complex representation, which may be regarded as a natural generalization of that used with monochromatic fields.

Let $V^{(r)}(t)$ $(-\infty < t < \infty)$ represent a real disturbance, for example, a Cartesian component of the electric vector, at a fixed point in space, and assume that $V^{(r)}(t)$ is square integrable. It may be expressed in the form of a Fourier integral

$$V^{(r)}(t) = \int_0^\infty a(\nu)\cos[\phi(\nu) - 2\pi\nu t]\mathrm{d}\nu. \tag{1}$$

With $V^{(r)}$ we associate the complex function

$$V(t) = \int_0^\infty a(\nu)\mathrm{e}^{\mathrm{i}[\phi(\nu) - 2\pi\nu t]}\mathrm{d}\nu. \tag{2}$$

Many of the basic papers on coherence properties of light are reprinted in L. Mandel and E. Wolf (eds.) *Selected Papers on Coherence and Fluctuations of Light*, (New York, Dover Publications, 1970), Vol. I (1850–1960), Vol. II (1961–1966); reprinted by SPIE (Optical Engineering Press, Bellingham, WA, 1990).

* See R. N. Bracewell, *Radio Astronomy Techniques* in *Encyclopedia of Physics*, Vol. 54, ed. S. Flügge (Berlin, Springer, 1959), Chapter V.

† Cf. J. A. Ratcliffe, *Rep. Progr. Phys.* (London, Physical Society), **19** (1956), 188.

‡ G. G. Stokes, *Trans. Cambr. Phil. Soc.*, **9** (1852), 399; reprinted in his *Mathematical and Physical Papers*, Vol. III (Cambridge, Cambridge University Press, 1901), p. 233; see also P. Soleillet, *Ann. de Physique* (10), **12** (1929), 23.

§ N. Wiener, *J. Math. and Phys.*, **7** (1928), 109; *J. Franklin Inst.*, **207** (1929), 525; *Acta Math.*, **55** (1930), §9, 182.

‖ F. Perrin, *J. Chem. Phys.*, **10** (1942), 415.

¶ E. Wolf, *Nuovo Cimento*, **12** (1954), 884; *Proc. Symp. Astr. Optics*, ed. Z. Kopal (Amsterdam, North Holland Publishing Company, 1956), p. 177; *Nuovo Cimento*, **13** (1959), 1165.

** S. Pancharatnam, *Proc. Ind. Acad. Sci.*, A, **44** (1956), 398; *ibid.*, **57** (1963), 218, 231.

Then

$$V(t) = V^{(r)}(t) + \mathrm{i}V^{(i)}(t), \tag{3}$$

where

$$V^{(i)}(t) = \int_0^\infty a(\nu)\sin[\phi(\nu) - 2\pi\nu t]\mathrm{d}\nu. \tag{4}$$

The functions $V^{(i)}(t)$ and $V(t)$ are uniquely specified by $V^{(r)}(t)$, $V^{(i)}$ being obtained from $V^{(r)}$ by replacing the phase $\phi(\nu)$ of each Fourier component by $\phi(\nu) - \pi/2$. The integrals (1) and (4) are said to be *allied Fourier integrals*, or *associated functions* (also called *conjugate functions*) and may be shown* to be Hilbert transforms of each other, i.e.

$$V^{(i)}(t) = \frac{1}{\pi}P\int_{-\infty}^{\infty}\frac{V^{(r)}(t')}{t'-t}\mathrm{d}t', \qquad V^{(r)}(t) = -\frac{1}{\pi}P\int_{-\infty}^{\infty}\frac{V^{(i)}(t')}{t'-t}\mathrm{d}t', \tag{5}$$

where P denotes the Cauchy principal value at $t' = t$.

This complex representation is used frequently in communication theory, where V is called the *analytic signal*† belonging to $V^{(r)}$. The name derives from the fact that, provided $V^{(r)}$ satisfies certain general regularity conditions, the function $V(z)$, considered as a function of a complex variable z is analytic in the lower half of the z-plane.‡

For future use we note the transition from $V^{(r)}$ to V when $V^{(r)}$ is represented as a Fourier integral of the form

$$V^{(r)}(t) = \int_{-\infty}^{\infty} \upsilon(\nu)\mathrm{e}^{-2\pi\mathrm{i}\nu t}\,\mathrm{d}\nu. \tag{6}$$

Since $V^{(r)}$ is real,

$$\upsilon(-\nu) = \upsilon^\star(\nu). \tag{7}$$

Using (7), we may re-write (6) in the form (1) and we obtain on comparison

$$\upsilon(\nu) = \tfrac{1}{2}a(\nu)\mathrm{e}^{\mathrm{i}\phi(\nu)}, \qquad \nu \geqslant 0. \tag{8}$$

In terms of υ, (2) becomes

$$V(t) = 2\int_0^\infty \upsilon(\nu)\mathrm{e}^{-2\pi\mathrm{i}\nu t}\,\mathrm{d}\nu. \tag{9}$$

Hence $V(t)$ may be derived from $V^{(r)}(t)$ by representing $V^{(r)}$ as a Fourier integral of

* See, for example, E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Oxford, Clarendon Press, 2nd edition, 1948), Chapter 5.

† The concept of an analytic signal was introduced by D. Gabor, *J. Inst. Electr. Engrs.*, **93** (1946), Pt. III, 429. See also V. I. Bunimovich, *J. Tech. Phys. USSR*, **19** (1949), 1231; J. Ville, *Câbles et Transmission*, **2** (1948), 61; *ibid.*, **4** (1950), 9; J. R. Oswald, *Trans. Inst. Radio Engrs.*, CT-3 (1956), 244.

Complex functions of a real variable whose real and imaginary parts are connected by the Hilbert transform relations play an important role in many branches of physics and engineering, in connection with causality. In physics, the Hilbert transform relations are often called *dispersion relations* as they made their first appearance in the theory of dispersion of light by atoms [H. A. Kramers, *Atti Congr. Internaz. Fisici, Como* (Sept. 1927), (V), Bologna, N. Zanichelli, 1928; see also J. S. Toll, *Phys. Rev.*, **104** (1956), 1760; H. M. Nussenzveig, *Causality and Dispersion Relations* (New York, Academic Press, 1972); *Selected Works of Emil Wolf with Commentary* (Singapore, World Scientific, 2001), Lecture 7.1, 577–584].

‡ Cf. E. C. Titchmarsh, *loc. cit.*, p. 128.

the form (6), suppressing the amplitudes belonging to the negative frequencies, and multiplying the amplitudes of the positive frequencies by two. For this reason V is also called *the complex half-range function* associated with $V^{(r)}$. Conversely it is evident that if the Fourier spectrum of a complex function V contains no amplitudes belonging to negative frequencies, then the real and imaginary parts of V are associated functions. We note the following relations which follow from (6), (7) and (9), by Parseval's theorem and by the use of the relation (3):

$$\int_{-\infty}^{\infty} V^{(r)2}(t)\mathrm{d}t = \int_{-\infty}^{\infty} V^{(i)2}(t)\mathrm{d}t = \tfrac{1}{2}\int_{-\infty}^{\infty} V(t)V^{\star}(t)\mathrm{d}t = \int_{-\infty}^{\infty} |v(\nu)|^2\,\mathrm{d}\nu = 2\int_{0}^{\infty} |v(\nu)|^2\,\mathrm{d}\nu. \tag{10}$$

In most of the applications with which we shall be concerned the spectral amplitudes will only have appreciable values in a frequency interval of width $\Delta\nu$ which is small compared to the mean frequency $\overline{\nu}$. The analytic signal then has a simple interpretation. We express V in the form

$$V(t) = A(t)\mathrm{e}^{\mathrm{i}[\Phi(t)-2\pi\overline{\nu}t]}, \tag{11}$$

where A ($\geqslant 0$) and Φ are real. According to (9) and (11),

$$\begin{aligned} A(t)\mathrm{e}^{\mathrm{i}\Phi(t)} &= 2\int_{0}^{\infty} v(\nu)\mathrm{e}^{-2\pi\mathrm{i}(\nu-\overline{\nu})t}\,\mathrm{d}\nu \\ &= \int_{-\overline{\nu}}^{\infty} g(\mu)\mathrm{e}^{-2\pi\mathrm{i}\mu t}\,\mathrm{d}\mu, \end{aligned} \tag{12}$$

where

$$g(\mu) = 2v(\overline{\nu}+\mu). \tag{13}$$

Now since the spectral amplitudes were assumed to differ appreciably from zero only in the neighbourhood of $\nu = \overline{\nu}$, $|g(\mu)|$ will be appreciable only near $\mu = 0$. Hence the integral (12) represents a superposition of harmonic components of low frequencies, and since $\Delta\nu/\overline{\nu} \ll 1$, $A(t)$ and $\Phi(t)$ will vary slowly* in comparison with $\cos 2\pi\overline{\nu}t$ and $\sin 2\pi\overline{\nu}t$. Since $V^{(r)}$ and $V^{(i)}$ are the real and imaginary parts of V, we have, in terms of A and Φ,

$$\left.\begin{aligned} V^{(r)}(t) &= A(t)\cos[\Phi(t) - 2\pi\overline{\nu}t], \\ V^{(i)}(t) &= A(t)\sin[\Phi(t) - 2\pi\overline{\nu}t]. \end{aligned}\right\} \tag{14}$$

These formulae express $V^{(r)}$ and $V^{(i)}$ in the form of modulated signals of carrier frequency $\overline{\nu}$, and we see that the complex analytic signal is intimately connected with the *envelope* of the real signal†. In terms of the analytic signal V, the envelope $A(t)$ and the associated phase factor $\Phi(t)$ are given by

* Under these circumstances one evidently has, according to (14)

$$V^{(i)}(t) \sim V^{(r)}\left(t + \frac{1}{4\overline{\nu}}\right).$$

† The envelope properties of analytic signals were studied by L. Mandel, *J. Opt. Soc. Amer.*, **57** (1967), 613.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$\left.\begin{aligned} A(t) &= \sqrt{V^{(r)2}+V^{(i)2}} = \sqrt{VV^\star} = |V|, \\ \Phi(t) &= 2\pi\bar{\nu}t + \tan^{-1}\frac{V^{(i)}}{V^{(r)}} = 2\pi\bar{\nu}t + \tan^{-1}\left(\mathrm{i}\frac{V^\star - V}{V^\star + V}\right). \end{aligned}\right\} \tag{15}$$

We see that $A(t)$ is independent of the exact choice of $\bar{\nu}$, and that $\Phi(t)$ depends on $\bar{\nu}$ only through the additive term $2\pi\bar{\nu}t$. Evidently we could have chosen in (14) any other frequency $\bar{\nu}'$ in place of $\bar{\nu}$ without affecting the value of A; the expression for the new phase factor would differ from that given by (15) only by having $\bar{\nu}'$ written in place of $\bar{\nu}$.

In deriving (14) and (15) we have not made use of the fact that we are dealing with a narrow-band signal ($\Delta\nu/\bar{\nu} \ll 1$), so that these relations hold quite generally. However, it is only when $\Delta\nu/\bar{\nu} \ll 1$ that the concept of the envelope is useful.

We have assumed that the 'disturbance' $V^{(r)}(t)$ is defined for all values of t. In practice the disturbance will exist only during a finite time interval $-T \leqslant t \leqslant T$, but this interval is as a rule so large compared to the physically significant time scales (the mean period $1/\bar{\nu}$ and the coherence time $1/\Delta\nu$) that we may idealize the situation by assuming $T \to \infty$. This idealization is mathematically desirable for reasons connected with the assumption of stationarity of the field (see §10.3.1). Evidently it is then also necessary to assume that the time average of the intensity (which is proportional to $V^{(r)2}$) tends to a finite value as the averaging interval is indefinitely increased, i.e. that

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} V^{(r)2}(t)\mathrm{d}t \tag{16}$$

is finite. Now if this limit is finite and not zero then obviously $\int_{-\infty}^{\infty} V^{(r)2}(t)\mathrm{d}t$ diverges. Nevertheless, it is possible to utilize the techniques of Fourier analysis.* We define the truncated functions

$$\left.\begin{aligned} V_T^{(r)}(t) &= V^{(r)}(t) \quad \text{when} \quad |t| \leqslant T, \\ &= 0 \qquad\quad\ \ \text{when} \quad |t| > T. \end{aligned}\right\} \tag{17}$$

Since each such truncated function may be assumed to be square integrable it may be expressed as a Fourier integral, say

$$V_T^{(r)}(t) = \int_{-\infty}^{\infty} v_T(\nu)\mathrm{e}^{-2\pi\mathrm{i}\nu t}\,\mathrm{d}\nu. \tag{18a}$$

Let ${V_T}^{(i)}$ be the associated function and V_T the corresponding analytic signal, i.e.

$$V_T(t) = V_T^{(r)}(t) + \mathrm{i}V_T^{(i)}(t) = 2\int_0^{\infty} v_T(\nu)\mathrm{e}^{-2\pi\mathrm{i}\nu t}\,\mathrm{d}\nu. \tag{18b}$$

* The problem of analysing functions of time which do not die down as t tends to infinity was encountered at the turn of the twentieth century by physicists who concerned themselves with the study of the nature of white light and noise (notably L. G. Gouy, Lord Rayleigh and A. Schuster). Rigorous mathematical techniques were developed chiefly by N. Wiener in his paper on generalized harmonic analysis (*Acta Math*., **55** (1930), 117). This paper also outlines the history of the problem and includes a very full bibliography.

Then the relations (10) hold with $V^{(r)}$ replaced by ${V_T}^{(r)}$, etc. Hence if we also divide each expression by $2T$, we obtain*

$$\frac{1}{2T}\int_{-\infty}^{\infty} V_T^{(r)2}(t)\mathrm{d}t = \frac{1}{2T}\int_{-\infty}^{\infty} V_T^{(i)2}(t)\mathrm{d}t = \frac{1}{2}\frac{1}{2T}\int_{-\infty}^{\infty} V_T(t)V_T^{\star}(t)\mathrm{d}t$$
$$= \int_{-\infty}^{\infty} S_T(\nu)\mathrm{d}\nu = 2\int_{0}^{\infty} S_T(\nu)\mathrm{d}\nu, \tag{19}$$

where

$$S_T(\nu) = \frac{|v_T(\nu)|^2}{2T}. \tag{20}$$

It would seem natural now to proceed to the limit $T \to \infty$. Unfortunately in many cases of practical interest, the function $S_T(\nu)$, known as the *periodogram*, does not tend to a limit but fluctuates† with increasing T. However, one may overcome this difficulty by an appropriate 'smoothing procedure'. For example, as customary in the theory of random processes, one regards the function $V^{(r)}(t)$ to be a typical member of an ensemble of functions which characterize the statistical properties of the process. Moreover, the ensembles that one normally encounters in optics are *stationary* and *ergodic*. Stationarity implies that all the ensemble averages are independent of the origin of time, whilst ergodicity implies that each ensemble average is equal to the corresponding time average involving a typical member of the ensemble. We will from now on assume that we are dealing with a stationary ergodic ensemble.‡ One may then show that the average of $S_T(\nu)$ taken over the ensemble of the functions $V^{(r)}(t)$ tends to a definite limit as $T \to \infty$. Thus if bar denotes the ensemble average, the 'smoothed periodogram'

$$\overline{S_T(\nu)} = \frac{\overline{|v_T(\nu)|^2}}{2T} \tag{21}$$

will possess a limit§

* Since the Hilbert transform of a truncated function is not necessarily a truncated function, ${V_T}^{(i)}$ and V_T do not, in general, vanish outside the range $-T \leqslant t \leqslant T$. For this reason, and also to avoid certain mathematical refinements, the limits of the time integrations in (19) and (23) are taken as $\pm\infty$ rather than $\pm T$.

† See, for example, W. B. Davenport and W. L. Root, *An Introduction to the Theory of Random Signals and Noise* (New York, McGraw-Hill, 1958), pp. 107–108. See also D. Middleton, *IRE Trans.*, CT-**3** (1956), 299.

‡ For a fuller discussion of these concepts, see, for example, W. B. Davenport and W. L. Root, *loc. cit.*; S. Goldman, *Information Theory* (New York, Prentice-Hall, Inc. 1953); D. Middleton, *An Introduction to Statistical Communication Theory* (New York, McGraw-Hill Co., 1960); A. M. Yaglom, *An Introduction to the Theory of Stationary Random Functions* (Englewood Cliffs, NJ, Prentice-Hall, 1962). See also E. Wolf, *J. Opt. Soc. Amer.*, **72** (1982), 343, §2.

§ A rigorous proof of the existence of this limit and of some of the relations introduced heuristically in this section would lead us far into ergodic theory and cannot, therefore, be given here. See J. L. Doob, *Stochastic Processes* (New York, J. Wiley & Sons, Inc., 1953), Chapt. XI; see also S. Goldman, *loc. cit.*, §8.4, A. M. Yaglom, *loc. cit.*, pp. 43–51 or D. Middleton, *loc. cit.*, §3.2.

Instead of taking the ensemble average, other smoothing operations may be used (see S. Goldman, *loc. cit.*, p. 244 or M. S. Bartlett, *An Introduction to Stochastic Processes* (Cambridge, Cambridge University Press, 1955), pp. 280–284).

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$S(\nu) = \lim_{T\to\infty} \overline{S_T(\nu)} = \lim_{T\to\infty} \frac{\overline{|v_T(\nu)|^2}}{2T}. \tag{22}$$

Now if angle brackets denote the time average,

$$\langle F(t)\rangle = \lim_{T\to\infty} \frac{1}{2T}\int_{-\infty}^{\infty} F_T(t)\mathrm{d}t, \tag{23}$$

then one obtains, in the limit as $T \to \infty$, the following relations analogous to (19):

$$\langle V^{(r)2}(t)\rangle = \langle V^{(i)2}(t)\rangle = \tfrac{1}{2}\langle V(t)V^{\star}(t)\rangle = \int_{-\infty}^{\infty} S(\nu)\mathrm{d}\nu = 2\int_{0}^{\infty} S(\nu)\mathrm{d}\nu. \tag{24}$$

In the theory of stationary random processes, the function $S(\nu)$, defined by (22), is called the *power spectrum* of the random process, characterized by the ensemble of the functions $V^{(r)}(t)$. In our considerations, where $V^{(r)}(t)$ represents the light disturbance, $S(\nu)\mathrm{d}\nu$ is proportional to the contribution to the intensity from the frequency range $(\nu,\ \nu + \mathrm{d}\nu)$; we shall refer to $S(\nu)$ as the *spectral density* or the *spectrum* of the light vibrations.

Since ${V_T}^{(r)}$ is the real part of V_T it follows that, when the operations on ${V_T}^{(r)}$ are linear, we may operate directly with V_T and take the real part at the end of the calculations. Moreover, just as in the monochromatic case, the relation $\langle V^{(r)2}\rangle = \frac{1}{2}\langle VV^{\star}\rangle$ allows us to calculate the time average of the square of the real disturbance directly in terms of the complex disturbance which we have associated with it.

10.3 The correlation functions of light beams

10.3.1 Interference of two partially coherent beams. The mutual coherence function and the complex degree of coherence

We have indicated in §10.1 that for a satisfactory treatment of problems involving light from a finite source and with a finite spectral range it is necessary to specify the correlation that may exist between the vibrations at two arbitrary points in the wave field. A suitable measure of this correlation is suggested by the analysis of a two-beam interference experiment.

Consider the wave field produced by an extended polychromatic source σ. For the present we neglect polarization effects, so that we may regard the light disturbance as a real scalar function $V^{(r)}(P,\ t)$ of position and time. With $V^{(r)}(P,\ t)$ we associate the analytic signal $V(P,\ t)$. By observation it is, of course, impossible to determine how these quantities vary with time, since any detector will only record averages over time intervals during which the disturbance will have changed sign very many times. The observable intensity $I(P)$ is proportional to the mean value of $V^{(r)2}(P,\ t)$, so that, apart from an inessential constant,

$$I(P) = 2\langle V^{(r)2}(P,\ t)\rangle = \langle V(P,\ t)V^{\star}(P,\ t)\rangle, \tag{1}$$

where the relation §10.2 (24) has been used.

Consider now two points P_1 and P_2 in the wave field. In addition to measuring the intensities $I(P_1)$ and $I(P_2)$ we may also determine experimentally the interference effects arising on superposition of the vibrations from these points. For this purpose imagine an opaque screen $\mathcal{A}$ to be placed across the field with pinholes at P_1 and P_2,

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

and consider the intensity distribution on a second screen $\mathcal{B}$ placed some distance from $\mathcal{A}$, on the side opposite the source (Fig. 10.1). For simplicity we assume that the medium between the two screens has refractive index unity. Let s_1 and s_2 be the distances of a typical point Q on the screen $\mathcal{B}$ from P_1 and P_2. We may regard P_1 and P_2 as centres of secondary disturbances, so that the complex disturbance at Q is given by

$$V(Q, t) = K_1 V(P_1, t - t_1) + K_2 V(P_2, t - t_2). \tag{2}$$

Here t_1 and t_2 are the times needed for light to travel from P_1 to Q and from P_2 to Q respectively, i.e.

$$t_1 = \frac{s_1}{c}, \qquad t_2 = \frac{s_2}{c}, \tag{3}$$

where c is the velocity of light in vacuum. The factors K_1 and K_2 are inversely proportional to s_1 and s_2, and depend also on the size of the openings and on the geometry of the arrangement (the angles of incidence and diffraction at P_1 and P_2). Since the secondary wavelets from P_1 and P_2 are out of phase with the primary wave by a quarter of a period (see §8.2, §8.3), K_1 and K_2 are pure imaginary numbers.

It follows from (1) and (2) that the intensity at Q is given by*

$$\begin{aligned} I(Q) = {} & K_1 K_1^\star \langle V_1(t - t_1) V_1^\star(t - t_1)\rangle + K_2 K_2^\star \langle V_2(t - t_2) V_2^\star(t - t_2)\rangle \\ & + K_1 K_2^\star \langle V_1(t - t_1) V_2^\star(t - t_2)\rangle + K_2 K_1^\star \langle V_2(t - t_2) V_1^\star(t - t_1)\rangle. \end{aligned} \tag{4}$$

Now the field was assumed to be stationary. We may shift the origin of time in all these expressions, and we have, therefore,

$$\langle V_1(t - t_1) V_1^\star(t - t_1)\rangle = \langle V_1(t) V_1^\star(t)\rangle = I_1, \tag{5}$$

and similarly for the other terms. If we also use (3), and remember that K_1 and K_2 are pure imaginary numbers, (4) may be simplified to give

$$I(Q) = |K_1|^2 I_1 + |K_2|^2 I_2 + 2|K_1 K_2| \Gamma_{12}^{(r)}\left(\frac{s_2 - s_1}{c}\right), \tag{6}$$

where $\Gamma_{12}^{(r)}(\tau)$ is the real part of the function

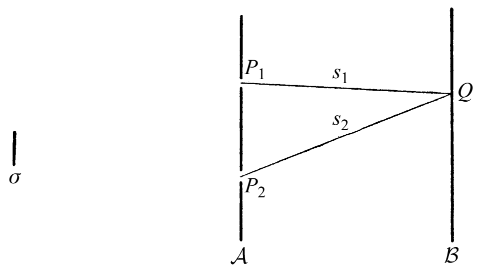


Fig. 10.1 An interference experiment with polychromatic light from an extended source σ.

* From now on, where convenient, we employ a shortened notation, writing $V_1(t)$ in place of $V(P_1, t)$, $\Gamma_{12}(\tau)$ in place of $\Gamma(P_1, P_2, \tau)$, etc.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$\Gamma_{12}(\tau) = \langle V_1(t+\tau)V_2^\star(t)\rangle. \tag{7}$$

The quantity represented by (7) is basic for the theory of partial coherence. We shall call it *the mutual coherence* of the light vibrations at P_1 and P_2, the vibrations at P_1 being considered at time τ later than at P_2; and we shall call $\Gamma_{12}(\tau)$ *the mutual coherence function** of the wave field. When the two points coincide ($P_1 = P_2$) we obtain

$$\Gamma_{11}(\tau) = \langle V_1(t+\tau)V_1^\star(t)\rangle \tag{8}$$

and we then speak of the *self-coherence* of the light vibrations at P_1; it reduces to ordinary intensity when $\tau = 0$:

$$\Gamma_{11}(0) = I_1, \qquad \Gamma_{22}(0) = I_2.$$

The term $|K_1|^2 I_1$ in (6) is evidently the intensity which would be observed at Q if the pinhole at P_1 alone were open ($K_2 = 0$) and the term $|K_2|^2 I_2$ has a similar interpretation. Let us denote these intensities by $I^{(1)}(Q)$ and $I^{(2)}(Q)$ respectively, i.e.,

$$I^{(1)}(Q) = |K_1|^2 I_1 = |K_1|^2\Gamma_{11}(0), \qquad I^{(2)}(Q) = |K_2|^2 I_2 = |K_2|^2\Gamma_{22}(0). \tag{9}$$

We also normalize $\Gamma_{12}(\tau)$:

$$\gamma_{12}(\tau) = \frac{\Gamma_{12}(\tau)}{\sqrt{\Gamma_{11}(0)}\sqrt{\Gamma_{22}(0)}} = \frac{\Gamma_{12}(\tau)}{\sqrt{I_1}\sqrt{I_2}}. \tag{10}$$

For reasons which will become apparent shortly, $\gamma_{12}(\tau)$ will be called *the complex degree of coherence* of the light vibrations. With the aid of (9) and (10), the formula (6) may finally be written in the form

$$I(Q) = I^{(1)}(Q) + I^{(2)}(Q) + 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}\gamma_{12}^{(r)}\left(\frac{s_2 - s_1}{c}\right), \tag{11}$$

where $\gamma_{12}^{(r)}$ denotes the real part of γ_{12}.

The formula (11) is the *general interference law for stationary optical fields*. It shows that, in order to determine the intensity arising from the superposition of two beams of light, we must know the intensity of each beam and the value of the real part $\gamma_{12}^{(r)}$ of the complex degree of coherence. We shall show later how $\gamma_{12}^{(r)}$ may be calculated from data that specify the source and the transmission properties of the medium.

If the light from P_1 and P_2 does not reach Q directly, but via an intervening optical system, and if dispersion effects are negligible, (11) retains its validity provided that $s_2 - s_1$ is replaced by the path difference $P_2Q - P_1Q$. With this generalization (11) also holds when the two interfering beams are derived from a primary beam, not by 'wave-front division' at P_1 and P_2, but by 'amplitude division' in the immediate neighbourhood of a single point P_1, for example in a Michelson interferometer. In this latter case (11) will involve $\gamma_{11}^{(r)}(\tau)$ in place of $\gamma_{12}^{(r)}(\tau)$.

Unlike the disturbance $V^{(r)}$, the correlation functions $\gamma_{12}^{(r)}$ and $\Gamma_{12}^{(r)}$ represent quantities which can be determined from experiment. To find the value of $\gamma_{12}^{(r)}$ for any

* In the general theory of stationary random processes $\Gamma_{12}(\tau)$ is called the *cross-correlation function* of $V_1(t)$ and $V_2(t)$ and $\Gamma_{11}(\tau)$ the *autocorrelation function* of $V_1(t)$.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

prescribed pair of points P_1 and P_2 and for any prescribed value of τ, one places an opaque screen across the light beam, with pinholes at P_1 and P_2, as in Fig. 10.1. One then measures the intensity $I(Q)$ at a point Q behind the screen, such that $P_2Q - P_1Q = c\tau$. Next one measures the intensities $I^{(1)}(Q)$ and $I^{(2)}(Q)$ of the light from each pinhole separately. In terms of these three observed values, $\gamma_{12}^{(r)}$ is, according to (11), given by

$$\gamma_{12}^{(r)} = \frac{I(Q) - I^{(1)}(Q) - I^{(2)}(Q)}{2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}}. \tag{12}$$

To determine $\Gamma_{12}^{(r)}$ one must also measure the intensities $I(P_1)$ and $I(P_2)$ at each pinhole. According to (10) and (12) $\Gamma_{12}^{(r)}$ is then given by

$$\Gamma_{12}^{(r)} = \sqrt{I(P_1)}\sqrt{I(P_2)}\gamma_{12}^{(r)} = \frac{1}{2}\sqrt{\frac{I(P_1)I(P_2)}{I^{(1)}(Q)I^{(2)}(Q)}}[I(Q) - I^{(1)}(Q) - I^{(2)}(Q)]. \tag{13}$$

Returning to (10), it is not difficult to see that our normalization ensures that $|\gamma_{12}(\tau)| \leqslant 1$. To show this we introduce, as in §10.2 (17), the truncated functions

$$\left.\begin{aligned} V_T^{(r)}(P,\, t) &= V^{(r)}(P,\, t) \quad \text{when} \quad |t| \leqslant T, \\ &= 0 \qquad\qquad\quad\ \text{when} \quad |t| > T \end{aligned}\right\} \tag{14}$$

and denote by $V_T(P,\, t)$ the associated analytic signal. By the Schwarz inequality*

$$\left|\int_{-\infty}^{\infty} V_T(P_1, t+\tau)V_T^{\star}(P_2,\, t)\mathrm{d}t\right|^2$$
$$\leqslant \int_{-\infty}^{\infty} V_T(P_1,\, t+\tau)V_T^{\star}(P_1,\, t+\tau)\mathrm{d}t\int_{-\infty}^{\infty} V_T(P_2,\, t)V_T^{\star}(P_2,\, t)\mathrm{d}t. \tag{15}$$

In the first integral on the right we may replace $t+\tau$ by t. Then dividing both sides by $4T^2$, and proceeding to the limit $T \to \infty$, it follows that

$$|\Gamma_{12}(\tau)|^2 \leqslant \Gamma_{11}(0)\Gamma_{22}(0), \tag{16}$$

or, by (10),

$$|\gamma_{12}(\tau)| \leqslant 1. \tag{17}$$

The significance of γ_{12} may best be seen by expressing (11) in a somewhat different form. Let $\bar{\nu}$ be a mean frequency of the light and write

$$\gamma_{12}(\tau) = |\gamma_{12}(\tau)|\mathrm{e}^{\mathrm{i}[\alpha_{12}(\tau) - 2\pi\bar{\nu}\tau]}, \tag{18}$$

where

$$\alpha_{12}(\tau) = 2\pi\bar{\nu}\tau + \arg\gamma_{12}(\tau). \tag{19}$$

Then (11) becomes

$$I(Q) = I^{(1)}(Q) + I^{(2)}(Q) + 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}|\gamma_{12}(\tau)|\cos[\alpha_{12}(\tau) - \delta], \tag{20}$$

* See, for example, H. Margenau and G. M. Murphy, *The Mathematics of Physics and Chemistry* (New York, D. van Nostrand Co., 1947), p. 131.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

where the parameter τ and the phase difference δ have the values

$$\tau = \frac{s_2 - s_1}{c}, \qquad \delta = 2\pi\bar{\nu}\tau = \frac{2\pi}{\bar{\lambda}}(s_2 - s_1), \tag{21}$$

and $\bar{\lambda}$ is the mean wavelength. If $|\gamma_{12}(\tau)|$ has the extreme value unity, the intensity at Q is the same as would be obtained with strictly monochromatic light of wavelength $\bar{\lambda}$, and with the phase difference between the vibrations at P_1 and P_2 equal to $\alpha_{12}(\tau)$. In this case the vibrations at P_1 and P_2 (with the appropriate time delay τ between them) may be said to be *coherent*.* If $\gamma_{12}(\tau)$ has the other extreme value, namely zero, the last term in (20) is absent; the beams do not give rise to any interference effects and the vibrations may then be said to be *incoherent*. If $|\gamma_{12}(\tau)|$ has neither of the two extreme values, i.e. if $0 < |\gamma_{12}(\tau)| < 1$, the vibrations are said to be *partially coherent*, $|\gamma_{12}(\tau)|$ representing their *degree of coherence*.†

Whatever the value of $|\gamma_{12}|$, the intensity $I(Q)$ may also be expressed in the form

$$\begin{aligned} I(Q) = {} & |\gamma_{12}(\tau)|\{I^{(1)}(Q) + I^{(2)}(Q) + 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}\cos[\alpha_{12}(\tau) - \delta]\} \\ & + [1 - |\gamma_{12}(\tau)|][I^{(1)}(Q) + I^{(2)}(Q)]. \end{aligned} \tag{22}$$

The terms in the first line may be considered to arise from *coherent* superposition of two beams of intensities $|\gamma_{12}(\tau)|I^{(1)}(Q)$ and $|\gamma_{12}(\tau)|I^{(2)}(Q)$ and of relative phase difference $\alpha_{12}(\tau) - \delta$; those in the second line from *incoherent* superposition of two beams of intensities $[1 - |\gamma_{12}(\tau)|]I^{(1)}(Q)$ and $[1 - |\gamma_{12}(\tau)|]I^{(2)}(Q)$. Thus the light which reaches Q from both pinholes may be regarded to be a mixture of coherent and incoherent light, with intensities in the ratio

$$\frac{I_{\text{coh}}}{I_{\text{incoh}}} = \frac{|\gamma_{12}(\tau)|}{1 - |\gamma_{12}(\tau)|}, \tag{23a}$$

or

$$\frac{I_{\text{coh}}}{I_{\text{tot}}} = |\gamma_{12}| \qquad (I_{\text{tot}} = I_{\text{coh}} + I_{\text{incoh}}). \tag{23b}$$

We have seen ((12)) that $\gamma_{12}{}^{(r)}$ may be determined from intensity measurements in an appropriate interference experiment. In §10.4.1 we shall see that in most cases of practical interest the modulus (and in principle also the phase) of γ_{12} can likewise be determined from such experiments.

10.3.2 Spectral representation of mutual coherence

Let

$$V_T^{(r)}(P,\ t) = \int_{-\infty}^{\infty} v_T(P,\ \nu)\mathrm{e}^{-2\pi\mathrm{i}\nu t}\,\mathrm{d}\nu \tag{24}$$

* General properties of coherent light have been investigated by L. Mandel and E. Wolf, *J. Opt. Soc. Amer.*, **51** (1961), 815.

† Various methods for measuring the degree of coherence are discussed by M. Françon and S. Mallick in *Progress in Optics*, Vol. 6, ed. E. Wolf (Amsterdam, North-Holland Publishing Company and New York, J. Wiley, 1967), p. 71.

be the Fourier integral representation of the truncated real function $V_T^{(r)}$. Then by the Fourier inversion formula

$$v_T(P, \nu) = \int_{-\infty}^{\infty} V_T^{(r)}(P, t)e^{2\pi i\nu t}\,dt, \tag{25}$$

and it follows that

$$\begin{aligned}
&\int_{-\infty}^{\infty} V_T^{(r)}(P_1, t+\tau)V_T^{(r)}(P_2, t)dt\\
&\quad= \int_{-\infty}^{\infty} V_T^{(r)}(P_2, t)\left[\int_{-\infty}^{\infty} v_T(P_1, \nu)e^{-2\pi i\nu(t+\tau)}\,d\nu\right]dt\\
&\quad= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} V_T^{(r)}(P_2, t)e^{-2\pi i\nu t}\,dt\right]v_T(P_1, \nu)e^{-2\pi i\nu\tau}\,d\nu\\
&\quad= \int_{-\infty}^{\infty} v_T(P_1, \nu)v_T^{\star}(P_2, \nu)e^{-2\pi i\nu\tau}\,d\nu.
\end{aligned}\tag{26}$$

Next we divide both sides of (26) by $2T$ and apply to the quantity $v_T(P_1, \nu)v_T^{\star}(P_2, \nu)/2T$ a 'smoothing operation' such as taking of the ensemble average (denoted by a bar) over the ensemble of the random functions $V^{(r)}$ as explained earlier in connection with §10.2 (20). Finally proceeding to the limit $T \to \infty$ one may then expect that*

$$\langle V^{(r)}(P_1, t+\tau)V^{(r)}(P_2, t)\rangle = \int_{-\infty}^{\infty} G_{12}(\nu)e^{-2\pi i\nu\tau}\,d\nu, \tag{27}$$

where

$$G_{12}(\nu) = \lim_{T\to\infty}\left[\frac{\overline{v_T(P_1, \nu)v_T^{\star}(P_2, \nu)}}{2T}\right]. \tag{28}$$

The function $G_{12}(\nu)$ may be called the *mutual spectral density*, or the *cross-spectral density* [often also denoted $W_{12}(\nu)$], of the light vibrations at P_1 and P_2. It is a generalization of the *spectral density*, $S(\nu)$, introduced earlier [§10.2 (22)] and reduces to it when the two points coincide. The mutual spectral density is the optical analogue of the concept of *cross-power spectrum* in the theory of stationary random processes. Eq. (27) shows that the real correlation function $\langle V^{(r)}(P_1, t+\tau)V^{(r)}(P, t)\rangle$ and the mutual spectral density $G_{12}(\nu)$ form a Fourier transform pair.†

Let us now pass to the complex representation. Let

$$V_T(P, t) = 2\int_0^{\infty} v_T(\nu)e^{-2\pi i\nu\tau}\,d\nu \tag{29}$$

be the analytic signal (see §10.2) associated with $V_T^{(r)}(P, t)$. It follows by an analysis similar to that which leads from (24) to (27) and again letting $T \to \infty$ that

* Similar remarks apply here as those made in footnote § on p. 561.

† N. Wiener, *Acta Math.*, **55** (1930), 117; A. Khintchine, *Math. Ann.*, **109** (1934), 604.] Essentially the same theorem was derived much earlier by A. Einstein, *Arch. Sci. Phys. Nat.*, **37** (1914), 254–256; English translation was published in *IEEE ASSP Mag.*, **4**, #4 (1987), 6, with commentary by A. M. Yaglom, *ibid.*, 7–11. There is a further discussion of this subject by Einstein in a manuscript published posthumously in *The Collected Papers of Albert Einstein*, Vol. 4 (Princeton, Princeton University Press, 1996) p. 302.

$$\Gamma_{12}(\tau) = \langle V(P_1,\ t+\tau)V^\star(P_2,\ t)\rangle = 4\int_0^\infty G_{12}(\nu)\mathrm{e}^{-2\pi\mathrm{i}\nu\tau}\,\mathrm{d}\nu. \tag{30}$$

In the case when $P_1 = P_2$ ($= P$ say), (27) and (30) imply that

$$\langle V^{(r)}(P,\ t+\tau)V^{(r)}(P,\ t)\rangle = \int_{-\infty}^{\infty} S(P,\ \nu)\mathrm{e}^{-2\pi\mathrm{i}\nu\tau}\,\mathrm{d}\nu, \tag{31a}$$

and

$$\langle V(P,\ t+\tau)V^\star(P,\ t)\rangle = 4\int_0^\infty S(P,\ \nu)\mathrm{e}^{-2\pi\mathrm{i}\nu\tau}\,\mathrm{d}\nu, \tag{31b}$$

where $S(P,\ \nu)$ is the spectral density at the point P, defined by §10.2 (22), viz.,

$$S(P,\ \nu) = \lim_{T\to\infty} \frac{\overline{|v_T(P,\ \nu)|^2}}{2T}. \tag{32}$$

Eq. (31a) shows that the real correlation function $\langle V^{(r)}(P,\ t+\tau)V^{(r)}(P,\ t)\rangle$ and the spectral density $S(P,\ \nu)$ form a Fourier transform pair. This result is the optical equivalent of the well-known *Wiener–Khintchine theorem* of the theory of stationary random processes.*

Since Γ_{12} does not contain spectral components belonging to negative frequencies, it is an analytic signal. Hence if $\Gamma_{12}^{(r)}$ and $\Gamma_{12}^{(i)}$ denote its real and imaginary parts, i.e.

$$\Gamma_{12}(\tau) = \Gamma_{12}^{(r)}(\tau) + \mathrm{i}\Gamma_{12}^{(i)}(\tau), \tag{33}$$

these functions are connected by the Hilbert transform relations

$$\Gamma_{12}^{(i)}(\tau) = \frac{1}{\pi} P\int_{-\infty}^{\infty} \frac{\Gamma_{12}^{(r)}(\tau')}{\tau'-\tau}\,\mathrm{d}\tau', \qquad \Gamma_{12}^{(r)}(\tau) = -\frac{1}{\pi} P\int_{-\infty}^{\infty} \frac{\Gamma_{12}^{(i)}(\tau')}{\tau'-\tau}\,\mathrm{d}\tau'. \tag{34}$$

It follows (see §10.2 (11)–(15)) that $|\Gamma_{12}|$, considered as a function of τ, is the envelope of $\Gamma_{12}^{(r)}$. Further, $G_{12}(-\nu) = G_{12}^\star(\nu)$ because, according to (27), $G_{12}(\nu)$ is the Fourier transform of a real function. Using this relation and (30) and (33) we find that

$$\Gamma_{12}^{(r)}(\tau) = 2\int_{-\infty}^{\infty} G_{12}(\nu)\mathrm{e}^{-2\pi\mathrm{i}\nu\tau}\,\mathrm{d}\nu = 2\langle V^{(r)}(P_1,\ t+\tau)V^{(r)}(P_2,\ t)\rangle. \tag{35}$$

Furthermore $|\gamma_{12}|$ is the envelope of the real correlation factor

$$\gamma_{12}^{(r)}(\tau) = \frac{\Gamma_{12}^{(r)}(\tau)}{\sqrt{\Gamma_{11}(0)}\sqrt{\Gamma_{22}(0)}} = \frac{\langle V^{(r)}(P_1,\ t+\tau)V^{(r)}(P_2,\ t)\rangle}{\sqrt{\langle V^{(r)2}(P_1,\ t)\rangle}\sqrt{\langle V^{(r)2}(P_2,\ t)\rangle}}. \tag{36}$$

Eq. (30) gives the spectral representation of the mutual coherence function $\Gamma_{12}(\tau)$. Eq. (35) shows that the real part of $\Gamma_{12}(\tau)$ is equal to twice the cross-correlation function of the real functions $V^{(r)}(P_1,\ t)$ and $V^{(r)}(P_2,\ t)$, and (34) shows the connection between the real and imaginary parts of $\Gamma_{12}(\tau)$.

* It is not difficult to show that $\Gamma_{12}^{(r)}(\tau)$ is also equal to $2\langle V^{(i)}(P_1,\ t+\tau)V^{(i)}(P_2,\ t)\rangle$ and $\Gamma_{12}^{(i)}(\tau) = 2\langle V^{(i)}(P_1,\ t+\tau)V^{(r)}(P_2,\ t)\rangle = -2\langle V^{(r)}(P_1,\ t+\tau)V^{(i)}(P_2,\ t)\rangle$. (See, for example, P. Roman and E. Wolf, *Nuovo Cimento*, **17** (1960), 474–476 or L. Mandel, *Progress in Optics*, Vol. 2, ed. E. Wolf (Amsterdam, North Holland Publishing Company and New York, J. Wiley and Sons, 1963), pp. 241–242.)

10.4 Interference and diffraction with quasi-monochromatic light

We have seen that, in order to describe adequately interference with partially coherent light, it is in general necessary to know the mutual coherence function $\Gamma_{12}(\tau)$ or, what amounts to the same thing, the ordinary intensities I_1 and I_2 and the complex degree of coherence $\gamma_{12}(\tau)$. We shall now restrict ourselves to the important case of quasi-monochromatic light, i.e. light consisting of spectral components that cover a frequency range $\Delta\nu$ which is small compared to the mean frequency $\bar{\nu}$. We shall see that the theory takes a simpler form in this case. In particular, we shall find that, under a certain additional assumption which is satisfied in many applications, it is possible to employ in place of $\Gamma_{12}(\tau)$ and $\gamma_{12}(\tau)$ correlation functions which are independent of the parameter τ.

10.4.1 Interference with quasi-monochromatic light. The mutual intensity

Let us again consider the interference experiment illustrated in Fig. 10.1. According to §10.3 (20), the intensity at a point Q in the interference pattern is given by

$$I(Q) = I^{(1)}(Q) + I^{(2)}(Q) + 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}|\gamma_{12}(\tau)|\cos[\alpha_{12}(\tau) - \delta], \tag{1}$$

where

$$\tau = \frac{s_2 - s_1}{c}, \qquad \delta = 2\pi\bar{\nu}\tau = \frac{2\pi}{\bar{\lambda}}(s_2 - s_1). \tag{2}$$

Suppose now that the light is quasi-monochromatic. Then it follows from §10.3 (18), in the same way as in connection with §10.2 (11), that $|\gamma_{12}(\tau)|$ and $\alpha_{12}(\tau)$, considered as functions of τ, will change slowly in comparison to $\cos 2\pi\bar{\nu}\tau$ and $\sin 2\pi\bar{\nu}\tau$. Moreover, if the openings at P_1 and P_2 are sufficiently small, the intensities $I^{(1)}(Q)$ and $I^{(2)}(Q)$ of the light diffracted from each opening separately will remain sensibly constant throughout a region of the pattern in which $\cos 2\pi\bar{\nu}\tau$ and $\sin 2\pi\bar{\nu}\tau$ change sign many times. It follows that the intensity distribution in the vicinity of any point Q consists of an almost uniform background $I^{(1)}(Q) + I^{(2)}(Q)$ on which a

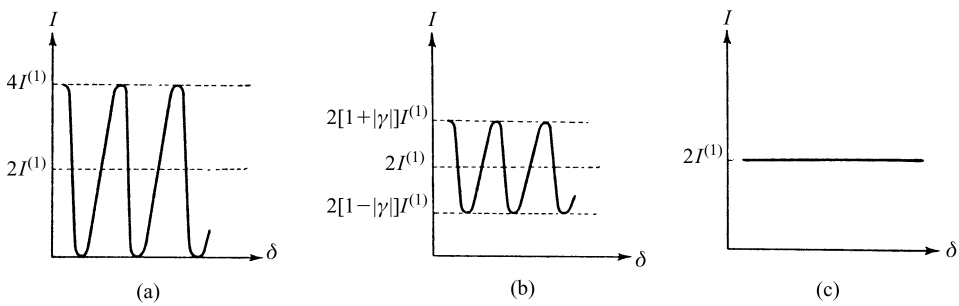


Fig. 10.2 Intensity distribution in the interference pattern produced by two quasi-monochromatic beams of equal intensity $I^{(1)}$ and with degree of coherence $|\gamma|$: (a) coherent superposition ($|\gamma| = 1$); (b) partially coherent superposition ($0 < |\gamma| < 1$); (c) incoherent superposition ($\gamma = 0$).

sinusoidal intensity distribution is superimposed, with almost constant amplitude $2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}|\gamma_{12}(\tau)|$. The behaviour of the total intensity distribution is shown for three typical cases in Fig. 10.2. The intensity maxima and minima near Q are to a good approximation given by

$$\left.\begin{aligned} I_{\max} &= I^{(1)}(Q) + I^{(2)}(Q) + 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}|\gamma_{12}(\tau)|, \\ I_{\min} &= I^{(1)}(Q) + I^{(2)}(Q) - 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}|\gamma_{12}(\tau)|. \end{aligned}\right\} \tag{3}$$

Hence the *visibility of the fringes at Q is*

$$\mathcal{V}(Q) = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \frac{2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}}{I^{(1)}(Q) + I^{(2)}(Q)}|\gamma_{12}(\tau)|. \tag{4}$$

This formula expresses the visibility of the fringes in terms of the intensity of the two beams and of their degree of coherence. If, as is often the case, the two beams are of equal intensity ($I^{(1)} = I^{(2)}$), (4) reduces to

$$\mathcal{V}(Q) = |\gamma_{12}(\tau)|, \tag{5}$$

i.e. *the visibility of the fringes is then equal to the degree of coherence*.

According to (1) and (2), the positions of the maxima of intensity near Q are given by

$$\frac{2\pi}{\bar{\lambda}}(s_2 - s_1) - \alpha_{12}(\tau) = 2m\pi \qquad (m = 0, \pm 1, \pm 2, \ldots),$$

just as if the opening were illuminated with strictly monochromatic light of wavelength $\bar{\lambda}$ and the phase at P_1 were retarded with respect to P_2 by $\alpha_{12}(\tau)$. Now according to §7.3 (7) a phase retardation of amount 2π corresponds to a displacement of the interference pattern in the direction parallel to P_1P_2 by an amount $a\bar{\lambda}/d$, where d is the distance between P_1 and P_2 and a is the distance between the screens $\mathcal{A}$ and $\mathcal{B}$. Hence the *quasi-monochromatic fringes are displaced relative to the fringes that would be formed with monochromatic and cophasal illumination of P_1 and P_2 by an amount*

$$x = \frac{\bar{\lambda}}{2\pi}\frac{a}{d}\alpha_{12}(\tau) \tag{6}$$

in the direction parallel to the line joining the openings.

We see that the amplitude and the phase of the complex degree of coherence of quasi-monochromatic light beams may be determined from measurements of the visibility and the position of interference fringes. These results have a close bearing on Michelson's method, described in §7.5.8, for determining the intensity distribution in spectral lines from measurements of visibility curves. It follows from §10.3 (10) and §10.3 (31b) that

$$\gamma_{11}(\tau) = \frac{\int_0^\infty S(\nu)\mathrm{e}^{-2\pi\mathrm{i}\nu\tau}\,\mathrm{d}\nu}{\int_0^\infty S(\nu)\mathrm{d}\nu},$$

where $S(\nu)$ is the spectral density. Hence, by the Fourier inversion theorem, $S(\nu)$ is

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

proportional to the Fourier transform of $\gamma_{11}(\tau)$. But we have just seen that the modulus of $\gamma_{11}(\tau)$ is essentially the visibility, and the phase of $\gamma_{11}(\tau)$ is simply related to the position of the fringes formed in an appropriate interference experiment. One is thus led to calculating S in exactly the same way as was done by Michelson. The visibility curves exhibited in Figs. 7.54 and 7.55 may be evidently interpreted as representing $|\gamma_{11}|$ as a function of the time delay between the two beams.

In practice the time delay τ introduced between the interfering beams is often very small, and it is then possible to simplify the formulae. According to §10.3 (30), §10.3 (18) and §10.3 (10) we have

$$|\Gamma_{12}(\tau)|\mathrm{e}^{\mathrm{i}\alpha_{12}(\tau)} = \sqrt{I_1}\sqrt{I_2}|\gamma_{12}(\tau)|\mathrm{e}^{\mathrm{i}\alpha_{12}(\tau)} = 4\int_0^\infty G_{12}(\nu)\mathrm{e}^{-2\pi\mathrm{i}(\nu-\overline{\nu})\tau}\,\mathrm{d}\nu. \tag{7}$$

If $|\tau|$ is so small that $|(\nu-\overline{\nu})\tau| \ll 1$ for all the frequencies for which $|G_{12}(\nu)|$ is appreciable, i.e. if

$$|\tau| \ll \frac{1}{\Delta\nu}, \tag{8}$$

then evidently only a small error is introduced if the exponential term of the integrand in (7) is replaced by unity. The condition (8) implies, according to §7.5 (105), that $|\tau|$ must be small compared to the coherence time of the light. When this condition is satisfied, $|\Gamma_{12}(\tau)|$, $|\gamma_{12}(\tau)|$ and $\alpha_{12}(\tau)$ differ inappreciably from $|\Gamma_{12}(0)|$, $|\gamma_{12}(0)|$ and $\alpha_{12}(0)$ respectively. It is useful to set*

$$J_{12} = \Gamma_{12}(0) = \langle V_1(t)V_2{}^\star(t)\rangle, \tag{9a}$$

$$j_{12} = \gamma_{12}(0) = \frac{\Gamma_{12}(0)}{\sqrt{\Gamma_{11}(0)}\sqrt{\Gamma_{22}(0)}} = \frac{J_{12}}{\sqrt{J_{11}}\sqrt{J_{22}}} = \frac{J_{12}}{\sqrt{I_1}\sqrt{I_2}}, \tag{9b}$$

$$\beta_{12} = \alpha_{12}(0) = \arg\gamma_{12}(0) = \arg j_{12}. \tag{9c}$$

§10.3 (18) and §10.3 (10) now give, subject to (8),

$$\gamma_{12}(\tau) \sim |j_{12}|\mathrm{e}^{\mathrm{i}(\beta_{12}-2\pi\overline{\nu}\tau)} = j_{12}\mathrm{e}^{-2\pi\mathrm{i}\overline{\nu}\tau}, \tag{10a}$$

$$\Gamma_{12}(\tau) \sim |J_{12}|\mathrm{e}^{\mathrm{i}(\beta_{12}-2\pi\overline{\nu}\tau)} = J_{12}\mathrm{e}^{-2\pi\mathrm{i}\overline{\nu}\tau}. \tag{10b}$$

Thus, *provided* (8) *is satisfied, we may replace* $\gamma_{12}(\tau)$ *and* $\Gamma_{12}(\tau)$ *in all our formulae by the quantities on the right-hand side of* (10a) *and* (10b) *respectively. In particular, the interference law* (1) *becomes*

$$I(Q) \sim I^{(1)}(Q) + I^{(2)}(Q) + 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}|j_{12}|\cos(\beta_{12}-\delta), \tag{11}$$

and is valid as long as the path difference $|s_2 - s_1| = c|\tau|$, introduced between the interfering beams, is small compared to the coherence length $c/\Delta\nu$, i.e. as long as

$$|\Delta\mathcal{S}| = |s_2 - s_1| = \frac{\overline{\lambda}}{2\pi}\delta \ll \frac{\overline{\lambda}^2}{\Delta\lambda}, \tag{12}$$

where the relation $c/\Delta\nu = \lambda^2/\Delta\lambda$ has been used.

* We again use the shortened notation where convenient, i.e. we write J_{12} in place of $J(P_1, P_2)$, etc.

Eq. (11) is the basic formula of an elementary (quasi-monochromatic) theory of partial coherence, which forms the subject matter of the rest of this section; some applications of this theory will be considered in §10.6. Within its range of validity [indicated by (8) or (12)] the correlation between the vibrations at any two points P_1 and P_2 in the wave field is characterized by J_{12} rather than by $\Gamma_{12}(\tau)$, i.e. by a quantity which depends on the positions of the two points, but not on the time difference τ. It follows from (10a) that, within the accuracy of this elementary theory,

$$|\gamma_{12}(\tau)| \sim |j_{12}|, \tag{13}$$

so that $|j_{12}|$ $(0 \leqslant |j_{12}| \leqslant 1)$ represents the degree of coherence of the vibrations at P_1 and P_2; and we see from (11) that the phase β_{12} of j_{12} represents their effective phase difference. j_{12}, just like $\gamma_{12}(\tau)$ of which it is a special case, is usually called *the complex degree of coherence* (sometimes the *complex coherence factor* or the *equal-time degree of coherence*); and J_{12} is called the *mutual intensity* or the *equal-time coherence function*.

10.4.2 Calculation of mutual intensity and degree of coherence for light from an extended incoherent quasi-monochromatic source

(a) The van Cittert–Zernike theorem

We shall now determine the mutual intensity J_{12} and the complex degree of coherence j_{12} for points P_1 and P_2 on a screen $\mathcal{A}$ illuminated by an extended quasi-monochromatic incoherent source σ. For simplicity σ will be taken to be a portion of a plane parallel to $\mathcal{A}$, and we will assume that the medium between the source and the screen is homogeneous. We also assume that the linear dimensions of σ are small compared to the distance OO' between the source and the screen (Fig. 10.3), and that the angles between OO' and the line joining a typical source point S to P_1 and P_2 are small.

Imagine the source to be divided into elements $\mathrm{d}\sigma_1$, $\mathrm{d}\sigma_2$, ... centred on points S_1, S_2, ... of linear dimensions small compared to the mean wavelength $\bar{\lambda}$. If $V_{m1}(t)$ and $V_{m2}(t)$ are the complex disturbances at P_1 and P_2 due to the element $\mathrm{d}\sigma_m$, the total disturbances at these points are

$$V_1(t) = \sum_m V_{m1}(t), \qquad V_2(t) = \sum_m V_{m2}(t). \tag{14}$$

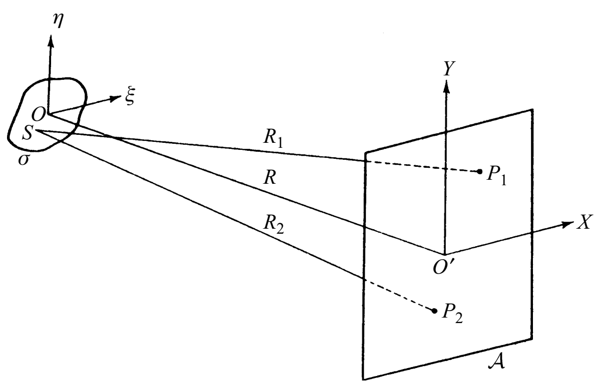


Fig. 10.3 Illustrating the van Cittert–Zernike theorem.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

Hence

$$J(P_1, P_2) = \langle V_1(t)V_2^\star(t)\rangle = \sum_m \langle V_{m1}(t)V_{m2}^\star(t)\rangle + \sum_{m\neq n}\sum \langle V_{m1}(t)V_{n2}^\star(t)\rangle. \tag{15}$$

Now the light vibrations arising from different elements of the source may be assumed to be statistically independent (mutually incoherent), and of zero mean value, so that*

$$\langle V_{m1}(t)V_{n2}^\star(t)\rangle = \langle V_{m1}(t)\rangle\langle V_{n2}^\star(t)\rangle = 0 \quad \text{when} \quad m \neq n. \tag{16}$$

If R_{m1} and R_{m2} are the distances of P_1 and P_2 from the source element $\mathrm{d}\sigma_m$, then

$$V_{m1}(t) = A_m\left(t - \frac{R_{m1}}{v}\right)\frac{\mathrm{e}^{-2\pi\mathrm{i}\bar{\nu}(t-R_{m1}/v)}}{R_{m1}}, \qquad V_{m2}(t) = A_m\left(t - \frac{R_{m2}}{v}\right)\frac{\mathrm{e}^{-2\pi\mathrm{i}\bar{\nu}(t-R_{m2}/v)}}{R_{m2}}, \tag{17}$$

where $|A_m|$ characterizes the strength and $\arg A_m$ the phase of the radiation from the mth element,† and v is the velocity of light in the medium between the source and the screen. Hence

$$\begin{aligned}\langle V_{m1}(t)V_{m2}^\star(t)\rangle &= \left\langle A_m\left(t - \frac{R_{m1}}{v}\right)A_{m1}^\star\left(t - \frac{R_{m2}}{v}\right)\right\rangle \frac{\mathrm{e}^{2\pi\mathrm{i}\bar{\nu}(R_{m1}-R_{m2})/v}}{R_{m1}R_{m2}} \\ &= \left\langle A_m(t)A_m^\star\left(t - \frac{R_{m2}-R_{m1}}{v}\right)\right\rangle \frac{\mathrm{e}^{2\pi\mathrm{i}\bar{\nu}(R_{m1}-R_{m2})/v}}{R_{m1}R_{m2}}.\end{aligned} \tag{18}$$

If the path difference $R_{m2} - R_{m1}$ is small compared to the coherence length of the light we may neglect the retardation $(R_{m2} - R_{m1})/v$ in the argument of $A_m^\star$, and we obtain from (15), (16) and (18)

$$J(P_1, P_2) = \sum_m \langle A_m(t)A_m^\star(t)\rangle \frac{\mathrm{e}^{2\pi\mathrm{i}\bar{\nu}(R_{m1}-R_{m2})/v}}{R_{m1}R_{m2}}. \tag{19}$$

The quantity $\langle A_m(t)A_m^\star(t)\rangle$ characterizes the intensity of the radiation from the source element $\mathrm{d}\sigma_m$. In any practical case the total number of the source elements may be assumed to be so large that we may regard the source to be effectively continuous. Denoting by $I(S)$ the intensity per unit area of the source, i.e. $I(S_m)d\sigma_m = \langle A_m(t)A_m^\star(t)\rangle$, (19) becomes‡

$$J(P_1, P_2) = \int_\sigma I(S)\frac{\mathrm{e}^{\mathrm{i}\bar{k}(R_1-R_2)}}{R_1R_2}\,\mathrm{d}S, \tag{20}$$

where R_1 and R_2 denote the distances between a typical source point S and the points P_1 and P_2, and $\bar{k} = 2\pi\bar{\nu}/v = 2\pi/\bar{\lambda}$ is the wave number in the medium. The complex degree of coherence $j(P_1, P_2)$ is, according to (20) and (9b), given by

* Incoherence always implies a finite (though not necessarily wide) spectral range, and (16) is, in fact, not valid for the idealized case of strictly monochromatic light. For monochromatic light one has $V_{m1}(t) = U_{m1}\mathrm{e}^{-2\pi\mathrm{i}\nu t}$, $V_{n2}(t) = U_{n2}\mathrm{e}^{-2\pi\mathrm{i}\nu t}$, where U_{m1} and U_{n2} are independent of time, so that $\langle V_{m1}(t)V_{n2}^\star(t)\rangle = U_{m1}U_{n2}^\star$ and this quantity is in general different from zero.

† In general A_m also depends on direction, but for simplicity we neglect this dependence.

‡ From now on we shall frequently use the notation $\mathrm{d}S$, $\mathrm{d}P_1$, ... for surface elements centred on the points S, P_1, ...

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$j(P_1, P_2) = \frac{1}{\sqrt{I(P_1)}\sqrt{I(P_2)}} \int_\sigma I(S) \frac{e^{i\bar{k}(R_1 - R_2)}}{R_1 R_2} \, dS, \tag{21}$$

where

$$I(P_1) = J(P_1, P_1) = \int_\sigma \frac{I(S)}{R_1^2} \, dS, \qquad I(P_2) = J(P_2, P_2) = \int_\sigma \frac{I(S)}{R_2^2} \, dS, \tag{21a}$$

are the average intensities at P_1 and P_2.

We note that the integral (21) is the same as that which occurs in quite a different connection; namely in the calculation, on the basis of the Huygens–Fresnel principle, of the complex disturbance in the diffraction pattern arising from diffraction of a spherical wave on an aperture in an opaque screen. More precisely, (21) implies that *the equal-time complex degree of coherence $j(P_1, P_2)$, which describes the correlation of vibrations at a fixed point P_2 and a variable point P_1 in a plane illuminated by an extended quasi-monochromatic incoherent source, is equal to the normalized complex amplitude at the corresponding point P_1 in a certain diffraction pattern, centred on P_2. This pattern would be obtained on replacing the source by a diffracting aperture of the same size and shape as the source, and on filling it with a spherical wave converging to P_2, the amplitude distribution over the wave-front in the aperture being proportional to the intensity distribution across the source.* This result was first established by van Cittert* and later more generally by Zernike.† We shall refer to it as the *van Cittert–Zernike theorem*.

In most applications the intensity $I(S)$ may be assumed to be independent of the position of S on the surface (uniform intensity). The corresponding diffraction problem is then that of diffraction of a spherical wave of uniform amplitude by an aperture of the same size and shape as the source.

Let (ξ, η) be the coordinates of a typical source point S, referred to axes at O, and let (X_1, Y_1) and (X_2, Y_2) be the coordinates of P_1 and P_2 referred to parallel axes at O' (Fig. 10.3). Then, if R denotes the distance OO'

$$R_1^2 = (X_1 - \xi)^2 + (Y_1 - \eta)^2 + R^2,$$

so that

$$R_1 \sim R + \frac{(X_1 - \xi)^2 + (Y_1 - \eta)^2}{2R}. \tag{22}$$

Here only the leading terms in X_1/R, Y_1/R, ξ/R and η/R have been retained. A strictly similar expression is obtained for R_2, so that

$$R_1 - R_2 \sim \frac{(X_1^2 + Y_1^2) - (X_2^2 + Y_2^2)}{2R} - \frac{(X_1 - X_2)\xi + (Y_1 - Y_2)\eta}{R}. \tag{23}$$

In the denominator of the integrands in (20) and (21), R_1 and R_2 may to a good approximation be replaced by R. We also set

* P. H. van Cittert, *Physica*, **1** (1934), 201. † F. Zernike, *Physica*, **5** (1938), 785.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$\frac{(X_1 - X_2)}{R} = p, \qquad \frac{(Y_1 - Y_2)}{R} = q, \tag{24}$$

$$\psi = \frac{\bar{k}[(X_1^2 + Y_1^2) - (X_2^2 + Y_2^2)]}{2R}. \tag{25}$$

Then (21) reduces to

$$j_{12} = \frac{\mathrm{e}^{\mathrm{i}\psi} \iint\limits_{\sigma} I(\xi, \eta)\mathrm{e}^{-\mathrm{i}\bar{k}(p\xi + q\eta)}\,\mathrm{d}\xi\,\mathrm{d}\eta}{\iint\limits_{\sigma} I(\xi, \eta)\mathrm{d}\xi\,\mathrm{d}\eta}. \tag{26}$$

Hence *if the linear dimensions of the source and the distance between* P_1 *and* P_2 *are small compared to the distance of these points from the source, the degree of coherence* $|j_{12}|$ *is equal to the absolute value of the normalized Fourier transform of the intensity function of the source.*

The quantity ψ defined by (25) has a simple interpretation. According to (23) it represents the phase difference $2\pi(OP_1 - OP_2)/\bar{\lambda}$, and may evidently be neglected when

$$OP_1 - OP_2 \ll \bar{\lambda}. \tag{27}$$

For a uniform circular source of radius ρ with its centre at O, (26) gives on integration (see §8.5.2)

$$j_{12} = \left(\frac{2J_1(v)}{v}\right)\mathrm{e}^{\mathrm{i}\psi}, \tag{28}$$

where

$$\left.\begin{aligned} v &= \bar{k}\rho\sqrt{p^2 + q^2} = \frac{2\pi}{\bar{\lambda}}\frac{\rho}{R}\sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}, \\ \psi &= \frac{2\pi}{\bar{\lambda}}\left[\frac{(X_1^2 + Y_1^2) - (X_2^2 + Y_2^2)}{2R}\right], \end{aligned}\right\} \tag{29}$$

J_1 being the Bessel function of the first kind and first order.* According to §8.5.2, $|2J_1(v)/v|$ decreases steadily from the value unity when $v = 0$ to the value zero when $v = 3.83$; thus as the points P_1 and P_2 are separated more and more, the degree of coherence steadily decreases and there is complete incoherence when P_1 and P_2 are separated by the distance

$$P_1P_2 = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} = \frac{0.61R\bar{\lambda}}{\rho}. \tag{30}$$

A further increase in v re-introduces a small amount of coherence, but the degree of coherence remains smaller than 0.14, and there is further complete incoherence for $v = 7.02$. Since $J_1(v)$ changes sign as v passes through each zero of $J_1(v)$, the phase

* No confusion should arise from the fact that the symbol J is also used for the mutual intensity, as the latter always appears with two suffixes or with several arguments.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$\beta_{12} = \arg j_{12}$ changes there by π; in consequence the position of the bright and dark fringes are interchanged after each disappearance of the fringes.

The function $2J_1(v)/v$ decreases steadily from the value 1 for $v = 0$ to 0.88 when $v = 1$, i.e. when

$$P_1 P_2 = \frac{0.16 R\bar{\lambda}}{\rho}. \tag{31}$$

Regarding a departure of 12 per cent from the ideal value unity as the maximum permissible departure, it follows that *the diameter of the circular area that is illuminated almost coherently by a quasi-monochromatic, uniform, incoherent source of angular radius* $\alpha = \rho/R$ is* $0.16\lambda/\alpha$. This result is useful in estimating the size of a source needed in experiments on interference and diffraction.

As an example consider the size of the 'area of coherence' around an arbitrary point on a screen illuminated directly by the sun. The angular diameter 2α which the sun's disc subtends on the surface of the earth is about $0° \, 32' \sim 0.0093$ radians. Hence, if the variation of brightness across the sun's disc is neglected, the diameter d of the area of coherence is approximately $0.16\bar{\lambda}/0.0047 \sim 34\bar{\lambda}$. Taking the mean wavelength $\bar{\lambda}$ as 5.5×10^{-5} cm this gives $d \sim 0.019$ mm.

In the present context Michelson's method of measuring angular diameters of stars (see §7.3.6) appears in a new light. According to (5) and (13), the visibility of the fringes is equal to the degree of coherence of the light vibrations at the two outer mirrors (M_1 and M_2 in Fig. 7.16) of the Michelson stellar interferometer. For a uniformly bright circular star disc of angular radius α the smallest separation of the mirrors for which the degree of coherence has zero value (first fringe disappearance) is, according to (30), equal to $0.61\bar{\lambda}/\alpha$ in agreement with §7.3 (43). Moreover, from the measurements of both the visibility and the position of the fringes, it is in principle possible to determine not only the stellar diameter, but also the distribution of the intensity over the stellar disc. For, according to §10.4.1, measurements of the visibility and the position of the fringes are equivalent to determining both the amplitude and the phase of the complex degree of coherence j_{12}, and, according to (26), the intensity distribution is proportional to the inverse Fourier transform of j_{12}.

We mentioned in §7.3.6 the important modification due to Hanbury Brown and Twiss of the Michelson's stellar interferometer known as the *intensity interferometer*.† In the Hanbury Brown–Twiss system light from the star is focused on two photoelectric detectors P_1, P_2 and information about the star is obtained from the study of the correlation in the fluctuations of their current outputs. A full analysis of the performance of this system must take into account the quantum nature of the photoelectric effect,‡ and requires also some knowledge of electronics, and is thus outside the scope of this book. The principle of the method may, however, be easily

* As early as 1865, E. Verdet estimated that the diameter of the 'circle of coherence' is somewhat smaller than $0.5R\bar{\lambda}/\rho$. (*Ann. Scientif. de l'École Normale Supérieure*, **2** (1865), 291; also his *Leçons d'Optique Physique*, Tome 1 (Paris, L'Imprimerie Impériale, 1869), p. 106.)

† The theory of this system and an account of its history is given in R. Hanbury Brown, *The Intensity Interferometer* (London, Taylor and Francis, 1974).

‡ See R. Hanbury Brown and R. Q. Twiss, *Proc. Roy. Soc.*, A, **242** (1957), 300; *ibid.*, A, **243** (1957), 291. See also E. M. Purcell, *Nature*, **178** (1956), 1449; L. Mandel, *Proc. Phys. Soc.*, **72** (1958), 1037; *Progress in Optics*, Vol. 2, ed. E. Wolf (Amsterdam, North Holland Publishing Company and New York, J. Wiley and Sons, 1963), p. 181.

understood. Under ideal experimental conditions (absence of noise), the current output of each photoelectric detector is proportional to the instantaneous intensity $I(t)$ of the incident light, and the fluctuation in the current output is proportional to $\Delta I(t) = I(t) - \langle I(t) \rangle$. Hence, in the interferometer of Hanbury Brown and Twiss, the quantity which is effectively being measured is proportional to $\Omega_{12} = \langle \Delta I_1 \Delta I_2 \rangle$. Detailed calculations show* that Ω_{12} is proportional to the square of the modulous of the degree of coherence, so that the knowledge of Ω_{12}, just like the knowledge of $|j_{12}|$, yields information about the size of the star.

(b) Hopkins' formula

In deriving the van Cittert–Zernike formula (21), it was assumed that the medium between the source σ and the points P_1 and P_2 is homogeneous. It is not difficult to generalize the formula to other cases, e.g. when the medium is heterogeneous or consists of a succession of homogeneous regions of different refractive indices.

We again imagine the source to be divided into small elements $d\sigma_1, d\sigma_2, \ldots,$ centred on points $S_1, S_2, \ldots,$ of linear dimensions small compared to the mean wavelength $\overline{\lambda}$. If, as before, $V_{m1}(t)$ and $V_{m2}(t)$ represent the disturbances at P_1 and P_2 due to the element $d\sigma_m$, (15) and (16) still hold, but in (17) we must replace each factor $e^{i\overline{k}R_{mj}}/R_{mj}$ $(j = 1, 2)$; $\overline{k} = 2\pi\overline{\nu}/v$ by a more general function. We introduce a *transmission function* $K(S, P, \nu)$ of the medium, defined in a similar way as in §9.5.1; it represents the complex disturbance at P, due to a monochromatic point source of frequency ν, of unit strength and of zero phase, situated at the element $d\sigma$ at S. For a homogeneous medium we have, from the Huygens–Fresnel principle, that $K(S, P, \nu) = -ie^{ikR}/\lambda R$, where R denotes the distance SP, it being assumed that the angle which SP makes with the normal to $d\sigma$ is sufficiently small. It follows that in the more general case the factor $e^{i\overline{k}R_{mj}/R_{mj}}$ must be replaced by $i\overline{\lambda}K(S_m, P, \overline{\nu})$, and we obtain, on passing to a continuous distribution, the following relation in place of (20):

$$J(P_1, P_2) = \overline{\lambda}^2 \int_\sigma I(S) K(S, P_1, \overline{\nu}) K^\star(S, P_2, \overline{\nu}) dS. \tag{32}$$

According to (32) and (9b),

$$j(P_1, P_2) = \frac{\overline{\lambda}^2}{\sqrt{I(P_1)}\sqrt{I(P_2)}} \int_\sigma I(S) K(S, P_1, \overline{\nu}) K^\star(S, P_2, \overline{\nu}) dS, \tag{33}$$

where $I(P_1) = J(P_1, P_1)$ and $I(P_2) = J(P_2, P_2)$ are the intensities at P_1 and P_2 respectively.

For the purpose of later applications it will be useful to express (32) and (33) in a slightly different form. We set

$$i\overline{\lambda}K(S, P_1, \overline{\nu})\sqrt{I(S)} = U(S, P_1), \qquad i\overline{\lambda}K(S, P_2, \overline{\nu})\sqrt{I(S)} = U(S, P_2). \tag{34}$$

Formulae (32) and (33) become

* L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge, Cambridge University Press, 1995), §9.9, §9.10 and §14.6.1. For elementary, nonrigorous treatment see E. Wolf, *Phil. Mag.*, **2** (1957), 351. See also J. A. Ratcliffe, *Rep. Progr. Phys.* (London, Physical Society), **19** (1956), 233.

$$J(P_1, P_2) = \int_\sigma U(S, P_1)U^\star(S, P_2)\mathrm{d}S, \tag{35a}$$

$$j(P_1, P_2) = \frac{1}{\sqrt{I(P_1)}\sqrt{I(P_2)}}\int_\sigma U(S, P_1)U^\star(S, P_2)\mathrm{d}S. \tag{35b}$$

We note that $U(S, P)$, defined by (34), is proportional to the disturbance which would arise at P from a strictly monochromatic point source of frequency $\bar{\nu}$, strength $\sqrt{I(S)}$ and zero phase, situated at S. Thus (35) may be interpreted as expressing the mutual intensity $J(P_1, P_2)$ and the complex degree of coherence $j(P_1, P_2)$ due to an extended quasi-monochromatic source, in terms of the *disturbances* produced at P_1 and P_2 by each source point of an 'associated' monochromatic source.*

Expression (35b) was first suggested by Hopkins† from heuristic considerations and is very useful in solving coherence problems of instrumental optics. The main usefulness of the formula arises from the fact that, like the van Cittert–Zernike theorem, it permits the calculation of the complex degree of coherence of light from an incoherent source without the explicit use of an averaging process.

10.4.3 An example

We shall illustrate the preceding considerations by the discussion of an experiment. A primary source σ_0 is imaged by a lens L_0 onto a pinhole σ_1 and the light that emerges from the pinhole is rendered parallel by a lens L_1. A second lens L_2, exactly similar to L_1, brings the interfering beams to a focus F in the focal plane $\mathcal{F}$ of the lens L_2. A plane mirror M is used to reduce the overall length of the instrument (Fig. 10.4). If a diffracting mask (a dark screen $\mathcal{A}$, for example a piece of uniformly blackened film), with apertures of any desired size, shape and distribution is placed in the parallel beam between L_1 and L_2, its Fraunhofer diffraction pattern is formed in the focal plane $\mathcal{F}$,

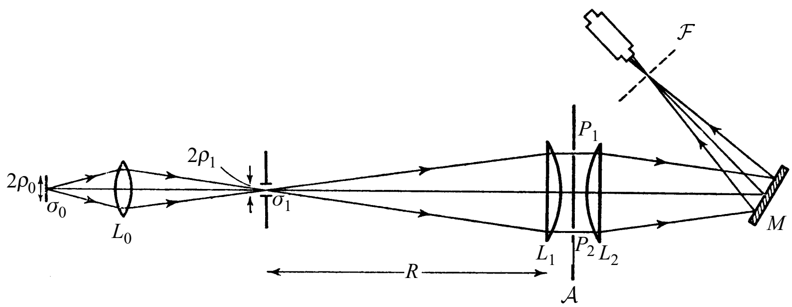


Fig. 10.4 The diffractometer.

* It must not be assumed that the *mutual intensity* and the *complex degree of coherence* of light from such a fictitious source are also given by (35). For, as already explained, the relation (16) used in the derivation of these formulae is not valid in the limiting case of monochromatic radiation; the degree of coherence of monochromatic light is, in fact, always equal to unity.

† H. H. Hopkins, *Proc. Roy. Soc.*, A, **208** (1951), 263.

and, under normal use, the operator may displace the mask while viewing the diffraction pattern through a microscope.*

If σ_0 were a quasi-monochromatic point source, it would give rise to coherent illumination in the neighbourhood of its geometrical image in the plane of σ_1. The size of this coherently illuminated area is of the order of the effective size of the Airy diffraction pattern σ_A, formed by the lens L_0, of the single source point. The light distribution in the plane of σ_1 due to a finite incoherent source may be regarded as arising from the incoherent superposition of many such patterns. If, as we assume, the image formed by L_0 of this extended source, and the pinhole σ_1, are both large compared to σ_A, the illuminated pinhole σ_1 will itself effectively act as an *incoherent source*.† According to the van Cittert–Zernike theorem, such a source will give rise to a correlation between vibrations at any two points on the first surface of the lens L_1 (and more generally in the plane $\mathcal{A}$); and with the usual approximations the complex degree of the coherence is given by the formula (28):

$$j_{12} = |j_{12}|\mathrm{e}^{\mathrm{i}\beta_{12}} = \frac{2J_1(v)}{v}\mathrm{e}^{\mathrm{i}\psi}, \tag{36a}$$

where

$$v = \frac{2\pi}{\bar{\lambda}}\frac{\rho_1 d}{R}, \qquad \psi = \frac{2\pi}{\bar{\lambda}}\left(\frac{r_1^2 - r_2^2}{2R}\right), \tag{36b}$$

where $d = P_1P_2$, ρ_1 is the radius of σ_1, R is the distance between σ_1 and L_1, and r_1 and r_2 are the distances of P_1 and P_2 from the axis.

If the diffracting mask $\mathcal{A}$ consists of two small circular apertures centred on P_1 and P_2, the pattern observed in the focal plane $\mathcal{F}$ results from the superposition of two partially coherent beams, with a degree of coherence $|j_{12}|$, emerging from these apertures. We shall investigate the changes in the structure of this pattern as the separation of P_1 and P_2 is gradually increased, i.e. as the degree of coherence between the two interfering beams is varied.

We assume that P_1 and P_2 are situated symmetrically about the axis. Then $\psi = 0$ and the intensities $I^{(1)}(Q)$ and $I^{(2)}(Q)$ at a point Q in the focal plane associated with either of the two beams are then equal, and are given by the Fraunhofer formula for diffraction by a circular aperture [§8.5 (14)]. If the point Q is the focus for rays diffracted in directions that make an angle ϕ with the normal to $\mathcal{A}$, and if a is the radius of each aperture (see Fig. 10.5) then, apart from a normalizing factor,

$$I^{(1)}(Q) = I^{(2)}(Q) = \left[\frac{2J_1(u)}{u}\right]^2, \qquad u = \frac{2\pi}{\bar{\lambda}}a\sin\phi. \tag{37}$$

The phase difference δ between the beams diffracted to Q is

$$\delta = \frac{2\pi}{\bar{\lambda}}P_2N = \frac{2\pi}{\bar{\lambda}}d\sin\phi = Cuv, \qquad C = \frac{\bar{\lambda}}{2\pi}\frac{R}{\rho_1 a}, \tag{38}$$

* This apparatus, known as the *diffractometer*, is mainly used in connection with optical diffraction methods for solutions of problems of X-ray structure analysis. (See C. A. Taylor, R. M. Hinde and H. Lipson, *Acta Cryst.*, **4** (1951), 261; A. W. Hanson, H. Lipson and C. A. Taylor, *Proc. Roy. Soc.* A, **218** (1953), 371; W. Hughes and C. A. Taylor, *J. Sci. Instr.*, **30** (1953), 105.)

† This point is discussed quantitatively in §10.6.1. See also A. T. Forrester, *Amer. J. Phys.*, **24** (1956), 194.

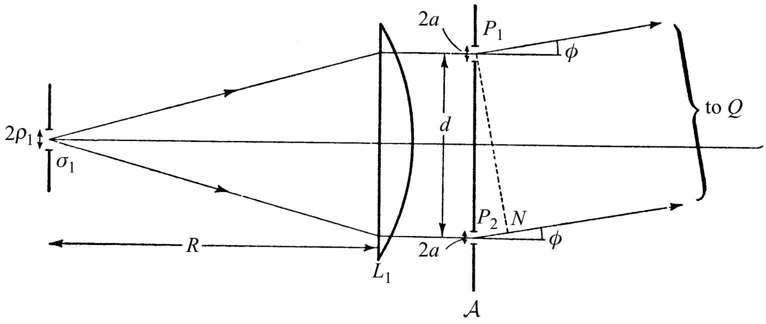


Fig. 10.5 Calculation of the intensity distribution in the focal plane of the diffractometer.

where N is the foot of the perpendicular dropped from P_1 on to the ray diffracted at P_2. On substituting from (36), (37) and (38) into (11) we finally obtain the following expression for the intensity at the point $Q(\phi)$ in the focal plane, when the apertures at P_1 and P_2 are separated by a distance d:

$$I(\phi,\, d) = 2\left[\frac{2J_1(u)}{u}\right]^2 \left\{1 + \left|\frac{2J_1(v)}{v}\right| \cos[\beta_{12}(v) - Cuv]\right\}, \tag{39}$$

where

$$\left.\begin{aligned} \beta_{12}(v) &= 0 \quad \text{when} \quad \frac{2J_1(v)}{v} > 0, \\ &= \pi \quad \text{when} \quad \frac{2J_1(v)}{v} < 0. \end{aligned}\right\} \tag{40}$$

In Fig. 10.6 are shown photographs of the patterns observed with such an arrangement, for various separations d. The corresponding theoretical curves, computed from the formula (39) are also shown. The chain lines represent the envelopes

$$\left.\begin{aligned} I_{\text{max}}(\phi,\, d) &= 2\left[\frac{2J_1(u)}{u}\right]^2 \left\{1 + \left|\frac{2J_1(v)}{v}\right|\right\}, \\ I_{\text{min}}(\phi,\, d) &= 2\left[\frac{2J_1(u)}{u}\right]^2 \left\{1 - \left|\frac{2J_1(v)}{v}\right|\right\}. \end{aligned}\right\} \tag{41}$$

It is of interest to note that when $\beta = \pi$ [cases (D) and (E)], the intensity at the centre of each pattern has a relative minimum, not a maximum, in agreement with our general considerations. The variation of the degree of coherence with the separation of the two apertures, based on experiments of Thompson and Wolf, is shown in Fig. 10.7, where also the six values corresponding to the photographs of Fig. 10.6 are indicated by the appropriate letters. A more recent two-beam interference experiment of this kind was performed using a high precision digital automated system.*

* D. Ambrosini, G. Schirripa Spagnolo, D. Paoletti and S. Vicalvi, *Pure Appl. Opt.* **7** (1998), 933.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

10.4.4 Propagation of mutual intensity

Consider a beam of quasi-monochromatic light from an extended incoherent source σ, and suppose that the mutual intensity is known for all pairs of points on a fictitious surface $\mathcal{A}$ intercepting the beam. We shall show that it is then possible to determine the mutual intensity for all pairs of points on any other second surface $\mathcal{B}$ illuminated by the light from $\mathcal{A}$ either directly or via an optical system.

We assume to begin with that the medium between $\mathcal{A}$ and $\mathcal{B}$ is homogeneous and of refractive index unity. Let $U(S, Q_1)$ and $U(S, Q_2)$ be the disturbances at points Q_1 and Q_2 on $\mathcal{B}$ (Fig. 10.8) due to a typical source point S of the associated monochromatic source. Then, according to (35), the mutual intensity $J(Q_1, Q_2)$ is given by

$$J(Q_1, Q_2) = \int_\sigma U(S, Q_1)U^\star(S, Q_2)\mathrm{d}S. \tag{42}$$

Now $U(S, Q_1)$ and $U(S, Q_2)$ may be expressed in terms of the disturbance at all points of $\mathcal{A}$ by means of the Huygens–Fresnel principle:

$$U(S, Q_1) = \int_{\mathcal{A}} U(S, P_1)\frac{\mathrm{e}^{\mathrm{i}\overline{k}s_1}}{s_1}\Lambda_1\,\mathrm{d}P_1. \tag{43}$$

Here s_1 is the distance from a typical point P_1 on $\mathcal{A}$ to Q_1, Λ_1 is the inclination factor (denoted by K in Chapter VIII) at P_1, and $\overline{k} = 2\pi\overline{\nu}/c$ is the mean wave number. For small obliquities, $\Lambda_1 \sim -\mathrm{i}/\lambda$. According to (43) and a similar expression for $U(Q_2)$, we have

$$U(S, Q_1)U^\star(S, Q_2) = \int_{\mathcal{A}}\int_{\mathcal{A}} U(S, P_1)U^\star(S, P_2)\frac{\mathrm{e}^{\mathrm{i}\overline{k}(s_1-s_2)}}{s_1 s_2}\Lambda_1\Lambda_2^\star\,\mathrm{d}P_1\,\mathrm{d}P_2, \tag{44}$$

where the points P_1 and P_2 take on independently all positions on the surface $\mathcal{A}$ of integration. Next we substitute from (44) into (42), and change the order of integration. The integration over σ gives precisely $J(P_1, P_2)$ and we obtain

$$J(Q_1, Q_2) = \int_{\mathcal{A}}\int_{\mathcal{A}} J(P_1, P_2)\frac{\mathrm{e}^{\mathrm{i}\overline{k}(s_1-s_2)}}{s_1 s_2}\Lambda_1\Lambda_2^\star\,\mathrm{d}P_1\,\mathrm{d}P_2. \tag{45}$$

This is the required formula, due to Zernike,* for the propagation of the mutual intensity. We have implicitly assumed in the derivation of (45) that light from every point of the surface $\mathcal{A}$ reaches the points Q_1 and Q_2. The presence of any diaphragm between the two surfaces may be taken care of by limiting the integration to those portions of the surface $\mathcal{A}$ which send light to Q_1 and Q_2, unless the diaphragm is so small that the effects of diffraction at its edges cannot be neglected. The diffraction can be taken into account by carrying out the transition from $\mathcal{A}$ to $\mathcal{B}$ in two steps, first from $\mathcal{A}$ to the plane of the diaphragm, and then from the plane of the diaphragm to the surface $\mathcal{B}$.

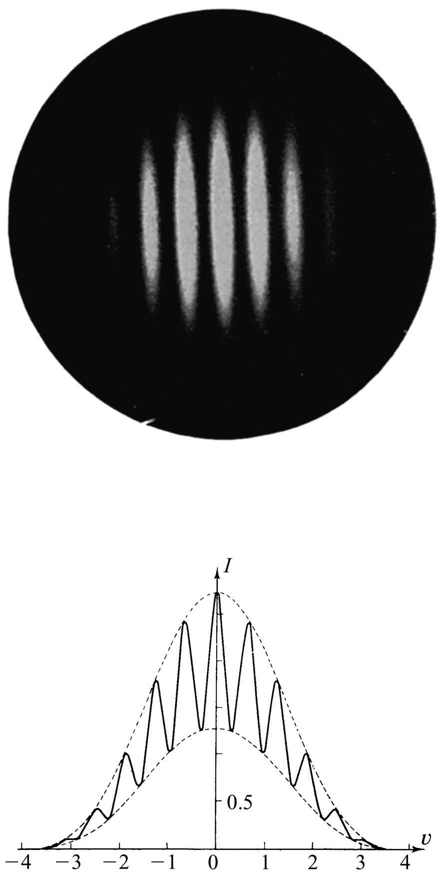
In the special case when the points Q_1 and Q_2 coincide, (45) reduces to the following expression for the intensity, when we also substitute for $J(P_1, P_2)$ in terms of the intensities $I(P_1)$, $I(P_2)$ and the complex degree of coherence $j(P_1, P_2)$:

$$I(Q) = \int_{\mathcal{A}}\int_{\mathcal{A}} \sqrt{I(P_1)}\sqrt{I(P_2)}j(P_1, P_2)\frac{\mathrm{e}^{\mathrm{i}\overline{k}(s_1-s_2)}}{s_1 s_2}\Lambda_1\Lambda_2^\star\,\mathrm{d}P_1\,\mathrm{d}P_2. \tag{46}$$

* F. Zernike, *Physica*, **5** (1938), 791.

(a)

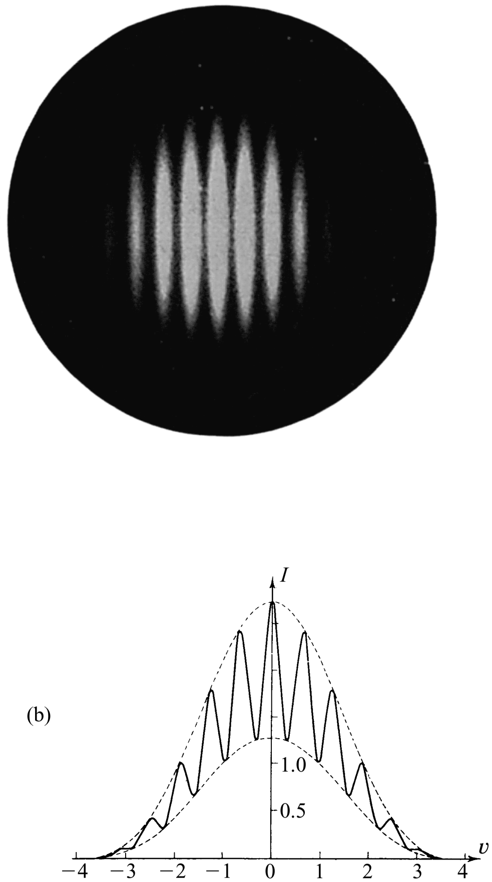
(b)



(A) $d = 0.6$ cm
$|j_{12}| = 0.593, \beta_{12} = 0$

(a)

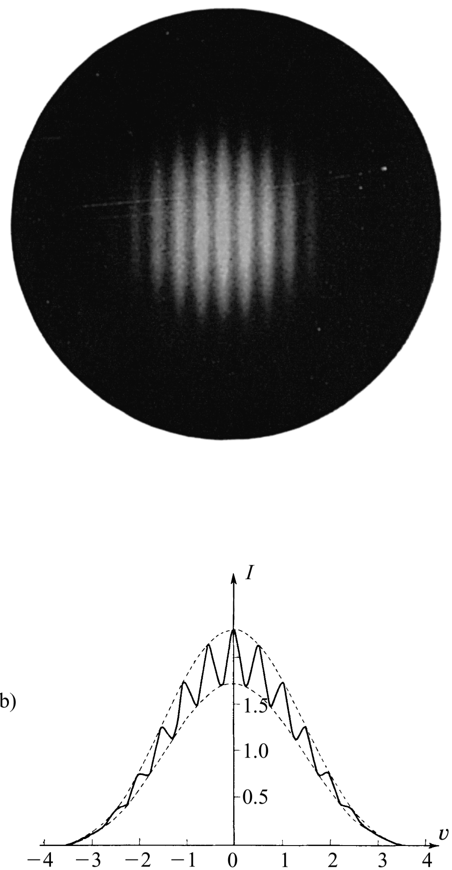
(b)



(B) $d = 0.8$ cm
$|j_{12}| = 0.361, \beta_{12} = 0$

(a)

(b)



(C) $d = 1$ cm
$|j_{12}| = 0.146, \beta_{12} = 0$

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

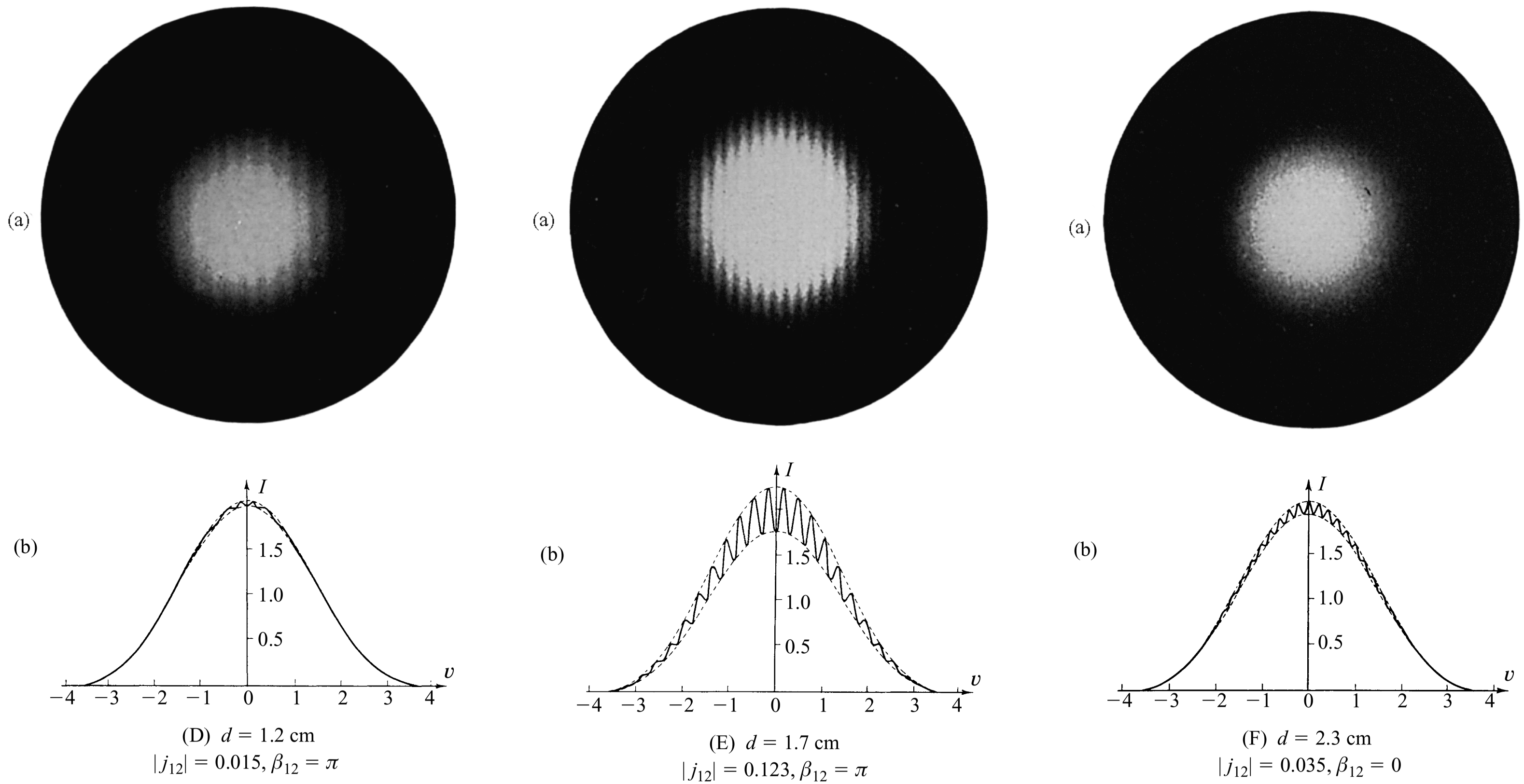


Fig. 10.6 Two-beam interference with partially coherent light. (a) Observed patterns, (b) theoretical intensity curves. Focal length of lenses L_0, L_1, and L_2 of diffractometer: $f_0 = 20$ cm, $f_1 = f_2 = R = 152$ cm. Diameter of $L_0 = 5$ cm. Distance from L_0 to σ_1: 40 cm. Separation of L_1 and L_2: 14 cm. Distance of mirror M from $L_2 = 85$ cm. Diameter $2\rho_1$ of pinhole σ_1: 0.9×10^{-2} cm. Diameter $2a$ of apertures at P_1 and P_2: 0.14 cm. Mean wavelength $\bar{\lambda} = 5790$ Å. (After B. J. Thompson and E. Wolf, *J. Opt. Soc. Amer.*, **47** (1957), 895.)

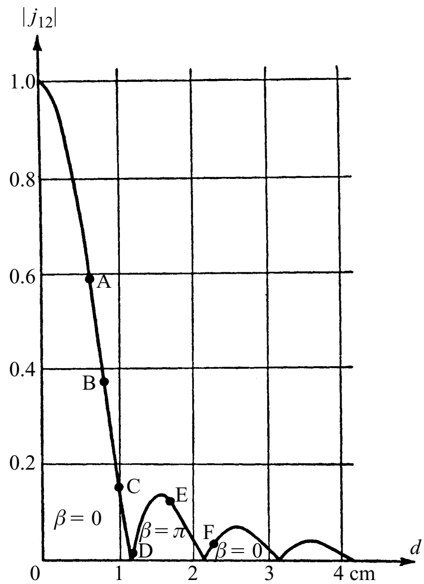


Fig. 10.7 Two-beam interference with partially coherent light. The degree of coherence as a function of the separation d of the two illuminated apertures in the diffractometer ($\rho_1 = 0.45 \times 10^{-2}$ cm, $R = 152$ cm, $\bar{\lambda} = 5790$ Å; incoherent illumination of σ_1 assumed).

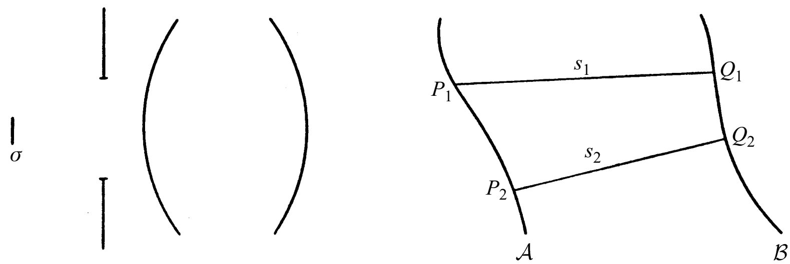


Fig. 10.8 Propagation of mutual intensity: illustrating formula (45).

This formula expresses the intensity at a point Q as the sum of contributions from each pair of elements dP_1, dP_2 of an arbitrary surface $\mathcal{A}$ intercepting the beam (see Fig. 10.9). The contribution from each pair of elements depends on the intensities at P_1 and P_2 and each contribution is weighted by the appropriate value of the complex degree of coherence factor $j(P_1, P_2)$. Formula (46) may be regarded as a kind of Huygens–Fresnel principle for the propagation of intensity in a partially coherent field. The resemblance between the formulae just derived and those of the more elementary Huygens–Fresnel theory has a deeper significance which will be brought out in our rigorous formulation of the theory of partial coherence (§10.8).

If the light from $\mathcal{A}$ reaches $\mathcal{B}$ via an optical system, factor $\Lambda e^{i\bar{k}s}/s$ must evidently be replaced by an appropriate transmission function $K(P, Q)$. Instead of (45) we then obtain the more general formula

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

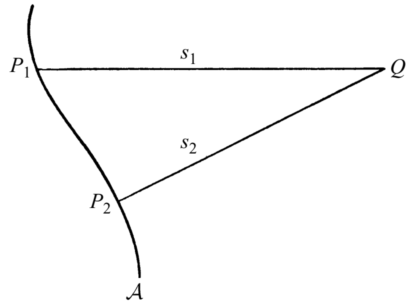


Fig. 10.9 Illustrating formula (46).

$$J(Q_1, Q_2) = \int_{\mathcal{A}}\int_{\mathcal{A}} J(P_1, P_2)K(P_1, Q_1)K^{\star}(P_2, Q_2)\mathrm{d}P_1\,\mathrm{d}P_2. \tag{47}$$

10.5 Interference with broad-band light and the spectral degree of coherence. Correlation-induced spectral changes

We next consider Young's interference experiment with two beams which are partially coherent with respect to each other. However, unlike in §10.3.1 we will now examine the effect of superposition of the beams not on the intensity but rather on the spectrum of the light in the region of superposition.

The experimental set up is the same as indicated in Fig. 10.1. With the same notation as used previously, the complex disturbance at a point Q in the detection plane $\mathcal{B}$ is given by §10.3 (2), viz.,

$$V(Q, t) = K_1 V(P_1, t - t_1) + K_2 V(P_2, t - t_2), \tag{1}$$

where

$$t_1 = \frac{s_1}{c}, \qquad t_2 = \frac{s_2}{c}, \tag{2}$$

again denote the times needed for light to travel to Q from the openings at P_1 and P_2 respectively. As we noted in connection with §10.3 (2) the factors K_1 and K_2 are pure imaginary numbers. They depend on the size of the openings at P_1 and P_2, on the angles of incidence and diffraction and are inversely proportional to the distances s_1 and s_2 and to the wavelength of the light transmitted through the pinholes. However, under usual experimental conditions K_1 and K_2 may, to a good approximation, be taken to be constants and approximately equal to each other. We will make this approximation and denote the constant by K.

Let

$$\Gamma(Q, Q, \tau) = \langle V(Q, t+\tau)V^{*}(Q, t)\rangle \tag{3}$$

be the self-coherence function of the light disturbance at Q. On substituting from (1) into (3) we obtain for $\Gamma(Q, Q, \tau)$ the expression

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$\Gamma(Q,\,Q,\,\tau)=|K|^2\{\langle V_1(t-t_1+\tau)V_1^{\star}(t-t_1)\rangle+\langle V_2(t-t_2+\tau)V_2^{\star}(t-t_2)\rangle$$
$$+\langle V_1(t-t_1+\tau)V_2^{\star}(t-t_2)\rangle+\langle V_2(t-t_2+\tau)V_1^{\star}(t-t_1)\rangle\},\qquad(4)$$

where $V_1(t-t_1)\equiv V(P,\,t-t_1)$ etc. Because the field was assumed to be stationary, we may shift the origin of time in all the averages and (4) then gives the following expression for $\Gamma(Q,\,Q,\,\tau)$:

$$\Gamma(Q,\,Q,\,\tau)=|K|^2[\Gamma_{11}(\tau)+\Gamma_{22}(\tau)+\Gamma_{12}(t_2-t_1+\tau)+\Gamma_{21}(t_1-t_2+\tau)].\qquad(5)$$

Here $\Gamma_{11}(\tau)$ and $\Gamma_{22}(\tau)$ are the self-coherence functions of the light vibrations at the points P_1 and P_2 respectively [§10.3 (8)] and $\Gamma_{12}(\tau)$ and $\Gamma_{21}(\tau)$ are the mutual coherence functions which characterize the correlation between the light vibrations at these two points [§10.3 (7)]. Let us take the Fourier transform of (5) and use the fact that the Fourier transforms of $\Gamma_{11}(\tau)$ and $\Gamma_{22}(\tau)$ are proportional to the spectral densities $S(P_1,\,\nu)$ and $S(P_2,\,\nu)$ respectively [§10.3 (31b)] and that the Fourier transforms of $\Gamma_{12}(\tau)$ and $\Gamma_{21}(\tau)$ are proportional to the cross-spectral densities $G(P_1,\,P_2,\,\nu)$ and $G(P_2,\,P_1,\,\nu)$ respectively [§10.3 (30)]. We then obtain the following expression for the spectral distribution of the light reaching the plane $\mathcal{B}$ of observation:

$$S(Q,\,\nu)=|K|^2[S(P_1,\,\nu)+S(P_2,\,\nu)+G(P_2,\,P_1,\,\nu)\mathrm{e}^{\mathrm{i}\delta}+G(P_1,\,P_2,\,\nu)\mathrm{e}^{-\mathrm{i}\delta}],\quad(\nu\geqslant 0),\qquad(6)$$

where

$$\delta=2\pi\nu(t_2-t_1)=\frac{2\pi}{\lambda}(s_2-s_1).\qquad(7)$$

It follows from the definition §10.3 (28) of the cross-spectral density that $G(P_2,\,P_1,\,\nu)$ and $G(P_1,\,P_2,\,\nu)$ are complex conjugates of each other and hence (6) may be expressed in the form

$$S(Q,\,\nu)=|K|^2\{S(P_1,\,\nu)+S(P_2,\,\nu)+2\mathcal{R}[G(P_1,\,P_2,\,\nu)\mathrm{e}^{-\mathrm{i}\delta}]\},\qquad(8)$$

where $\mathcal{R}$ denotes the real part. This formula may be re-written in a physically more significant form by expressing the last term on the right-hand side of (8) in a somewhat different form. For this purpose we introduce the normalized cross-spectral density

$$\mu_{12}(\nu)=\frac{G(P_1,\,P_2,\,\nu)}{\sqrt{S(P_1,\,\nu)}\sqrt{S(P_2,\,\nu)}},\qquad(9)$$

known as the *spectral degree of coherence*. It has been shown that it is a measure of correlations of the spectral components of frequency ν of the light vibrations at the points* P_1 and P_2. We show in Appendix VIII that its upper bound is unity, i.e., that for all values of its argument

$$0\leqslant|\mu_{12}(\nu)|\leqslant 1.\qquad(10)$$

The extreme value unity represents complete correlation (complete spectral coherence), while the extreme value zero represents complete absence of correlation (complete spectral incoherence).

* E. Wolf, *J. Opt. Soc. Amer.*, **A3** (1986), 76, Sec. 3. See also L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge, Cambridge University Press, 1995), Sec. 4.7.2.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

One must distinguish between the (temporal) complex degree of coherence $\gamma_{12}(\tau)$ and the spectral degree of coherence $\mu_{12}(\nu)$; $\gamma_{12}(\tau)$ is a measure of coherence in the space-time domain, whilst $\mu_{12}(\nu)$ is a measure of coherence in the space-frequency domain.*

On using the definition (9), formula (8) for the spectral density at Q may be expressed in the form

$$S(Q, \nu) = |K|^2\{S(P_1, \nu) + S(P_2, \nu) + 2\sqrt{S(P_1, \nu)}\sqrt{S(P_2, \nu)}|\mu_{12}(\nu)|\cos[\beta_{12}(\nu) - \delta]\}, \tag{11}$$

where

$$\beta_{12}(\nu) = \arg \mu_{12}(\nu). \tag{12}$$

The term $|K|^2 S(P_1, \nu)$ evidently represents the spectral density which would be observed at Q if the small aperture at P_1 alone was open and the term $|K|^2 S(P_2, \nu)$ has a similar interpretation. Denoting these spectral densities by $S^{(1)}(Q, \nu)$ and $S^{(2)}(Q, \nu)$ respectively, i.e.

$$S^{(1)}(Q, \nu) = |K|^2 S(P_1, \nu), \qquad S^{(2)}(Q, \nu) = |K|^2 S(P_2, \nu), \tag{13}$$

Eq. (11) becomes

$$S(Q, \nu) = S^{(1)}(Q, \nu) + S^{(2)}(Q, \nu) + 2\sqrt{S^{(1)}(Q, \nu)}\sqrt{S^{(2)}(Q, \nu)}|\mu_{12}(\nu)|\cos[\beta_{12}(\nu) - \delta]. \tag{14}$$

Formula (14) is called the *spectral interference law for superposition of beams of any state of coherence*. It is to be noted that it is of the same mathematical form as the interference law §10.4 (1) for superposition of beams of any state of coherence, but the physical significance of the two laws is quite different. §10.4 (1) describes the variation of the average intensity observed in the region of superposition, whilst (14) describes how superposition of the two beams affects the spectrum.

In most cases of interest one may assume that $S^{(1)}(Q, \nu) \approx S^{(2)}(Q, \nu)$. The spectral interference law then becomes

$$S(Q, \nu) = 2S^{(1)}(Q, \nu)\{1 + |\mu_{12}(\nu)|\cos[\beta_{12}(\nu) - \delta]\}. \tag{15}$$

This formula implies two results. First, that at any fixed frequency ν, the spectral density varies sinusoidally with the position of the point Q across the observation plane $\mathcal{B}$, with the amplitude and the phase depending on the value of the (generally complex) spectral degree of coherence $\mu_{12}(\nu)$. Secondly, at any fixed point Q in the observation plane $\mathcal{B}$, the spectral density $S(Q, \nu)$ will, in general, differ from the spectral density $S^{(1)}(Q, \nu)$ of the light which would reach Q through only one of the openings, the difference depending on the spectral degree of coherence.

In a sense the 'intensity' interference law §10.4 (1) and the spectral interference law (14) above are complementary to each other. The former shows that appreciable

* Simple relationships between $\gamma_{12}(\tau)$ and $\mu_{12}(\nu)$ exist under special circumstances. For example, when the spectral range of the light is sufficiently narrow or is made sufficiently narrow by filtering and is centred at frequency $\overline{\nu}$ one can show that $\mu_{12}(\overline{\nu}) \approx \gamma_{12}(0)$ [E. Wolf, *Opt. Lett.*, **8** (1983), 250]. The relationship between the two degrees of coherence is discussed, under more general circumstances, by A. T. Friberg and E. Wolf, *Opt. Lett.*, **20** (1995), 623.

modifications of the averaged intensity take place when *narrow-band* quasi-monochromatic light beams are superposed. The latter indicates that appreciable changes of spectra may take place when two *broad-band* beams are superposed. More detailed analysis reveals that in the former case no appreciable spectral changes take place, whilst in the latter case no appreciable intensity variations are produced. Moreover, no interference fringes are formed when the path difference between the two beams exceeds the coherence length whereas spectral modulation takes place irrespective of the path difference $s_2 - s_1$ $(= \lambda\delta/2\pi)$, as is evident from (14).*

If sufficient information is available about the coherence properties of the light incident on the pinholes, the spectral degree of coherence may be calculated. In other cases the spectral degree of coherence may be determined from the spectral interference law (14) if the spectra $S(Q, \nu)$ and $S^{(1)}(Q, \nu)$ are measured.†

In Fig. 10.10 spectra are shown, calculated from the spectral interference law (4), at different points Q in the plane of observation, for various separations d of the pinholes. The source of light illuminating the pinholes was assumed to be a uniform circular blackbody source at temperature 3000 K. It is seen that the spectrum changes drastically on superposition of the two beams and that with increasing distance x from the axis, it exhibits more and more rapid oscillations. Experimental demonstration of such spectral changes has been reported in several papers.‡ Results of one such set of experiments are shown in Fig. 10.11.

Because the spectral modulation which we have just discussed is a consequence of the wave nature of light and its coherence properties, one might expect that similar spectral modifications may arise with other kinds of waves. They have indeed been demonstrated in acoustics§ and in neutron interferometry.‖

The spectral changes produced by superposition of two broad-band light beams which we discussed are the simplest example of the effect of spatial coherence on spectra of optical fields. It has been shown more generally that the spectrum of light generated by partially coherent sources may change on propagation, even in free space,¶ sometimes quite appreciably. The changes may be manifested as shifts of spectral lines either towards the longer or the shorter wavelengths (see, for example, Fig. 10.12), they may result in broadening or narrowing of lines, or in the suppression of some of the spectral lines and in generation of new ones. Such correlation-induced spectral changes have been extensively studied, both theoretically and experimentally** and are discussed in some detail elsewhere.††

* Interference in complementary spaces has been discussed by G. S. Agarwal, *Found. Phys.*, **25** (1995), 219.

† Detailed description of the procedure is described in D. F. V. James and E. Wolf, *Opt. Commun.*, **145** (1998), 1.

‡ M. Santarsiero and F. Gori, *Phys. Lett. A*, **167** (1992), 123; H. C. Kandpal, J. S. Vaishya, M. Chander, K. Saxena, D. S. Mehta and K. Joshi, *Phys. Lett. A* **167** (1992), 114.

§ M. F. Bocko, D. H. Douglass and R. S. Knox, *Phys. Rev. Lett.*, **58** (1987), 2649.

‖ See, for example, H. Rauch, *Phys. Lett. A*, **173** (1993), 240 and D. L. Jacobson, S. A. Werner and H. Rauch, *Phys. Rev. A* **49** (1994), 3196.

¶ E. Wolf, *Phys. Rev. Lett.* **56** (1986), 1370.

** For a review of this subject see E. Wolf and D. F. V. James, *Rep. Progr. Phys.*, **59** (1996), 771.

†† L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge, Cambridge University Press, 1995), Sec. 5.8.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

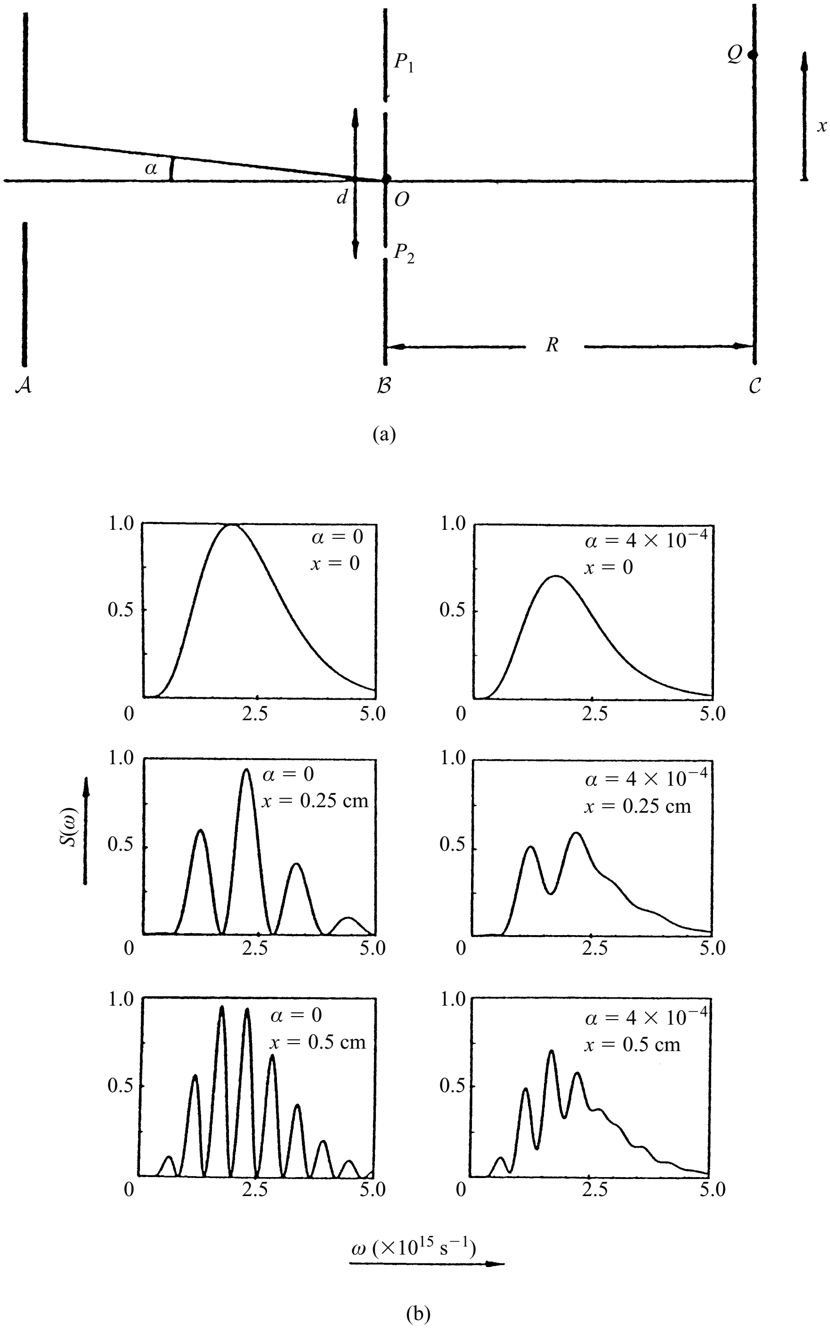


Fig. 10.10 (a) Notation relating to Young's interference experiment with broad-band light, originating in a uniform circular source whose diameter subtends an angle 2α at the axial point O in plane $\mathcal{B}$ of the pinholes. The curves in (b) below were calculated for the case when $d = 0.1$ cm and $R = 150$ cm. (b) The spectra at different points Q, located at distance x from the axis. The spectrum of the incident light was assumed to be a black-body spectrum at temperature 3000 K. (After D. F. V. James and E. Wolf, *Phys. Lett. A*, **157** (1991), 6.)

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

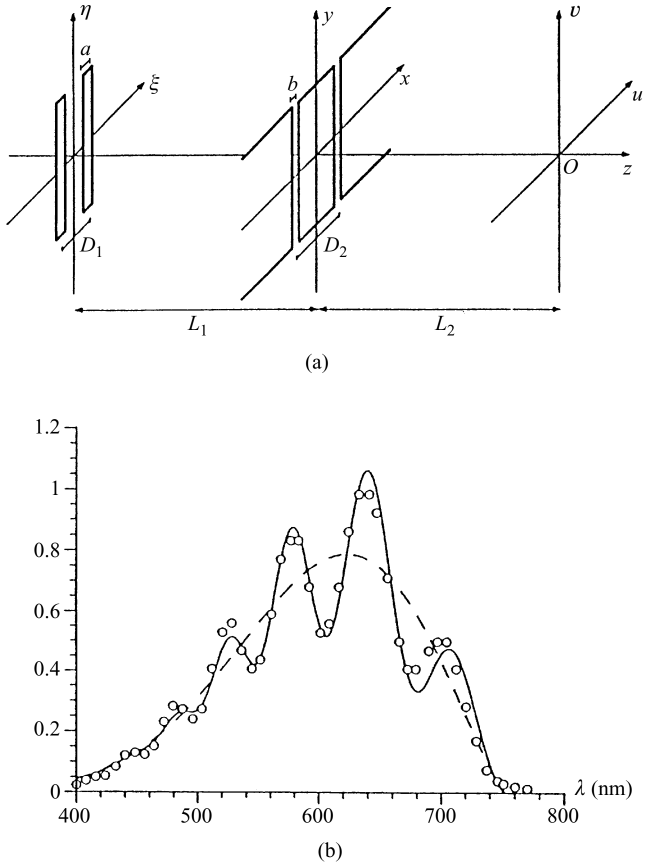


Fig. 10.11 Correlation-induced spectral changes in a Young two-slit interference experiment with broad-band light. (a) Notation and layout: The source consists of two mutually incoherent radiating strips parallel to the η direction in the ξ,η-plane. The x,y-plane contains two slits along the y direction in the x,y-plane. (b) The dashed curve, taken to be the best-fitting curve from measured values, represents the spectrum $S(0, \nu)$ at the origin O in the u,v-plane formed by light which reaches O through one of the slits only. The full line represents the calculated spectrum $S(0, \nu)$ at O, formed by light which reaches O via both slits. The circles show measured values. The calculations and measurements apply to the situation when $D_1 = 0.68$ mm, $D_2 = 3.4$ mm, $a = 0.026$ mm and $b = 0.11$ mm. The spectra are plotted as function of the wavelength rather than frequency. (After M. Santarsiero and F. Gori, *Phys. Lett. A*, **167** (1992), 123.)

10.6 Some applications

10.6.1 The degree of coherence in the image of an extended incoherent quasi-monochromatic source

As a preliminary to the study of image formation with partially coherent light, it will be useful to consider the degree of coherence in the image of an extended incoherent source, formed by a centred optical system. A finite degree of correlation among

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

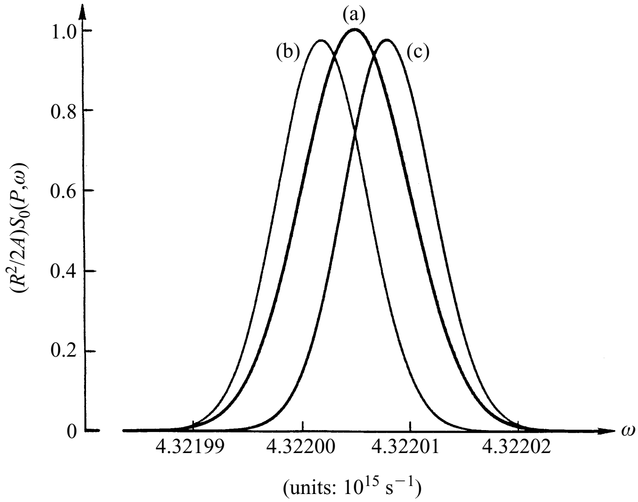


Fig. 10.12 Correlation-induced shifts of spectral lines produced on superposition of light originating in two small sources at a point P at equal distance R from the sources. Apart from a proportionality constant the source spectrum $S_0(\omega) = \exp[-(\omega - \omega_0)^2/2\delta_0{}^2]$, $(\omega = 2\pi\nu)$, $\omega_0 = 4.32201 \times 10^{15}\ \mathrm{s}^{-1}$ (mercury line $\lambda = 4538.33$ Å), $\delta_0 = 5 \times 10^9\ \mathrm{s}^{-1}$. (a) The spectrum at P when the sources are uncorrelated ($\mu_{12} \equiv 0$). (b), (c) The spectra at P when the two sources are correlated, with the spectral degree of coherence $\mu_{12} = a \exp[-(\omega - \omega_0)^2/2\delta_1{}^2] - 1$, $a = 1.8$, $\delta_1 = 7.5 \times 10^9\ \mathrm{s}^{-1}$ and (b) $\omega_1 = \omega_0 - 2\delta_0$ ('red-shifted' line), (c) $\omega_1 = \omega_0 + 2\delta_0$ ('blue-shifted' line). After E. Wolf, *Phys. Rev. Lett.*, **58** (1987), 2646.

vibrations in the image plane arises from the fact that, because of diffraction (and in general also because of aberrations), the light from each source point is not concentrated into a point, but spreads over a finite area. Some of these 'image patterns' overlap and in consequence points in the image plane that are close enough to each other receive coherent as well as incoherent contributions.

Suppose that σ is a uniform quasi-monochromatic incoherent circular source of radius ρ emitting light with mean (vacuum) wavelength $\bar{\lambda}_0$, situated in a homogeneous object space of refractive index n. Further let D be the distance between the object plane and the plane of the entrance pupil. Corresponding quantities in the image will be denoted by primed symbols of the same type.

Let d be the distance between two points P_1 and P_2 in the entrance pupil. We assume that $\rho/D \ll 1$, $d/D \ll 1$ and $OP_1 - OP_2 \ll \bar{\lambda}_0$, where O is the axial point of the source*; then the complex degree of coherence $j(P_1, P_2)$ is, according to §10.4 (28), given by

$$j(P_1, P_2) = \frac{2J_1(v)}{v}, \tag{1}$$

* If this last condition is not satisfied, $|j(P_1, P_2)|$ remains unchanged, but according to §10.4 (28) the phase of $j(P_1, P_2)$ is increased by an amount $\psi = 2\pi[OP_1 - OP_2]/\bar{\lambda}_0$.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$v = \frac{2\pi}{\bar{\lambda}} d \sin\alpha = \frac{2\pi n}{\bar{\lambda}_0} d \sin\alpha. \tag{2}$$

where $\alpha \sim \sin\alpha \sim \rho/D$ is the angular radius of the source as seen from the centre of the entrance pupil (Fig. 10.13).

To determine the complex degree of coherence for any pair of points in the plane of the exit pupil, we could apply the propagation law §10.4 (47). However, we may now proceed in a simpler manner as follows:

If $U(S, P_1)$ and $U(S, P_2)$ are the complex disturbances at P_1 and P_2 due to a source point S of the associated monochromatic source (see p. 578), the disturbances due to S at the conjugate points P_1' and P_2' in the exit pupil are, to a good approximation, given by the geometrical optics formulae

$$U(S, P_1') = K_{11} U(S, P_1), \qquad U(S, P_2') = K_{22} U(S, P_2). \tag{3}$$

Here K_{11} and K_{22} are transmission factors that characterize, within the accuracy of geometrical optics, the propagation from P_1 to P_1' and from P_2 to P_2' respectively, along the rays that originate in the source point S. By Hopkins' formula, §10.4 (35b)

$$j(P_1', P_2') = \frac{1}{\sqrt{I(P_1')}\sqrt{I(P_2')}} \int_\sigma U(S, P_1') U^\star(S, P_2') \mathrm{d}S. \tag{4}$$

Now the intensities $I(P_1{}')$ and $I(P_1)$ in the two pupil planes are related by

$$I(P_1') = \int_\sigma |U(S, P_1')|^2 \, \mathrm{d}S = |K_{11}|^2 \int_\sigma |U(S, P_1)|^2 \, \mathrm{d}S = |K_{11}|^2 I(P_1), \tag{5}$$

with a similar relation between $I(P_2{}')$ and $I(P_2)$. From (3), (4) and (5) we have

$$\begin{aligned} j(P_1', P_2') &= \frac{K_{11} K_{22}^\star}{|K_{11}|\,|K_{22}|} \frac{1}{\sqrt{I(P_1)}\sqrt{I(P_2)}} \int_\sigma U(S, P_1) U^\star(S, P_2) \mathrm{d}S \\ &= \mathrm{e}^{\mathrm{i}(\Phi_{11} - \Phi_{22})} j(P_1, P_2), \end{aligned} \tag{6}$$

where Φ_{11} and Φ_{22} are the phases of K_{11} and K_{22} respectively. This relation implies that *the degree of coherence* $|j|$ *for any two points in the exit pupil is equal to the degree of coherence for the conjugate points in the entrance pupil; and the phases of the corresponding values of the complex degree of coherence for corresponding point*

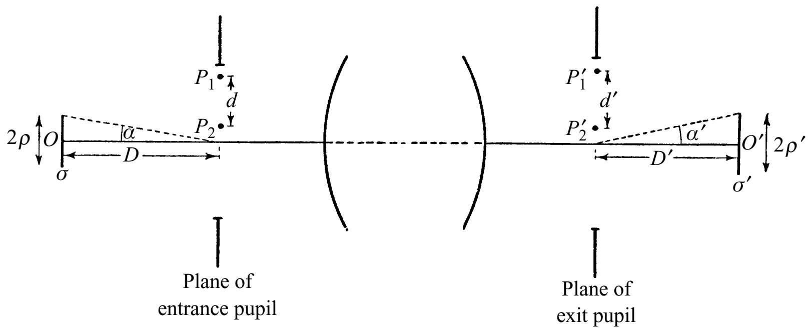


Fig. 10.13 Calculation of the degree of coherence in the image of an incoherent light source.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

pairs differ by the amount $\Phi_{11} - \Phi_{22}$*, i.e. by the geometrical phase difference* $2\pi\{[P_1P_1'] - [P_2P_2']\}/\bar{\lambda}_0$.

Let

$$v' = \frac{2\pi n'}{\bar{\lambda}_0} d' \sin\alpha'. \tag{7}$$

Since $P_1{}'$ is the conjugate of P_1 and $P_2{}'$ the conjugate of P_2, it follows by the Smith–Helmholtz theorem [§4.4 (49)] that, within the accuracy of Gaussian optics,* $v' = v$. Hence, by (1) and (6), the complex degree of coherence for pairs of points in the exit pupil may be written as

$$j(P_1', P_2') = \left[\frac{2J_1(v')}{v'}\right] \mathrm{e}^{\mathrm{i}(\Phi_{11}-\Phi_{22})}. \tag{8}$$

If, as in §10.4 (31), we regard the values $|j| \geqslant 0.88$ as sufficiently close approximations to full coherence, and remember that $|2J(v)/v| \geqslant 0.88$ when $v \leqslant 1$, it follows that *an incoherent quasi-monochromatic uniform circular source will give rise in the exit pupil to coherently illuminated areas of diameter*

$$d'_{\mathrm{coh}} \sim \frac{0.16\bar{\lambda}_0}{n' \sin\alpha'}, \tag{9}$$

where $2\alpha' \sim 2\rho'/\mathrm{D}'$ *is the angle which the diameter of the image of the source subtends at the centre of the exit pupil and* $\bar{\lambda}_0/n' = \bar{\lambda}$ *is the mean wavelength of the light in the image space.*

We shall express (9) in a somewhat different form. Let r_A' denote the radius of the first dark ring in the Airy pattern associated with the system,

$$r_A' = \frac{0.61\bar{\lambda}_0}{n' \sin\theta'}, \tag{10}$$

where $n' \sin\theta' \sim n'a'/\mathrm{D}'$ is the numerical aperture on the image side. Then according to (9) and (10), $d'_{\mathrm{coh}}/r_A' \sim 0.16 \sin\theta'/0.61 \sin\alpha'$, so that

$$d'_{\mathrm{coh}} \sim 0.26a'\left(\frac{r_A'}{\rho'}\right). \tag{11}$$

This formula gives an estimate for the size of the coherently illuminated areas of the exit pupil in terms of the 'physical parameters', viz. the radius r_A' of the first dark ring of the associated Airy pattern, the radius ρ' of the geometrical image of the source, and the radius a' of the exit pupil.

The exit pupil and hence the image plane will be illuminated almost *coherently* if $d'_{\mathrm{coh}} \geqslant 2a'$, i.e. if

$$\rho' \leqslant 0.13 r_A'. \tag{12}$$

When $d'_{\mathrm{coh}} \ll 2a'$, i.e. when

$$\rho' \gg 0.13 r_A', \tag{13}$$

the coherently illuminated areas of the exit pupil will be small compared to the exit

* This means that v and v' represent a particular choice of the Seidel variables (§5.2).

pupil itself so that in this case the illumination of the exit pupil is effectively *incoherent*. The complex degree of coherence for pairs of points Q_1', Q_2' in the image plane will then be essentially the same as that due to an incoherent source; this source has the same size, shape, and position as the exit pupil, and the intensity distribution across this source is the same as the intensity distribution across the exit pupil. Hence according to the van Cittert–Zernike theorem, §10.4 (21),

$$j(Q_1', Q_2') = \frac{1}{\sqrt{I(Q_1')}\sqrt{I(Q_2')}} \int_{\mathcal{A}'} I(P') \frac{e^{i\bar{k}(s_1 - s_2)}}{s_1 s_2} \, dP', \tag{14}$$

$$I(Q_1') = \int_{\mathcal{A}'} \frac{I(P')}{s_1^2} \, dP', \qquad I(Q_2') = \int_{\mathcal{A}'} \frac{I(P')}{s_2^2} \, dP'. \tag{15}$$

The integration is taken over the exit pupil $\mathcal{A}'$ and s_1 and s_2 denote the distances from the typical point P' in $\mathcal{A}'$ to the points Q_1' and Q_2' respectively [Fig. 10.14(a)]. The intensity $I(P')$ may be calculated from the intensity $I(P)$ at the conjugate point in the entrance pupil by means of the relation (5). Since the phase of the transmission function does not appear in this relation, $j(Q_1', Q_2')$ *is independent of the aberrations of the system*. As a rule $I(P')$ is effectively constant; if, moreover, the points Q_1' and Q_2' are sufficiently close to each other the expression (14) then reduces to

$$j(Q_1', Q_2') = \frac{2J_1(u')}{u'}, \qquad u' = \frac{2\pi n'}{\bar{\lambda}_0} \frac{a'}{D'} h', \tag{16}$$

where h' is the distance between Q_1' and Q_2'.

In the general case, when neither condition (12) nor condition (13) holds, the exit pupil is illuminated with partially coherent light, characterized by the complex degree of coherence (8). The value of the complex degree of coherence for pairs of points in the image plane must then be calculated with the help of the propagation law §10.4 (45) and leads to the expression

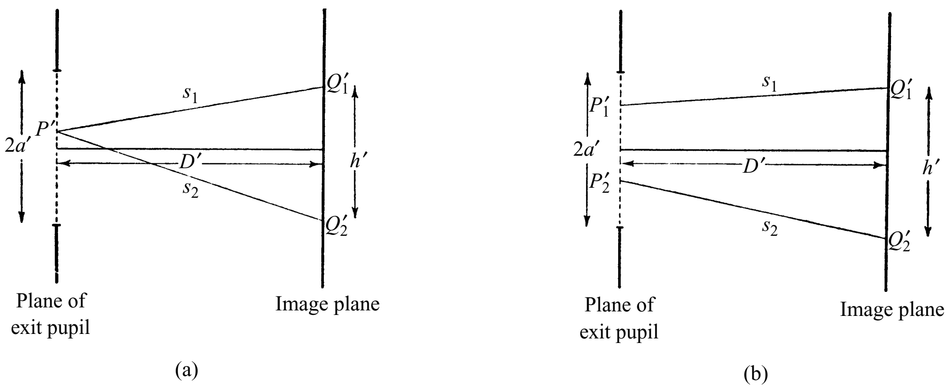


Fig. 10.14 Calculation of the equal time complex degree of coherence $j(Q_1', Q_2')$ in the image plane: (a) incoherent illumination of the exit pupil; (b) partially coherent illumination of the exit pupil.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$j(Q_1', Q_2') = \frac{1}{\sqrt{I(Q_1')}\sqrt{I(Q_2')}} \int_{\mathcal{A}'}\int_{\mathcal{A}'} \sqrt{I(P_1')}\sqrt{I(P_2')} \left[\frac{2J_1(v')}{v'}\right] \frac{e^{i[\Phi_{11}-\Phi_{22}+\bar{k}(s_1-s_2)]}}{s_1 s_2} \times \Lambda_1\Lambda_2^\star \, dP_1' \, dP_2'. \tag{17}$$

The intensities $I(Q_1')$ and $I(Q_2')$ may also be calculated from this formula, if use is made of the fact that $j(Q_1', Q_1') = j(Q_2', Q_2') = 1$. We note that, since the integrand contains the phases Φ_{11} and Φ_{22} of the transmission function, the complex degree of coherence now depends on the aberrations of the system.

10.6.2 The influence of the condenser on resolution in a microscope

In order to examine a small nonluminous object under a microscope, the object has to be illuminated. If, as is usually the case, the object is almost transparent, it is illuminated from behind, or *transilluminated*, and the light which has passed through the object is then focused on to the image plane of the microscope objective. To obtain a sufficient concentration of light, an auxiliary lens system — a condenser — is usually used. Various methods of illumination are employed. We shall briefly describe two commonly used methods, the so-called *critical illumination* and *Köhler's illumination*, and discuss the resolving power which can be achieved with them.

(a) Critical illumination

In this method of illumination a uniformly bright source is placed close behind the field stop and is imaged by the condenser on to the object plane of the microscope objective (Fig. 10.15). The size of the field stop aperture is adjusted so that its image by the condenser just covers the field.

The illuminated region in the image plane of the condenser (the object plane of the objective) is very much larger than the effective size of the Airy pattern due to a single source point ($\rho' \gg r_A'$ in the notation of §10.6.1). It follows from §10.6.1 that under these circumstances *the equal time complex degree of coherence* j_{12} *for any pair of points in the object plane of the objective is the same as that due to an incoherent source filling the condenser aperture*; moreover, it is independent of the aberrations of the condenser. Now it is evident that the resolving power depends only on the degree of coherence (characterized by this factor) of the light incident upon the object and on the properties of the microscope objective. Hence *the aberrations of the condenser have no influence on the resolving power of a microscope*. This important result was

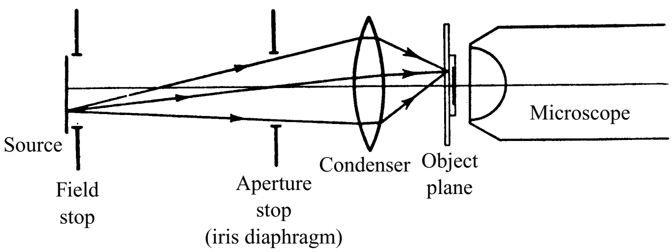


Fig. 10.15 Critical illumination.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

first established in a different manner by Zernike,* and shows that the widely held view, according to which a highly corrected condenser is of advantage for attaining high resolving power, is incorrect.

To estimate the effect of the size of the condenser on the resolution, consider two pinholes $P_1(X_1, Y_1)$ and $P_2(X_2, Y_2)$ in the plane of the object. With the same assumptions as before, the complex degree of coherence of the light that reaches these pinholes is given by a formula of the form (16):

$$j(P_1, P_2) = \frac{2J_1(u_{12})}{u_{12}}, \qquad u_{12} = \frac{2\pi}{\bar{\lambda}_0}\sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}\, n'_c \sin\theta'_c, \tag{18}$$

where $n'_c \sin\theta'_c$ is the numerical aperture of the condenser on the side of the microscope objective.

Let $P(X, Y)$ be any other point in the object plane, and P' its image by the objective. If we assume that the objective is effectively free of aberrations, the intensity distribution, in the image plane of the objective, of the light arriving from P_1 alone is an Airy pattern centred on the image P'_1 of P_1. Hence if $n_0 \sin\theta_0$ is the numerical aperture of the objective, the intensity $I^{(1)}(P')$ due to light that reaches P' from P_1 alone is, apart from a constant factor, equal to

$$I^{(1)}(P') = \left[\frac{2J_1(v_1)}{v_1}\right]^2, \qquad v_1 = \frac{2\pi}{\bar{\lambda}_0}\sqrt{(X - X_1)^2 + (Y - Y_1)^2}\, n_0 \sin\theta_0. \tag{19a}$$

The intensity $I^{(2)}(P')$, due to the light reaching it from the pinhole P_2, is given by a similar expression:

$$I^{(2)}(P') = \left[\frac{2J_2(v_2)}{v_2}\right]^2, \qquad v_2 = \frac{2\pi}{\bar{\lambda}_0}\sqrt{(X - X_2)^2 + (Y - Y_2)^2}\, n_0 \sin\theta_0. \tag{19b}$$

It follows that when the two pinholes are illuminated via the condenser, the intensity $I(P')$ in the image plane of the microscope objective arises from the superposition of two partially coherent beams. The intensity of each beam is given by (19) and the complex degree of coherence of the two beams is given by (18). An expression for $I(P')$ is immediately obtained on substituting from these equations into §10.4 (11). This gives, if we also assume that P' is very close to the geometrical images of P_1 and P_2 (more precisely that $\delta = [P_1 P'] - [P_2 P'] \ll \bar{\lambda}$),

$$I(P') = \left[\frac{2J_1(v_1)}{v_1}\right]^2 + \left[\frac{2J_1(v_2)}{v_2}\right]^2 + 2\left[\frac{2J_1(mv_{12})}{mv_{12}}\right]\left[\frac{2J_1(v_1)}{v_1}\right]\left[\frac{2J_1(v_2)}{v_2}\right], \tag{20}$$

where

$$m = \frac{n'_c \sin\theta'_c}{n_0 \sin\theta_0}, \qquad v_{12} = \frac{u_{12}}{m} = \frac{2\pi}{\bar{\lambda}_0}\sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}\, n_0 \sin\theta_0. \tag{21}$$

Some interesting conclusions may be drawn from (20). When mv_{12} is a root other than $mv_{12} = 0$ of the equation $J_1(mv_{12}) = 0$, the product term is absent and (20) reduces to

* F. Zernike, *Physica*, **5** (1938), 794.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$I(P') = \left[\frac{2J_1(v_1)}{v_1}\right]^2 + \left[\frac{2J_1(v_2)}{v_2}\right]^2. \tag{22}$$

The distribution of the intensity in the image plane is now the same as if P_1 and P_2 were illuminated *incoherently*. In particular this will be the case when $m = 1$ and v_{12} is a nonzero root of $J_1(v_{12}) = 0$; that is, when the numerical apertures are equal and the geometrical images of the pinholes are separated by a distance equal to the radius of any dark ring of the Airy pattern of the objective.

When the numerical aperture of the condenser is very small ($m \to 0$), then $2J_1(mv_{12})/mv_{12} \sim 1$ and (20) reduces to

$$I(P') = \left[\frac{2J_1(v_1)}{v_1} + \frac{2J_1(v_2)}{v_2}\right]^2. \tag{23}$$

The distribution of the intensity is now the same as with perfectly *coherent* illumination, whatever the separation of the pinholes.

Formula (20) makes it possible to study the dependence of the intensity distribution in the image plane of the microscope objective on the ratio m of the numerical apertures. In particular consider the intensity at the midpoint between P_1' and P_2'. We shall regard the pinholes to be just resolved when the intensity at the midpoint is 26.5 per cent smaller than the intensity at either of the two points. The value 26.5 per cent corresponds to Rayleigh's criterion for a circular aperture in incoherent illumination (see §8.6.2). We shall express this limiting separation $(P_1P_2)_{\text{lim}}$ in the same form as for incoherent [§8.6 (32)] and coherent [§8.6 (55)] illuminations:

$$(P_1P_2)_{\text{lim}} = L(m)\frac{\bar{\lambda}_0}{n_0 \sin\theta_0}. \tag{24}$$

The curve $L(m)$ computed from (20) on the basis of this criterion is shown in Fig. 10.16. It is seen that the best resolving power is obtained with $m \sim 1.5$, i.e. when the numerical aperture of the condenser is about 1.5 times that of the objective. The value of L is then slightly smaller than the value 0.61 obtained with incoherent illumination.

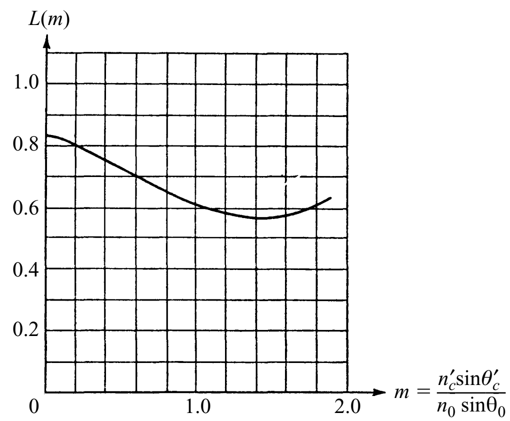


Fig. 10.16 Effect of the condenser aperture on the resolution of two pinholes of equal brightness. (After H. H. Hopkins and P. M. Barham, *Proc. Phys. Soc.*, **63** (1950), 72.)

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

(b) Köhler's illumination

In a method of illumination due to Köhler,* which is illustrated in Fig. 10.17, a converging lens is placed close to the field stop and forms an image of the source σ in the focal plane of the condenser, which now contains the condenser diaphragm. The rays from each source point then emerge from the condenser as a parallel beam. This arrangement has the advantage that the irregularities in the brightness distribution on the source do not cause irregularities in the intensity of the field illumination.

To estimate the limit of resolution that is attained with Köhler's method of illumination we must first determine the complex degree of coherence j_{12} for pairs of points in the object plane of the microscope objective. Let

$$U(S, P_1) = A_1 e^{i\phi_1}, \qquad U(S, P_2) = A_2 e^{i\phi_2}, \tag{25}$$

be the complex disturbances at points $P_1(X_1, Y_1)$ and $P_2(X_2, Y_2)$ of the object plane of the microscope objective, due to a source point S of the monochromatic source associated with σ (see p. 578). Evidently

$$\phi_1 - \phi_2 = \frac{2\pi}{\bar{\lambda}_0}[p(X_1 - X_2) + q(Y_1 - Y_2)], \tag{26}$$

where p and q are the first two ray components of the two parallel rays through P_1 and P_2 from the source point S. If the condenser system suffers from aberrations, the two rays will not be strictly parallel, but as we are only considering points which are close to each other, this effect may be neglected. On substituting from (25) and (26) into Hopkins' formula §10.4 (35b) it follows that

$$\left.\begin{aligned} j(P_1, P_2) &= \frac{1}{\sqrt{I(P_1)}\sqrt{I(P_2)}}\int_\sigma A_1 A_2 e^{i\bar{k}_0[p(X_1-X_2)+q(Y_1-Y_2)]}\, dS, \\ I(P_1) &= \int_\sigma A_1^2\, dS, \qquad I(P_2) = \int_\sigma A_2^2\, dS. \end{aligned}\right\} \tag{27}$$

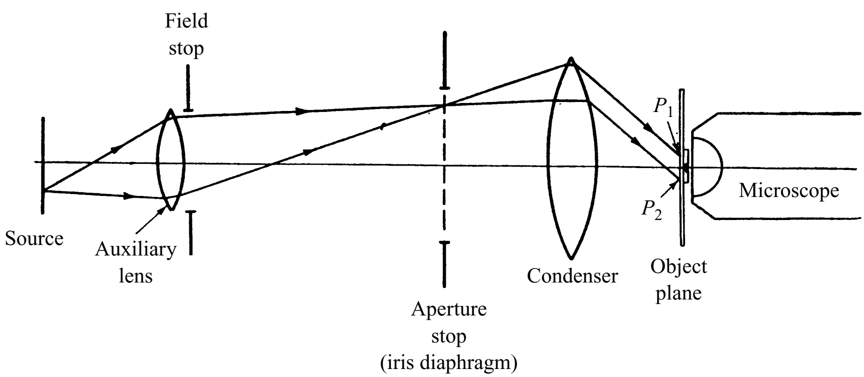


Fig. 10.17 Köhler's method of illumination.

* A. Köhler, *Zs. f. wiss. Mikrosk.*, **10** (1893), 433; **16** (1899), 1.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

Since to every source point $S(\xi, \eta)$ there corresponds a pair of ray components (p, q), we may transform the integrals over σ into integrals over the solid angle

$$p^2 + q^2 \leqslant n_c'^2 \sin^2 \theta_c', \tag{28}$$

formed by the rays that are incident upon the object. Now within the accuracy of Gaussian optics, the relations $\xi = \xi(p, q)$, $\eta = \eta(p, q)$ are linear; in fact, as is easily seen from the formulae of §4.3 (10), $\xi = fp$, $\eta = fq$, where f is the focal length of the condenser system. Hence the Jacobian $\partial(\xi, \eta)/\partial(p, q)$ of the transformation is constant. Outside the domain of geometrical optics, the Jacobian will in general vary over the region of integration, but this variation may be assumed to be slow compared to the variation of the exponential term and may therefore be neglected. If we also neglect the slow variation of A_1 and A_2, (27) reduces to

$$j(P_1, P_2) = \frac{\displaystyle\iint_\Omega e^{i\bar{k}_0[p(X_1 - X_2) + q(Y_1 - Y_2)]}\, dp\, dq}{\displaystyle\iint_\Omega dp\, dq}, \tag{29}$$

where Ω denotes the domain (28). Evaluation of (29) leads to the expression

$$j(P_1, P_2) = \frac{2J_1(u_{12})}{u_{12}}, \qquad u_{12} = \frac{2\pi}{\bar{\lambda}_0}\sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}\, n_c' \sin\theta_c'. \tag{30}$$

This formula is identical with the formula (18) for critical illumination. Hence *the complex degree of coherence of the light incident upon the object plane of a microscope is the same whether critical or Köhler's illumination is employed*. In view of this result it is somewhat unfortunate that critical illumination is often designated as 'incoherent' and Köhler's illumination as 'coherent'. It follows that formula (20) holds with both types of illumination, and that Fig. 10.16 likewise applies in both cases.

*10.6.3 Imaging with partially coherent quasi-monochromatic illumination**

(a) Transmission of mutual intensity through an optical system

In §9.5 some general methods were described for the study of imagery of extended objects. The cases of completely coherent illumination (§9.5.1) and completely incoherent illumination (§9.5.2) were studied. In the first case the transmission of the complex amplitude through the system was considered, in the second case the transmission of the intensity. We shall now investigate the more general case of partially coherent quasi-monochromatic illumination. The appropriate quantity to consider in this case is the mutual intensity.

We shall employ the same scale-normalized (Seidel) coordinates as in §9.5 (1), so that the object point and its Gaussian image have the same coordinate numbers. Let $J_0(x_0, y_0; x_0', y_0')$ be the mutual intensity for points (x_0, y_0), (x_0', y_0') in the object plane. If $K(x_0, y_0; x_1, y_1)$ is the transmission function of the system (§9.5.1), the

* The considerations of this section are based in parts on investigations of H. H. Hopkins, *Proc. Roy. Soc.*, A, **217** (1953), 408 and P. Dumontet, *Publ. Sci. Univ. d'Alger*, B, **1** (1955), 33.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

mutual intensity in the image plane is, according to the propagation law §10.4 (47), given by

$$J_1(x_1, y_1; x_1', y_1') = \iiiint\limits_{-\infty}^{+\infty} J_0(x_0, y_0; x_0', y_0')K(x_0, y_0; x_1, y_1)K^\star(x_0', y_0'; x_1', y_1')\mathrm{d}x_0\,\mathrm{d}y_0\,\mathrm{d}x_0'\,\mathrm{d}y_0'. \tag{31a}$$

The integration extends only formally over an infinite domain, since J_0 is zero for all points in the object plane from which no light proceeds to the image plane.

As in §9.5, we assume that the object is so small that it forms an isoplanatic region of the system, i.e. that for all points on it $K(x_0, y_0; x_1, y_1)$ may be replaced to a good approximation by a function depending on the differences $x_1 - x_0$ and $y_1 - y_0$ only, say $K(x_1 - x_0, y_1 - y_0)$. Eq. (31a) then becomes

$$J_1(x_1, y_1; x_1', y_1') = \iiiint\limits_{-\infty}^{+\infty} J_0(x_0, y_0; x_0', y_0')K(x_1 - x_0, y_1 - y_0) \times K^\star(x_1' - x_0', y_1' - y_0')\mathrm{d}x_0\,\mathrm{d}y_0\,\mathrm{d}x_0'\,\mathrm{d}y_0'. \tag{31b}$$

We represent J_0, J_1 and the product $KK^\star$ in the form of four-dimensional Fourier integrals:

$$J_0(x_0, y_0; x_0', y_0') = \iiiint\limits_{-\infty}^{+\infty} \mathcal{J}_0(f, g; f', g')\mathrm{e}^{-2\pi\mathrm{i}(fx_0+gy_0+f'x_0'+g'y_0')}\,\mathrm{d}f\,\mathrm{d}g\,\mathrm{d}f'\,\mathrm{d}g', \tag{32a}$$

$$J_1(x_1, y_1; x_1', y_1') = \iiiint\limits_{-\infty}^{+\infty} \mathcal{J}_1(f, g; f', g')\mathrm{e}^{-2\pi\mathrm{i}(fx_1+gy_1+f'x_1'+g'y_1')}\,\mathrm{d}f\,\mathrm{d}g\,\mathrm{d}f'\,\mathrm{d}g', \tag{32b}$$

$$K(x, y)K^\star(x', y') = \iiiint\limits_{-\infty}^{+\infty} \mathcal{M}(f, g; f', g')\mathrm{e}^{-2\pi\mathrm{i}(fx+gy+f'x'+g'y')}\,\mathrm{d}f\,\mathrm{d}g\,\mathrm{d}f'\,\mathrm{d}g'. \tag{32c}$$

Then by the Fourier inversion formula

$$\mathcal{J}_0(f, g; f', g') = \iiiint\limits_{-\infty}^{+\infty} J_0(x_0, y_0; x_0', y_0')\mathrm{e}^{2\pi\mathrm{i}(fx_0+gy_0+f'x_0'+g'y_0')}\,\mathrm{d}x_0\,\mathrm{d}y_0\,\mathrm{d}x_0'\,\mathrm{d}y_0', \tag{33}$$

and there are strictly analogous relations for $\mathcal{J}_1$ and $\mathcal{M}$.

On applying the convolution theorem to (31b) we obtain the relation

$$\mathcal{J}_1(f, g; f', g') = \mathcal{J}_0(f, g; f', g')\mathcal{M}(f, g; f', g'). \tag{34}$$

This formula implies that, if the mutual intensity in the object and image planes are represented as a superposition of four-dimensional space-harmonic components of all possible spatial frequencies (f, g, f', g'), then each component in the image depends only on the corresponding component in the object, and the ratio of the components is equal to $\mathcal{M}$. Thus within the accuracy of the present approximation *the action of the optical system on the mutual intensity is equivalent to the action of a four-dimensional linear filter*. $\mathcal{M}$ is called *the frequency response function for partially coherent quasi-monochromatic illumination*.

The frequency response function $\mathcal{M}$ is related in a simple way to the pupil function of the system. If as in §9.5 (10c), we represent K in the form of a two-dimensional Fourier integral

$$K(x, y) = \iint_{-\infty}^{+\infty} \mathcal{K}(f, g) e^{-2\pi i(fx+gy)} \, df \, dg, \tag{35}$$

and substitute for K into the inverse of (32c), we find that

$$\mathcal{M}(f, g; f', g') = \mathcal{K}(f, g)\mathcal{K}^{\star}(-f', -g'). \tag{36}$$

But, according to §9.5 (13), $\mathcal{K}(f, g)$ is equal to the value of the pupil function $G(\xi, \eta)$ of the system at the point

$$\xi = \lambda R f, \qquad \eta = \lambda R g, \tag{37}$$

on the Gaussian reference sphere (radius R). Hence *the frequency response function for partially coherent quasi-monochromatic illumination is connected with the pupil function of the system by the formula*

$$\mathcal{M}\left(\frac{\xi}{\lambda R}, \frac{\eta}{\lambda R}; \frac{\xi'}{\lambda R}, \frac{\eta'}{\lambda R}\right) = G(\xi, \eta)G^{\star}(-\xi', -\eta'). \tag{38}$$

Since the pupil function is zero for points outside the area of the exit pupil, it follows that spectral components belonging to frequencies above certain values are not transmitted. If the exit pupil is a circle of radius a, then $G(\xi, \eta)G^{\star}(-\xi', -\eta')$ vanishes if $\xi^2 + \eta^2 > a^2$ or $\xi'^2 + \eta'^2 > a^2$. Hence spectral components of the mutual intensity belonging to frequencies (f, g, f', g') for which either

$$f^2 + g^2 > \left(\frac{a}{\bar{\lambda} R}\right)^2 \quad \text{or} \quad f'^2 + g'^2 > \left(\frac{a}{\bar{\lambda} R}\right)^2, \tag{39}$$

are not transmitted.* Here $\bar{\lambda}$ denotes the mean wavelength in the image space.

In Table 10.1, the basic formulae relating to imaging with partially coherent illumination are displayed, together with the corresponding formulae of §9.5, relating to coherent and incoherent illumination. The formulae relating to incoherent illumination may be derived from the general formulae (34), (36) and (38) by assuming J_0 to be of the form $J_0(x_0, y_0; x_0', y_0') = I_0(x_0, y_0)\delta(x_0' - x_0)\delta(y_0' - y_0)$, where δ is the Dirac delta function (see Appendix IV). The calculations are straightforward but somewhat lengthy and will not be given here. Formulae relating to the special case of

* If the angular aperture of the system is small and the sine condition is obeyed, we have, as in §9.5, p. 547, $a/\bar{\lambda}R \sim n_0 \sin\theta_0/M\bar{\lambda}_0$, where $n_0 \sin\theta_0$ is the numerical aperture of the system, M the Gaussian magnification, and $\bar{\lambda}_0$ is the mean wavelength in vacuum.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

Table 10.1. *The action of an optical system from the standpoint of its response to spatial frequencies. (Isoplanatic object region assumed.)*

Illumination	Basic quantity	Transition from object to image	Frequency response function
Coherent	Complex disturbance $U(x, y)$	$\mathcal{U}_1(f, g) = \mathcal{U}_0(f, g)\mathcal{K}(f, g)$	$\mathcal{K}(f, g)$
Incoherent	Intensity $I(x, y)$	$\mathcal{J}_1(f, g) = \mathcal{J}_0(f, g)\mathcal{L}(f, g)$	$\mathcal{L}(f, g) = \iint_{-\infty}^{+\infty} \mathcal{K}(f'+f, g'+g) \times \mathcal{K}^\star(f', g')\mathrm{d}f'\,\mathrm{d}g'$
Partially coherent	Mutual intensity $J(x, y; x', y')$	$\mathcal{J}_1(f, g; f', g') = \mathcal{J}_0(f, g; f', g')\mathcal{M}(f, g; f', g')$	$\mathcal{M}(f, g; f', g') = \mathcal{K}(f, g)\mathcal{K}^\star(-f', -g')$

$\mathcal{K}(f, g)$ is the Fourier inverse of the transmission function $K(x, y)$ of the system and is related to the pupil function $G(\xi, \eta)$ by the formula $\mathcal{K}(\xi/\lambda R, \eta/\lambda R) = G(\xi, \eta)$, where R is the radius of the Gaussian reference sphere and λ the mean wavelength in the image space.

perfectly monochromatic (and hence completely coherent) illumination may be derived somewhat more easily by noting that in this case the mutual intensity is of the form $J_0(x_0, y_0; x_0', y_0') = U_0(x_0, y_0)U_0^\star(x_0', y_0')$. The Fourier transform $\mathcal{J}_0$ of J_0 now likewise splits into the product of two factors and the application of (34) and (36) shows that each spectral component of the complex disturbance U_0 is transmitted through the system in accordance with the formulae displayed in the first row of Table 10.1.

(b) Images of transilluminated objects

Suppose that a portion of the object plane is occupied by a transparent or semitransparent object which is illuminated with partially coherent quasi-monochromatic light. It will be assumed that this light originates in an incoherent source and reaches the object plane after the passage through some illuminating system (condenser).

As in §8.6.1, we specify the object by an appropriate transmission function $F_0(x_0, y_0)$. If $U_0^-(S; x_0, y_0)$ represents the disturbance at the point (x_0, y_0) of the object plane due to a source point S of the associated monochromatic source (see p. 578), the disturbance from this source point after the passage through the object is given by

$$U_0(S; x_0, y_0) = U_0^-(S; x_0, y_0)F(x_0, y_0). \tag{40}$$

Now according to §10.4 (35a) the mutual intensity of the light incident on the object is given by

$$J_0^-(x_0, y_0; x_0', y_0') = \int_\sigma U_0^-(S; x_0, y_0)U_0^{-\star}(S; x_0', y_0')\mathrm{d}S, \tag{41a}$$

and the mutual intensity of the light emerging from the object is

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$J_0(x_0, y_0; x_0', y_0') = \int_\sigma U_0(S; x_0, y_0) U_0^\star(S; x_0', y_0') \mathrm{d}S, \tag{41b}$$

so that, because of (40),

$$J_0(x_0, y_0; x_0', y_0') = F(x_0, y_0) F^\star(x_0', y_0') J_0^-(x_0, y_0; x_0', y_0'). \tag{42}$$

We shall confine our attention to the important case when the mutual intensity J_0^- of the incident light depends on the four coordinates x_0, y_0, x_0', y_0' through the differences $x_0 - x_0', y_0 - y_0'$ only, i.e. when J_0^- is of the form

$$J_0^-(x_0, y_0; x_0', y_0') = J_0^-(x_0 - x_0', y_0 - y_0'). \tag{43}$$

We learned in §10.6.2 that this will be the case for both the critical and the Köhler illumination. We retain the earlier assumptions that the object is so small that it forms an isoplanatic area of the system. It then follows from (31b) that the intensity $I_1(x_1, y_1) = J_1(x_1, y_1; x_1, y_1)$ in the image plane is given by

$$I_1(x_1, y_1) = \iiiint\limits_{-\infty}^{+\infty} J_0^-(x_0 - x_0', y_0 - y_0') F(x_0, y_0) F^\star(x_0', y_0') K(x_1 - x_0, y_1 - y_0)$$
$$\times K^\star(x_1 - x_0', y_1 - y_0') \mathrm{d}x_0 \, \mathrm{d}y_0 \, \mathrm{d}x_0' \, \mathrm{d}y_0'. \tag{44}$$

We represent F and $J_0{}^-$ in the form of two-dimensional Fourier integrals

$$F(x, y) = \iint\limits_{-\infty}^{+\infty} \mathcal{F}(f, g) \mathrm{e}^{-2\pi \mathrm{i}(fx + gy)} \, \mathrm{d}f \, \mathrm{d}g, \tag{45a}$$

$$J_0^-(x, y) = \iint\limits_{-\infty}^{+\infty} \mathcal{J}_0^-(f, g) \mathrm{e}^{-2\pi \mathrm{i}(fx + gy)} \, \mathrm{d}f \, \mathrm{d}g. \tag{45b}$$

If we substitute for F and $F^\star$ from (45a) into (44), use the identity $f'x_0 - f''x_0' = (f' - f'')x_1 - f'(x_1 - x_0) + f''(x_1 - x_0')$ and a similar identity involving g and y, and introduce new variables of integration $u' = x_1 - x_0$, $u'' = x_1 - x_0'$, we obtain the following expression for I_1:

$$I_1(x_1, y_1) = \iiiint\limits_{-\infty}^{+\infty} \mathcal{T}(f', g'; f'', g'') \mathcal{F}(f', g') \mathcal{F}^\star(f'', g'')$$
$$\times \mathrm{e}^{-2\pi \mathrm{i}[(f' - f'')x_1 + (g' - g'')y_1]} \, \mathrm{d}f' \, \mathrm{d}g' \, \mathrm{d}f'' \, \mathrm{d}g'', \tag{46}$$

where

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$\mathcal{T}(f', g'; f'', g'') = \iiiint\limits_{-\infty}^{+\infty} J_0^-(u''-u', v''-v')K(u', v')K^\star(u'', v'')$$
$$\times\, e^{2\pi i[(f'u'+g'v')-(f''u''+g''v'')]}\,du'\,dv'\,du''\,dv''$$
$$= \int\!\!\int\!\!\int\!\!\int\!\!\int\!\!\int\limits_{-\infty}^{+\infty} \mathcal{J}_0^-(f, g)K(u', v')K^\star(u'', v'')$$
$$\times\, e^{2\pi i[(f+f')u'+(g+g')v'-(f+f'')u''-(g+g'')v'']}\,df\,dg\,du'\,dv'\,du''\,dv''$$
$$= \iint\limits_{-\infty}^{+\infty} \mathcal{J}_0^-(f, g)\mathcal{K}(f+f', g+g')\mathcal{K}^\star(f+f'', g+g'')df\,dg. \tag{47}$$

In passing from the second to the third lines in (47) we substituted for J_0^- from (45b) and in passing from the third to the fourth line we used the inverse of (35).

We see that in (46) the influence of the object (characterized by $\mathcal{F}$) and the combined effect of the illumination ($\mathcal{J}_0^-$) and of the system ($\mathcal{K}$) are separated. With a uniform illumination ($I_0^- =$ constant), the intensity of the light emerging from the object would be proportional to $|F|^2$, and were the imaging perfect, the intensity in the image plane would be given (apart from a constant factor) by

$$\tilde{I}_1(x_1, y_1) = F(x_1, y_1)F^\star(x_1, y_1)$$
$$= \iiiint\limits_{-\infty}^{+\infty} \mathcal{F}(f', g')\mathcal{F}^\star(f'', g'')e^{-2\pi i[(f'-f'')x_1+(g'-g'')y_1]}\,df'\,dg'\,df''\,dg''. \tag{48}$$

Eqs. (46) and (48) represent the true intensity I_1 and the ideal intensity $\tilde{I}_1$ as sums of contributions from all pairs of frequencies (f', g'), (f'', g'') of the spatial spectrum of the object. Each contribution in the former case is $\mathcal{T}$ times that of the latter, and it follows that unless $\mathcal{T}$ is constant for all values f', g', f'', g'' for which both the spectral components $\mathcal{F}(f', g')$, $\mathcal{F}(f'', g'')$ differ from zero, some information about the object will be lost or falsified. The function $\mathcal{T}$ is called *the transmission cross-coefficient* of the system, working with the given transillumination.

Instead of the intensity itself, let us now consider its spatial spectrum $\mathcal{J}(f, g)$. To derive the appropriate formula we multiply both sides of (46) by $e^{2\pi i(fx_1+gy_1)}$ and integrate with respect to x_1 and y_1. If next we use the Fourier integral theorem (or more shortly the Fourier integral representation of the Dirac delta function (see Appendix IV), we obtain

$$\mathcal{J}_1(f, g) = \iint\limits_{-\infty}^{+\infty} \mathcal{T}(f'+f, g'+g, f', g')\mathcal{F}(f'+f, g'+g)\mathcal{F}^\star(f', g')df'\,dg'. \tag{49}$$

For the ideal case represented by (48) we have

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$\tilde{\mathcal{J}}_1(f, g) = \iint_{-\infty}^{+\infty} \mathcal{F}(f' + f, g' + g)\mathcal{F}^{\star}(f', g')\mathrm{d}f'\,\mathrm{d}g'. \tag{50}$$

These formulae express $\mathcal{J}_1$ and $\tilde{\mathcal{J}}_1$ as the sums of contributions from each spatial frequency (f', g') of the object structure. It is seen that $\mathcal{T}$ plays a similar role as before. It characterizes the changes which arise in each contribution from the mode of illumination of the object and from the transmission characteristics of the image-forming system.

Since the response function $\mathcal{K}(f, g)$ is zero when the point $\xi = \bar{\lambda}Rf$, $\eta = \bar{\lambda}Rg$ lies outside the area of the exit pupil, it follows from (47) that $\mathcal{T}$ vanishes for sufficiently high frequencies. If the exit pupil is a circle of radius a, the product $\mathcal{K}(f + f', g + g')\mathcal{K}^{\star}(f + f'', g + g'')$ and consequently $\mathcal{T}(f', g'; f''g'')$ can only differ from zero if the two circles C' and C'', each of radius $\sqrt{f^2 + g^2} = a/\bar{\lambda}R$, centred on the points $O'(-f', -g')$, $O''(-f'', -g'')$ in the f,g-plane have a domain in common (Fig. 10.18). To illustrate the effect of the illumination, suppose that the illumination is either critical or Köhler, and that the numerical aperture $n_c' \sin\theta_c'$ of the condenser system is m times the numerical aperture $n_0 \sin\theta_0$ of the system that images the object. Then according to (18) or (30) the mutual intensity of the illuminating beam is

$$J_0^-(x_0 - x_0', y_0 - y_0') = \left[\frac{2J_1(mv)}{mv}\right] I_0^-, \tag{51}$$

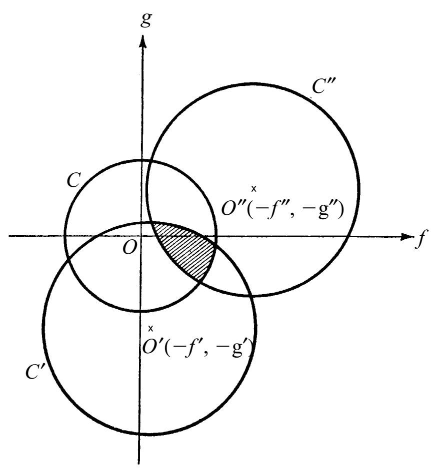


Fig. 10.18 The effective domain of integration (shown shaded) for the transmission cross-coefficient $\mathcal{T}(f', g'; f'', g'')$ of an image-forming system with a circular exit pupil of radius a. The object is assumed to be transilluminated by quasi-monochromatic light of mean wavelength $\bar{\lambda}_0$ via a condenser of numerical aperture $mn_0 \sin\theta_0$, where $n_0 \sin\theta_0$ is the numerical aperture of the image-forming system. C' and C'' are circles of radius $a/\bar{\lambda}R \sim n_0 \sin\theta_0/M\bar{\lambda}_0$ centred at the points $O'(-f', -g')$ and $O''(-f'', -g'')$ respectively. C is a circle of radius $mn_0 \sin\theta_0/\bar{\lambda}_0$ centred on the origin. (R = radius of Gaussian reference sphere, M = Gaussian magnification.)

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

where

$$v = \frac{2\pi}{\bar{\lambda}_0}\sqrt{(x_0' - x_0)^2 + (y_0' - y_0)^2}\, n_0 \sin\theta_0, \tag{52}$$

and I_0^- is the intensity (assumed uniform) of the incident light. Now the function on the right of (51) is a Fourier transform of the function (see §8.5.2)

$$\left.\begin{aligned} \mathcal{J}_0^-(f,\, g) &= \text{constant} = \left(\frac{\bar{\lambda}_0^2}{\pi m^2 n_0^2 \sin^2\theta_0}\right) I_0^- \quad \text{when} \quad f^2 + g^2 < \frac{m^2 n_0^2 \sin^2\theta_0}{\bar{\lambda}_0^2}, \\ &= 0 \qquad\qquad\qquad\qquad\qquad\qquad \text{when} \quad f^2 + g^2 > \frac{m^2 n_0^2 \sin^2\theta_0}{\bar{\lambda}_0^2}. \end{aligned}\right\} \tag{53}$$

In Fig. 10.17 the circle outside which $\mathcal{J}_0^-$ vanishes is denoted by C and it follows that for given $(f',\, g')$ and $(f'',\, g'')$ only those points of the f,g-plane which lie within the area (shown shaded) common to the three circles C, C' and C'' contribute to the integral (47) for $\mathcal{T}$.

10.7 Some theorems relating to mutual coherence

In §10.4 and §10.6 we considered interference and diffraction of quasi-monochromatic light, and confined ourselves to situations where the time delay τ was small compared to the coherence time of the light. We have seen that to a good approximation the correlation functions then depend on τ only through a harmonic term, i.e. they are of the form

$$\Gamma(P_1,\, P_2,\, \tau) \sim J(P_1,\, P_2)\mathrm{e}^{-2\pi \mathrm{i}\bar{\nu}\tau}, \qquad \gamma(P_1,\, P_2,\, \tau) \sim j(P_1,\, P_2)\mathrm{e}^{-2\pi \mathrm{i}\bar{\nu}\tau}.$$

The elementary theory which makes use of this approximation is adequate to take into account the decrease in the visibility at the 'centre' of the pattern ($\tau = 0$), due to a finite extension of the light source. However, it does not take into account the change in the visibility with increasing path difference. To describe adequately situations in which the time delay τ is not negligible in comparison with the coherence time, it is necessary to use more accurate expressions for the correlation functions. We shall now discuss the appropriate generalization of some of our formulae.

10.7.1 Calculation of mutual coherence for light from an incoherent source

Let $V_1(t)$ and $V_2(t)$ be the disturbances at points P_1 and P_2 in a wave field produced by an extended (not necessarily quasi-monochromatic) incoherent source σ. To begin with we assume that the medium between σ and the points P_1 and P_2 is homogeneous.

As in §10.4.2, we imagine the source to be divided into elements $\mathrm{d}\sigma_1, \mathrm{d}\sigma_2 \ldots$ centred about points $S_1, S_2, \ldots,$ with linear dimensions that are small compared to the effective wavelengths. If $V_{m1}(t)$ and $V_{m2}(t)$ are the contributions to V_1 and V_2 from the element $\mathrm{d}\sigma_m$, then

$$V_1(t) = \sum_m V_{m1}(t), \qquad V_2(t) = \sum_m V_{m2}(t), \tag{1}$$

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

and the mutual coherence function is given by

$$\Gamma(P_1, P_2, \tau) = \langle V_1(t+\tau)V_2^\star(t)\rangle = \sum_m \langle V_{m1}(t+\tau)V_{m2}^\star(t)\rangle. \tag{2}$$

Terms of the type $\langle V_{m1}(t+\tau)V_{n2}^\star(t)\rangle$ $(m \neq n)$ have been omitted on the right-hand of (2), since the contributions from the different source elements may be assumed to be mutually incoherent.

We now proceed somewhat differently than in §10.4. According to §10.3 (32), each term under the summation sign in (2) may be expressed in the form

$$\langle V_{m1}(t+\tau)V_{m2}^\star(t)\rangle = 4\int_0^\infty G_m(P_1, P_2, \nu)\mathrm{e}^{-2\pi \mathrm{i}\nu\tau}\,\mathrm{d}\nu \tag{3}$$

where

$$G_m(P_1, P_2, \nu) = \operatorname*{Lim}_{T\to\infty}\left[\frac{\overline{v_{mT}(P_1, \nu)v_{mT}^\star(P_2, \nu)}}{2T}\right] \tag{4}$$

is the mutual spectral density of the disturbances $V_{m1} = V_m(P_1, t)$, $V_{m2} = V_m(P_2, t)$. Now v_m represents the contribution of the appropriate frequency to the disturbance arising from the element $\mathrm{d}\sigma_m$, and, since the medium is assumed to be homogeneous this contribution is propagated in the form of a spherical wave. Hence

$$v_{mT}(P_1, \nu) = a_{mT}(\nu)\frac{\mathrm{e}^{\mathrm{i}kR_{m1}}}{R_{m1}}, \qquad v_{mT}(P_2, \nu) = a_{mT}(\nu)\frac{\mathrm{e}^{\mathrm{i}kR_{m2}}}{R_{m2}}, \tag{5}$$

where R_{m1} and R_{m2} (assumed to be large compared to the effective wavelengths) represent the distances of P_1 and P_2 from the source point S_m, and $k = 2\pi\nu/v = 2\pi/\lambda$. The amplitude $|a_m(\nu)|$ of $a_m(\nu)$ represents the strength of the component of frequency ν from the element $\mathrm{d}\sigma_m$, and $\arg a_m(\nu)$ represents its phase. From (4) and (5),

$$G_m(P_1, P_2, \nu) = \left[\operatorname*{Lim}_{T\to\infty}\frac{\overline{|a_{mT}(\nu)|^2}}{2T}\right]\frac{\mathrm{e}^{\mathrm{i}k(R_{m1}-R_{m2})}}{R_{m1}R_{m2}}. \tag{6}$$

The term in the large bracket on the right represents the spectral density of the light from the source element $\mathrm{d}\sigma_m$. As in §10.4.2, we assume that the number of the source elements is so large that we may treat the source as being effectively continuous. Hence if $I(S_m, \nu)\mathrm{d}\sigma_m\,\mathrm{d}\nu = 4\,\mathrm{Lim}_{T\to\infty}\,[\overline{|a_{mT}(\nu)|^2}/2T]\mathrm{d}\nu$, i.e. if $I(S, \nu)$ denotes the intensity per unit area of the source, per unit frequency range, then according to (2), (3) and (6),

$$\Gamma(P_1, P_2, \tau) = \sqrt{I(P_1)}\sqrt{I(P_2)}\gamma(P_1, P_2, \tau) = \int_0^\infty \mathrm{d}\nu\,\mathrm{e}^{-2\pi\mathrm{i}\nu\tau}\int_\sigma I(S, \nu)\frac{\mathrm{e}^{\mathrm{i}k(R_1-R_2)}}{R_1R_2}\,\mathrm{d}S, \tag{7}$$

where

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$\left.\begin{aligned} I(P_1) = \Gamma(P_1, P_1, 0) = \int_0^\infty \mathrm{d}\nu \int_\sigma \frac{I(S, \nu)}{R_1^2}\,\mathrm{d}S, \\ I(P_2) = \Gamma(P_2, P_2, 0) = \int_0^\infty \mathrm{d}\nu \int_\sigma \frac{I(S, \nu)}{R_2^2}\,\mathrm{d}S, \end{aligned}\right\} \tag{8}$$

are the intensities at P_1 and P_2, and R_1 and R_2 are the distances of these points from the source point S. Eqs. (7) are generalizations of the van Cittert–Zernike formulae, §10.4 (20), §10.4 (21).

If the medium between the source and the points P_1 and P_2 is not homogeneous, we may proceed as in §10.4. We only have to replace the factors $\mathrm{e}^{\mathrm{i}kR_j}/R_j$ by $\mathrm{i}cK(S, P_j, \nu)/\nu$, where K is the appropriate transmission function of the medium. In place of (7) we then obtain

$$\begin{aligned} \Gamma(P_1, P_2, \tau) &= \sqrt{I(P_1)}\sqrt{I(P_2)}\gamma(P_1, P_2, \tau) \\ &= c^2\int_0^\infty \frac{\mathrm{d}\nu}{\nu^2}\mathrm{e}^{-2\pi\mathrm{i}\nu\tau}\int_\sigma I(S, \nu)K(S, P_1, \nu)K^\star(S, P_2, \nu)\mathrm{d}S, \end{aligned} \tag{9}$$

where

$$\left.\begin{aligned} I(P_1) = c^2\int_0^\infty \frac{\mathrm{d}\nu}{\nu^2}\int_\sigma I(S, \nu)|K(S, P_1, \nu)|^2\,\mathrm{d}S, \\ I(P_2) = c^2\int_0^\infty \frac{\mathrm{d}\nu}{\nu^2}\int_\sigma I(S, \nu)|K(S, P_2, \nu)|^2\,\mathrm{d}S. \end{aligned}\right\} \tag{10}$$

By analogy with §10.4, we re-write (9) or (10) in a slightly different form. We set

$$\frac{\mathrm{i}c}{\nu}K(S, P_1, \nu)\sqrt{I(S, \nu)} = U(S, P_1, \nu), \qquad \frac{\mathrm{i}c}{\nu}K(S, P_2, \nu)\sqrt{I(S, \nu)} = U(S, P_2, \nu). \tag{11}$$

Then (9) and (10) become

$$\begin{aligned} \Gamma(P_1, P_2, \tau) &= \sqrt{I(P_1)}\sqrt{I(P_2)}\gamma(P_1, P_2, \tau) \\ &= \int_0^\infty \mathrm{d}\nu\mathrm{e}^{-2\pi\mathrm{i}\nu\tau}\int_\sigma U(S, P_1, \nu)U^\star(S, P_2, \nu)\mathrm{d}S, \end{aligned} \tag{12}$$

where

$$I(P_1) = \int_0^\infty \mathrm{d}\nu\int_\sigma |U(S, P_1, \nu)|^2\,\mathrm{d}S, \qquad I(P_2) = \int_0^\infty \mathrm{d}\nu\int_\sigma |U(S, P_2, \nu)|^2\,\mathrm{d}S. \tag{13}$$

Formula (12), which is a generalization of Hopkins' formula §10.4 (35), expresses the mutual coherence function and the complex degree of coherence in terms of the light distribution arising from an associated fictitious source. For, according to (11), $U(S, P, \nu)$ may be regarded as the disturbance at P due to a monochromatic point source of frequency ν, of zero phase, and of strength proportional to $\sqrt{I(S, \nu)}$, situated at S.

10.7.2 Propagation of mutual coherence

As in §10.4.4 let $\mathcal{A}$ be a fictitious surface which intercepts a beam of light from an extended incoherent source σ (Fig. 10.8). We shall show how, from knowledge of the mutual coherence for all pairs of points on $\mathcal{A}$, its value on any other surface $\mathcal{B}$ illuminated by the light from $\mathcal{A}$ can be determined. For simplicity we assume that the medium between $\mathcal{A}$ and $\mathcal{B}$ has refractive index equal to unity.

The mutual coherence for any two points Q_1 and Q_2 on $\mathcal{B}$ may be calculated by the use of (12). In this formula U represents a monochromatic disturbance and hence may be determined from the knowledge of the disturbance at all points of the surface $\mathcal{A}$ by means of the Huygens–Fresnel principle:

$$U(S,\, Q_1,\, \nu) = \int_{\mathcal{A}} U(S,\, P_1,\, \nu)\frac{\mathrm{e}^{\mathrm{i}k_0 s_1}}{s_1}\Lambda_1\,\mathrm{d}P_1. \tag{14}$$

Here s_1 denotes the distance between a typical point P_1 on $\mathcal{A}$ and the point Q_1, and Λ_1 is the usual inclination factor. There is a strictly analogous formula for $U(S,\, Q_2,\, \nu)$. If we substitute from these formulae into (12) we obtain, after changing the order of integration, the following expression for the mutual coherence:

$$\Gamma(Q_1,\, Q_2,\, \tau) = \int_{\mathcal{A}}\int_{\mathcal{A}} \frac{\mathrm{d}P_1\,\mathrm{d}P_2}{s_1 s_2}\int_0^\infty J(P_1,\, P_2,\, \nu)\Lambda_1(\nu)\Lambda_2^\star(\nu)\mathrm{e}^{-2\pi\mathrm{i}\nu[\tau-\frac{s_1-s_2}{c}]}\,\mathrm{d}\nu, \tag{15}$$

where

$$J(P_1,\, P_2,\, \nu) = \int_\sigma U(S,\, P_1,\, \nu)U^\star(S,\, P_2,\, \nu)\mathrm{d}S. \tag{16}$$

The repeated integration in (15) means that the points P_1 and P_2 explore the surface $\mathcal{A}$ independently. Now the inclination factors $\Lambda_1(\nu)$ and $\Lambda_2(\nu)$ depend on the frequency through a multiplicative factor ν and change slowly with ν in comparison with the other terms. If the effective spectral range of the light is sufficiently small, we may replace these factors by $\overline{\Lambda}_1 = \Lambda_1(\overline{\nu})$, $\overline{\Lambda}_2 = \Lambda_2(\overline{\nu})$, where $\overline{\nu}$ denotes the mean frequency of the light. The rest of the ν-integral is, according to (12), precisely $\Gamma(P_1,\, P_2,\, \tau - (s_1 - s_2)/c)$. We thus finally obtain the formula

$$\Gamma(Q_1,\, Q_2,\, \tau) = \int_{\mathcal{A}}\int_{\mathcal{A}} \frac{\Gamma\left(P_1,\, P_2,\, \tau - \dfrac{s_1 - s_2}{c}\right)}{s_1 s_2}\overline{\Lambda}_1\overline{\Lambda}_2^\star\,\mathrm{d}P_1\,\mathrm{d}P_2. \tag{17}$$

This is the required formula for the mutual coherence function at points Q_1 and Q_2 on the surface $\mathcal{B}$ in terms of the mutual intensity at all pairs of points on the surface $\mathcal{A}$.

Of particular interest is the case when the two points Q_1 and Q_2 coincide and $\tau = 0$. Denoting the common point by Q, the left-hand side of (17) reduces to the intensity $I(Q)$. Also, if we substitute for Γ on the right in terms of the intensities and the correlation factor γ, the formula reduces to

$$I(Q) = \int_{\mathcal{A}}\int_{\mathcal{A}} \frac{\sqrt{I(P_1)}\sqrt{I(P_2)}}{s_1 s_2}\gamma\left(P_1,\, P_2,\, \frac{s_2 - s_1}{c}\right)\overline{\Lambda}_1\overline{\Lambda}_2^\star\,\mathrm{d}P_1\,\mathrm{d}P_2. \tag{18}$$

This formula expresses the *intensity* at an arbitrary point Q as the sum of contributions

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

from all pairs of elements of the arbitrary surface $\mathcal{A}$. Each contribution is directly proportional to the geometrical mean of the intensities at the two elements, inversely proportional to the product of their distances from Q, and is weighted by the appropriate value of the correlation factor γ.

The formulae (17) and (18) are generalizations, due to Wolf, of the propagation law of Zernike (§10.4 (45)) and of the formula §10.4 (46) for the intensity in a partially coherent wave field.

10.8 Rigorous theory of partial coherence*

10.8.1 Wave equations for mutual coherence

Some of the theorems relating to the correlation functions, which we derived in the preceding sections are in certain respects similar to theorems relating to the complex disturbance itself. For example, the van Cittert–Zernike formula §10.4 (21) for the complex degree of coherence in a plane illuminated by an extended quasi-monochromatic incoherent source was seen to be identical with a formula relating to the complex disturbance in a diffraction pattern arising from the diffraction by an aperture of the same size and shape as the source. Other examples are the laws for the propagation of the mutual intensity [§10.4 (45)] which was seen to resemble the Huygens–Fresnel principle. Now the results relating to the complex disturbance may be regarded as approximate deductions from certain rigorous theorems, namely the formulae of Helmholtz and Kirchhoff [§8.3 (7), §8.3 (13)], which are a consequence of the fact that the light disturbance is propagated as a wave. This analogy suggests that correlation is also propagated as a wave, and that our theorems are approximate formulations of some associated theorems of the Helmholtz–Kirchhoff type. It is not difficult to show that this in fact is the case.

Consider now a stationary wave field in vacuum and let $V(P_1, t)$ and $V(P_2, t)$ represent the disturbances at points P_1 and P_2 respectively. It will be convenient to begin by expressing the mutual coherence function in the more symmetrical form

$$\begin{aligned}\Gamma(P_1, P_2, t_1, t_2) &= \langle V(P_1, t_1 + t)V^\star(P_2, t_2 + t)\rangle \\ &= \lim_{T\to\infty} \frac{1}{2T}\int_{-\infty}^{\infty} V_T(P_1, t + t_1)V_T^\star(P_2, t + t_2)\mathrm{d}t. \qquad (1)\end{aligned}$$

Further let

$$\nabla_1^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2} \qquad (2)$$

be the Laplacian operator with respect to the Cartesian rectangular coordinates of P_1. On applying this operator to (1) and interchanging the order of the various operations, we obtain

$$\nabla_1^2\Gamma(P_1, P_2, t_1, t_2) = \lim_{T\to\infty} \frac{1}{2T}\int_{-\infty}^{\infty} \{[\nabla_1^2 V_T(P_1, t + t_1)]V_T^\star(P_2, t + t_2)\}\mathrm{d}t. \qquad (3)$$

* The analysis of this section is in the main part based on investigations of E. Wolf, *Proc. Roy. Soc.*, A, **230** (1955), 246; and *Proc. Phys. Soc.*, **71** (1958), 257.

Now the real part $V_T^{(r)}$ of V_T represents the true physical wave field (e.g. a Cartesian component of the electric vector wave) and hence satisfies the wave equation

$$\nabla_1^2 V_T^{(r)}(P_1,\, t+t_1) = \frac{1}{c^2}\frac{\partial^2}{\partial {t_1}^2} V_T^{(r)}(P_1,\, t+t_1). \tag{4a}$$

The imaginary part $V_T^{(i)}$ of V_T, and hence V_T itself, also satisfies the wave equation; this result follows immediately on taking the Hilbert transforms of both sides of (4a) and on using the fact that if two functions are Hilbert transforms of each other so are their derivatives. Hence*

$$\nabla_1^2 V_T(P_1,\, t+t_1) = \frac{1}{c^2}\frac{\partial^2}{\partial {t_1}^2} V_T(P_1,\, t+t_1). \tag{4b}$$

It follows that we may replace ∇_1^2 by $\partial^2/c^2\partial {t_1}^2$ on the right of (3), and we obtain after again changing the order of the operations,

$$\nabla_1^2\Gamma(P_1,\, P_2,\, t_1,\, t_2) = \frac{1}{c^2}\frac{\partial^2}{\partial {t_1}^2}\lim_{T\to\infty}\frac{1}{2T}\int_{-\infty}^{\infty} V_T(P_1,\, t+t_1)V_T^{\star}(P_2,\, t+t_2)\mathrm{d}t,$$

i.e.

$$\nabla_1^2\Gamma(P_1,\, P_2,\, t_1,\, t_2) = \frac{1}{c^2}\frac{\partial^2\Gamma(P_1,\, P_2,\, t_1,\, t_2)}{\partial t_1^2}. \tag{5a}$$

In a strictly similar manner it also follows that

$$\nabla_2^2\Gamma(P_1,\, P_2,\, t_1,\, t_2) = \frac{1}{c^2}\frac{\partial^2\Gamma(P_1,\, P_2,\, t_1,\, t_2)}{\partial t_2^2}, \tag{5b}$$

where ∇_2^2 is the Laplacian operator with respect to the coordinates of the point P_2.

Now for a stationary field Γ depends on t_1 and t_2 only through the difference $t_1 - t_2 = \tau$, and we then may write as before, $\Gamma(P_1,\, P_2,\, t_1,\, t_2) = \Gamma(P_1,\, P_2,\, \tau)$. Then $\partial^2/\partial t_1^2 = \partial^2/\partial t_2^2 = \partial^2/\partial\tau^2$ and we have from (5),

$$\nabla_1^2\Gamma(P_1,\, P_2,\, \tau) = \frac{1}{c^2}\frac{\partial^2\Gamma(P_1,\, P_2,\, \tau)}{\partial\tau^2}, \tag{6a}$$

$$\nabla_2^2\Gamma(P_1,\, P_2,\, \tau) = \frac{1}{c^2}\frac{\partial^2\Gamma(P_1,\, P_2,\, \tau)}{\partial\tau^2}. \tag{6b}$$

We see that *in vacuum the mutual coherence obeys two wave equations*.† Each of them decribes the variation of the mutual coherence when one of the points (P_2 or P_1) is fixed, while the other point and the parameter τ change. Now τ represents a time difference between the instants at which the correlation at the two points is considered,

* The following alternative proof may be noted: since $V_T^{(r)}$ satisfies the wave equation, each of its spectral components $v_T(\nu)$ $(-\infty \leqslant \nu \leqslant \infty)$ satisfies the Helmholtz equation. Now according to §10.2 (18b) the spectrum of $V_T = V_T^{(r)} + \mathrm{i}V_T^{(i)}$ is $2v_T(\nu)$ or 0 according as $\nu \gtrless 0$. Hence each spectral component of V_T also satisfies the Helmholtz equation, and consequently V_T satisfies the wave equation.

† When τ is small compared to the coherence time, we have, according to §10.4 (10), $\Gamma(P_1,\, P_2,\, \tau) \sim J(P_1,\, P_2)\mathrm{e}^{-2\pi\mathrm{i}\bar{\nu}\tau}$. It follows from (6) that within the range of validity of the quasi-monochromatic theory, the mutual intensity J_{12} in vacuum obeys to a good approximation the Helmholtz equations

$$\nabla_1^2 J(P_1,\, P_2) + \bar{k}^2 J(P_1,\, P_2) = 0, \qquad \nabla_2^2 J(P_1,\, P_2) + \bar{k}^2 J(P_1,\, P_2) = 0,$$

where $\bar{k} = 2\pi\bar{\nu}/c$.

and in all experiments enters only in the combination $c\tau = \Delta\mathcal{S}$, i.e. as a path difference. The time itself has thus effectively been eliminated from our final description of the field. This is a particularly attractive feature of the theory of partial coherence, since in optical wave fields true time variations entirely escape detection. The basic entity in this theory, the mutual coherence function $\Gamma(P_1, P_2, \tau)$ is directly measurable, for example by means of the interference experiments described in §10.3 and §10.4.

10.8.2 Rigorous formulation of the propagation law for mutual coherence

We again consider a stationary wave field in vacuum. Let Q_1 and Q_2 be any two points in the field and let $\mathcal{A}$ be any fictitious surface which surrounds these points. If ∇_1^2 denotes the Laplacian operator with respect to the coordinates of Q_1, we have, according to (5a),

$$\nabla_1^2\Gamma(Q_1, Q_2, t_1, t_2) = \frac{1}{c^2}\frac{\partial^2\Gamma(Q_1, Q_2, t_1, t_2)}{\partial t_1^2}. \tag{7}$$

It follows that we may apply to Γ the Kirchhoff integral formula §8.3 (13). Thus $\Gamma(Q_1, Q_2, t_1, t_2)$ may be expressed in terms of the values of $[\Gamma(P_1, Q_2, t_1, t_2)]_1$, where P_1 takes on all positions on $\mathcal{A}$ and $[\ldots]_1$ denotes retardation with respect to the first time argument, i.e.

$$[\Gamma(P_1, Q_2, t_1, t_2)]_1 = \Gamma\left(P_1, Q_2, t_1 - \frac{s_1}{c}, t_2\right) \tag{8}$$

and s_1 is the distance between P_1 and Q_1 (Fig. 10.19). Written out explicitly, Kirchhoff's formula gives

$$\Gamma(Q_1, Q_2, t_1, t_2) = \frac{1}{4\pi}\int_{\mathcal{A}}\left\{f_1[\Gamma(P_1, Q_2, t_1, t_2)]_1 + g_1\left[\frac{\partial}{\partial t_1}\Gamma(P_1, Q_2, t_1, t_2)\right]_1 \right.$$
$$\left. + h_1\left[\frac{\partial}{\partial n_1}\Gamma(P_1, Q_2, t_1, t_2)\right]_1\right\}\mathrm{d}P_1. \tag{9}$$

Here $\partial/\partial n_1$ denotes differentiation along the inward normal to $\mathcal{A}$ at P_1, and

$$f_1 = \frac{\partial}{\partial n_1}\left(\frac{1}{s_1}\right), \qquad g_1 = -\frac{1}{cs_1}\frac{\partial s_1}{\partial n_1}, \qquad h_1 = -\frac{1}{s_1}. \tag{10}$$

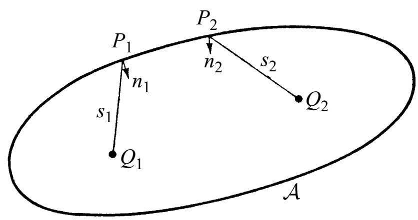


Fig. 10.19 Notation used in the rigorous formulation of the propagation law for mutual coherence.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

Now, according to (5b), we also have

$$\nabla_2^2\Gamma(P_1, Q_2, t_1, t_2) = \frac{1}{c^2}\frac{\partial^2\Gamma(P_1, Q_2, t_1, t_2)}{\partial t_2^2}, \tag{11}$$

where ∇_2^2 is the Laplacian operator with respect to the coordinates of Q_2. Hence $\Gamma(P_1, Q_2, t_1, t_2)$ which appears on the right of (9) may be expressed in the form of a Kirchhoff integral involving the values of $[\Gamma(P_1, P_2, t_1, t_2)]_2$, where P_2 takes on all possible positions on $\mathcal{A}$ and $[\ldots]_2$ denotes retardation with respect to the second time argument, e.g.

$$[\Gamma(P_1, P_2, t_1, t_2)]_2 = \Gamma\left(P_1, P_2, t_1, t_2 - \frac{s_2}{c}\right), \tag{12}$$

and s_2 is the distance between P_2 and Q_2. Written out explicitly, the appropriate formula is

$$\begin{aligned}\Gamma(P_1, Q_2, t_1, t_2) = \frac{1}{4\pi}\int_{\mathcal{A}}\bigg\{ & f_2[\Gamma(P_1, P_2, t_1, t_2)]_2 + g_2\left[\frac{\partial}{\partial t_2}\Gamma(P_1, P_2, t_1, t_2)\right]_2 \\ & + h_2\left[\frac{\partial}{\partial n_2}\Gamma(P_1, P_2, t_1, t_2)\right]_2\bigg\}\mathrm{d}P_2.\end{aligned} \tag{13}$$

Here $\partial/\partial n_2$ denotes differentiation along the inward normal at P_2, and f_2, g_2, h_2 are the same quantities as in (10) but with the suffix 1 replaced by 2. We next differentiate (13) with respect to t_1 and n_1 and obtain

$$\begin{aligned}&\frac{\partial}{\partial t_1}\Gamma(P_1, Q_2, t_1, t_2) \\ &\quad = \frac{1}{4\pi}\int_{\mathcal{A}}\bigg\{f_2\left[\frac{\partial}{\partial t_1}\Gamma(P_1, P_2, t_1, t_2)\right]_2 + g_2\left[\frac{\partial^2}{\partial t_1\partial t_2}\Gamma(P_1, P_2, t_1, t_2)\right]_2 \\ &\quad + h_2\left[\frac{\partial^2}{\partial t_1\partial n_2}\Gamma(P_1, P_2, t_1, t_2)\right]_2\bigg\}\mathrm{d}P_2,\end{aligned} \tag{14}$$

$$\begin{aligned}&\frac{\partial}{\partial n_1}\Gamma(P_1, Q_2, t_1, t_2) \\ &\quad = \frac{1}{4\pi}\int_{\mathcal{A}}\bigg\{f_2\left[\frac{\partial}{\partial n_1}\Gamma(P_1, P_2, t_1, t_2)\right]_2 + g_2\left[\frac{\partial^2}{\partial n_1\partial t_2}\Gamma(P_1, P_2, t_1, t_2)\right]_2 \\ &\quad + h_2\left[\frac{\partial^2}{\partial n_1\partial n_2}\Gamma(P_1, P_2, t_1, t_2)\right]_2\bigg\}\mathrm{d}P_2.\end{aligned} \tag{15}$$

We now substitute from (13), (14) and (15) into (9), and obtain the following expression for $\Gamma(Q_1, Q_2, t_1, t_2)$:

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$\Gamma(Q_1,\,Q_2,\,t_1,\,t_2) = \frac{1}{(4\pi)^2}\int_{\mathcal{A}}\int_{\mathcal{A}}\Bigg\{f_1 f_2[\Gamma]_{1,2} + f_1 g_2\left[\frac{\partial}{\partial t_2}\Gamma\right]_{1,2} + f_1 h_2\left[\frac{\partial}{\partial n_2}\Gamma\right]_{1,2}$$
$$+ g_1 f_2\left[\frac{\partial}{\partial t_1}\Gamma\right]_{1,2} + g_1 g_2\left[\frac{\partial^2}{\partial t_1\partial t_2}\Gamma\right]_{1,2}$$
$$+ g_1 h_2\left[\frac{\partial^2}{\partial t_1\partial n_2}\Gamma\right]_{1,2} + h_1 f_2\left[\frac{\partial}{\partial n_1}\Gamma\right]_{1,2}$$
$$+ h_1 g_2\left[\frac{\partial^2}{\partial n_1\partial t_2}\Gamma\right]_{1,2} + h_1 h_2\left[\frac{\partial^2}{\partial n_1\partial n_2}\Gamma\right]_{1,2}\Bigg\}\mathrm{d}P_1\,\mathrm{d}P_2, \tag{16}$$

where the first two arguments in Γ on the right are P_1 and P_2 and $[\ldots]_{1,2}$ denotes retardation with respect to both the time arguments, e.g.

$$[\Gamma]_{1,2} = \Gamma\left(P_1,\,P_2,\,t_1 - \frac{s_1}{c},\,t_2 - \frac{s_2}{c}\right). \tag{17}$$

Finally, we make use of the assumption of stationarity which ensures that Γ depends on the two time arguments through their difference only. We write as before $\Gamma(P_1,\,P_2,\,t_1,\,t_2) = \Gamma(P_1,\,P_2,\,\tau)$, $\tau = t_1 - t_2$. Then $\partial/\partial t_1 = -\partial/\partial t_2 = \partial/\partial\tau$, and (16) becomes*

$$\Gamma(Q_1,\,Q_2,\,\tau) = \frac{1}{(4\pi)^2}\int_{\mathcal{A}}\int_{\mathcal{A}}\Bigg\{f_1 f_2[\Gamma] - f_1 g_2\left[\frac{\partial}{\partial\tau}\Gamma\right] + f_1 h_2\left[\frac{\partial}{\partial n_2}\Gamma\right]$$
$$+ g_1 f_2\left[\frac{\partial}{\partial\tau}\Gamma\right] - g_1 g_2\left[\frac{\partial^2}{\partial\tau^2}\Gamma\right]$$
$$+ g_1 h_2\left[\frac{\partial^2}{\partial\tau\partial n_2}\Gamma\right] + h_1 f_2\left[\frac{\partial}{\partial n_1}\Gamma\right]$$
$$- h_1 g_2\left[\frac{\partial^2}{\partial n_1\partial\tau}\Gamma\right] + h_1 h_2\left[\frac{\partial^2}{\partial n_1\partial n_2}\Gamma\right]\Bigg\}\mathrm{d}P_1\,\mathrm{d}P_2. \tag{18}$$

The first two arguments in Γ on the right are P_1 and P_2, and $[\ldots]$ denotes 'retardation' by the amount $(s_1 - s_2)/c$, e.g.

$$[\Gamma] = \Gamma\left(P_1,\,P_2,\,\tau - \frac{s_1 - s_2}{c}\right). \tag{19}$$

Formula (18) may be regarded as a rigorous formulation of the law for the propagation of the mutual coherence [§10.7 (17)]. It expresses the value of the mutual coherence function for any two points Q_1 and Q_2 in terms of the values of this function and of some of its derivatives at all pairs of points on an arbitrary closed surface which surrounds both these points.

In the special case when Q_1 and Q_2 coincide and $\tau = 0$, we obtain from (18), on substituting $\Gamma_{12}(\tau) = \sqrt{I_1}\sqrt{I_2}\gamma_{12}(\tau)$, the following expressions for the intensity:

* Eq. (18) applies for propagation from a closed surface $\mathcal{A}$ of arbitrary form. A much simpler formula exists for propagation from a plane surface. [See M. J. Beran and G. B. Parrent, *Theory of Partial Coherence* (Englewood Cliffs, NJ, Prentice Hall, 1964), §3.3.]

$$
\begin{aligned}
I(Q) = \frac{1}{(4\pi)^2}\int_{\mathcal{A}}\int_{\mathcal{A}}\Bigg(&\sqrt{I_1}\sqrt{I_2}\left\{f_1f_2[\gamma] + (f_2g_1 - f_1g_2)\left[\frac{\partial}{\partial\tau}\gamma\right] - g_1g_2\left[\frac{\partial^2}{\partial\tau^2}\gamma\right]\right\} \\
&+ \sqrt{I_1}\left\{f_1h_2\frac{\partial}{\partial n_2}(\sqrt{I_2}[\gamma]) + g_1h_2\frac{\partial}{\partial n_2}\left(\sqrt{I_2}\left[\frac{\partial}{\partial\tau}\gamma\right]\right)\right\} \\
&+ \sqrt{I_2}\left\{f_2h_1\frac{\partial}{\partial n_1}(\sqrt{I_1}[\gamma]) - g_2h_1\frac{\partial}{\partial n_1}\left(\sqrt{I_1}\left[\frac{\partial}{\partial\tau}\gamma\right]\right)\right\} \\
&+ h_1h_2\frac{\partial^2}{\partial n_1\partial n_2}(\sqrt{I_1}\sqrt{I_2}[\gamma])\Bigg)\,\mathrm{d}P_1\,\mathrm{d}P_2, \qquad (20)
\end{aligned}
$$

where I_1 and I_2 are the intensities at P_1 and P_2 respectively, $[\gamma] = \gamma(P_1, P_2, (s_2 - s_1)/c)$, etc. Formula (20) may be regarded as a rigorous formulation of the theorem expressed by §10.7 (18). It gives the intensity at an arbitrary point Q in terms of the distribution of the intensity and of the complex degree of coherence (and of some of the derivatives of these quantities) on an arbitrary surface surrounding Q.

10.8.3 The coherence time and the effective spectral width

The concept of the coherence time, which was found useful in many considerations involving polychromatic light, was introduced in §7.5.8 from the study of the disturbance resulting from a superposition of identical wave trains of finite duration. We showed from a simple example (a random sequence of periodic wave trains) that the coherence time* $\Delta\tau$ and the effective spectral width $\Delta\nu = c\Delta\lambda/\bar{\lambda}^2$ of the resulting disturbance are connected by the order of magnitude relation

$$\Delta\tau\Delta\nu \sim 1. \qquad (21)$$

We also mentioned that a relation of this type holds under more general conditions, provided that $\Delta\tau$ and $\Delta\nu$ are defined as suitable averages. In this section we shall define these quantities and establish the required reciprocity relation rigorously.

Suppose that a beam of light is divided at a point P into two beams which are brought together after the introduction of a path difference $c\tau$ between them. The resulting interference effects are characterized by the self-coherence function

$$\Gamma(\tau) = \langle V(t+\tau)V^\star(t)\rangle = 4\int_0^\infty S(\nu)\mathrm{e}^{-2\pi\mathrm{i}\nu\tau}\,\mathrm{d}\nu, \qquad (22)$$

where $V(t)$ is the complex disturbance at P and $S(\nu)$ is the spectral density.

Since the degree of coherence of the two interfering beams is represented by $|\gamma(\tau)| = |\Gamma(\tau)|/\Gamma(0)$ it is reasonable, and mathematically convenient, to define *the coherence time* $\Delta\tau$ of the light at P as the normalized root-mean-square width (r.m.s.) of the squared modulus of $\Gamma(\tau)$, i.e.†

* To conform to the notation of the present chapter we now write $\Delta\tau$ in place of Δt.

† The average value $\bar{\tau} = \int_{-\infty}^{+\infty}\tau|\Gamma(\tau)|^2\,\mathrm{d}\tau/\int_{-\infty}^{+\infty}|\Gamma(\tau)|^2\,\mathrm{d}\tau$ is zero since $|\Gamma(\tau)|$ is an even function of τ.

For another definition of the coherence time, see L. Mandel, *Proc. Phys. Soc.*, **74** (1959), 233. See also L. Mandel and E. Wolf, *ibid.*, **80** (1962), 894.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$(\Delta\tau)^2 = \frac{\int_{-\infty}^{\infty} \tau^2 |\Gamma(\tau)|^2 \, \mathrm{d}\tau}{\int_{-\infty}^{+\infty} |\Gamma(\tau)|^2 \, \mathrm{d}\tau}. \tag{23}$$

Next we define the *effective spectral width* $\Delta\nu$ of the light at P as the normalized r.m.s. width of the spectrum of Γ, i.e. as the normalized r.m.s. width of the square of the spectral density $S(\nu)$, taken over the range $\nu \geqslant 0$. Thus

$$(\Delta\nu)^2 = \frac{\int_0^{\infty} (\nu - \overline{\nu})^2 S^2(\nu) \mathrm{d}\nu}{\int_0^{\infty} S^2(\nu) \mathrm{d}\nu}, \qquad \overline{\nu} = \frac{\int_0^{\infty} \nu S^2(\nu) \mathrm{d}\nu}{\int_0^{\infty} S^2(\nu) \mathrm{d}\nu}. \tag{24}$$

To establish the required reciprocity relation we set

$$\left.\begin{aligned}
\xi &= \nu - \overline{\nu}, && && \text{(a)}\\
\Phi(\xi) &= 4S(\overline{\nu} + \xi) && \text{when} \quad \xi > -\overline{\nu}, \\
&= 0 && \text{when} \quad \xi < -\overline{\nu}, && \text{(b)}\\
\Psi(\tau) &= \Gamma(\tau) \mathrm{e}^{2\pi \mathrm{i} \overline{\nu} \tau}. && && \text{(c)}
\end{aligned}\right\} \tag{25}$$

We shall assume that $\Phi(\xi)$ is continuous everywhere $(-\infty < \xi < \infty)$; consequently* $\Phi(-\overline{\nu}) = S(0) = 0$. From (22) it follows that Ψ and Φ form a Fourier transform pair,

$$\Psi(\tau) = \int_{-\infty}^{\infty} \Phi(\xi) \mathrm{e}^{-2\pi \mathrm{i} \xi \tau} \, \mathrm{d}\xi, \qquad \Phi(\xi) = \int_{-\infty}^{\infty} \Psi(\tau) \mathrm{e}^{2\pi \mathrm{i} \xi \tau} \, \mathrm{d}\tau. \tag{26}$$

The expressions for $\Delta\tau$ and $\Delta\nu$ become

$$(\Delta\tau)^2 = \frac{1}{N} \int_{-\infty}^{\infty} \tau^2 |\Psi(\tau)|^2 \, \mathrm{d}\tau, \tag{27}$$

$$(\Delta\nu)^2 = \frac{1}{N} \int_{-\infty}^{\infty} \xi^2 \Phi^2(\xi) \mathrm{d}\xi, \tag{28}$$

where

$$N = \int_{-\infty}^{\infty} |\Psi(\tau)|^2 \, \mathrm{d}\tau = \int_{-\infty}^{\infty} \Phi^2(\xi) \mathrm{d}\xi. \tag{29}$$

Next we express the integral (28) in terms of Ψ. We have, on using the second relation in (26),

* This condition is, in fact, necessary for the integral in the numerator of (23) to converge. When it is satisfied, as assumed here, the value of (23) remains unchanged when $\Gamma(\tau) = \langle V(t+\tau)V^{\star}(t)\rangle$ is replaced by the real correlation function $\Gamma^{(r)}(\tau) = \mathcal{R}\Gamma(\tau) = 2\langle V^{(r)}(t+\tau)V^{(r)}(t)\rangle$. (See E. Wolf, *Proc. Phys. Soc.*, B, **71** (1958), 257; R. Silverman, *Trans. Inst. Rad. Engrs.*, CT-5 (1958), 84.) The more general case when $S(0) \neq 0$ was investigated by A. G. Mayer and E. A. Leontovich, *Doklady Akad. Nauk SSSR*, **4** (1934), 353 and by I. Kay and R. Silverman, *Information and Control*, **1** (1957), 64, 396. See also A. A. Kharkevich, *Spectra and Analysis*, translated from Russian (New York, Consultants Bureau, 1960), §12.

$$
\begin{aligned}
(\Delta\nu)^2 &= \frac{1}{N}\int_{-\infty}^{\infty}\xi^2\Phi(\xi)\mathrm{d}\xi\int_{-\infty}^{\infty}\Psi(\tau)\mathrm{e}^{2\pi\mathrm{i}\xi\tau}\,\mathrm{d}\tau \\
&= \frac{1}{N}\int_{-\infty}^{\infty}\Psi(\tau)\mathrm{d}\tau\left(\frac{1}{2\pi i}\right)^2\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\int_{-\infty}^{\infty}\Phi(\xi)\mathrm{e}^{2\pi\mathrm{i}\xi\tau}\,\mathrm{d}\xi \\
&= -\frac{1}{4\pi^2}\frac{1}{N}\int_{-\infty}^{\infty}\Psi(\tau)\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\Psi^\star(\tau)\mathrm{d}\tau \\
&= \frac{1}{4\pi^2}\frac{1}{N}\int_{-\infty}^{\infty}\left|\frac{\mathrm{d}\Psi}{\mathrm{d}\tau}\right|^2\mathrm{d}\tau.
\end{aligned}
\tag{30}
$$

In passing from the second to the third line, the first relation in (26) was used, together with the relation $\Psi(-\tau) = \Psi^\star(\tau)$. The last line follows from the preceding one on integrating by parts and using the fact that $\Psi \to 0$ as $\tau \to \pm\infty$; this is so because the integral $\int_{-\infty}^{\infty}|\Psi(\tau)|^2\,\mathrm{d}\tau$ in (29) is assumed to be convergent.

It follows from (27), (29), and (30) that

$$
(\Delta\tau)^2(\Delta\nu)^2 = \frac{1}{16\pi^2}\left[\frac{4\left(\displaystyle\int_{-\infty}^{\infty}\tau^2|\Psi(\tau)|^2\,\mathrm{d}\tau\right)\left(\displaystyle\int_{-\infty}^{\infty}\left|\frac{\mathrm{d}\Psi}{\mathrm{d}\tau}\right|^2\mathrm{d}\tau\right)}{\left(\displaystyle\int_{-\infty}^{\infty}|\Psi(\tau)|^2\,\mathrm{d}\tau\right)^2}\right]. \tag{31}
$$

Now by a straightforward algebraical argument given in Appendix IX, the term in the large brackets on the right-hand side of (31) is greater than or equal to unity for any function Ψ for which the integrals exist. Hence we have established the following *reciprocity inequality** for the coherence time and the effective spectral width:

$$
\Delta\tau\Delta\nu \geqslant \frac{1}{4\pi}. \tag{32}
$$

We recall that when the light is quasi-monochromatic and the intensities of the two interfering beams are equal, the degree of coherence $|\gamma_{11}(\tau)| = |\Gamma_{11}(\tau)|/\Gamma_{11}(0)$ is according to §10.4 (5) equal to the visibility $\mathcal{V}(\tau)$ of the fringes at a point corresponding to the path difference $c\tau$ between the two beams. Hence (23) may then be written in the form

$$
(\Delta\tau)^2 = \frac{\displaystyle\int_{-\infty}^{\infty}\tau^2\mathcal{V}^2(\tau)\mathrm{d}\tau}{\displaystyle\int_{-\infty}^{\infty}\mathcal{V}^2(\tau)\mathrm{d}\tau} = \frac{\displaystyle\int_{0}^{\infty}\tau^2\mathcal{V}^2(\tau)\mathrm{d}\tau}{\displaystyle\int_{0}^{\infty}\mathcal{V}^2(\tau)\mathrm{d}\tau}; \tag{33}
$$

thus *when the interfering beams are of equal intensity, the coherence time $\Delta\tau$ is equal to the normalized r.m.s. width of the square of the visibility function*.

The present definition of the coherence time is more satisfactory than that given in §7.5.8, for we have now made no special assumptions about the nature of the elementary fields which give rise to the disturbance. In fact we now no longer require

* Our derivation is modelled on that given by H. Weyl and W. Pauli in connection with the Heisenberg uncertainty relation. [H. Weyl, *The Theory of Groups and Quantum Mechanics* (London, Methuen, 1931; also New York, Dover Publications), pp. 77 and 393.]

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

the knowledge of the detailed behaviour of the rapidly fluctuating function $V(t)$, our definition being based on the measurable correlation function $\Gamma(\tau)$. If we wish to retain the description of interference phenomena in terms of elementary wave trains, we may regard $\Delta\tau$ as the duration of an *average* wave train; this interpretation must, however, be employed with caution.

Returning to (32) we see that the equality sign only holds when the term in the large brackets on the right-hand side of (31) is equal to unity, and this, according to Appendix IX, is only possible when $\Psi(\tau)$ is a Gaussian function. Now the Fourier transform of a Gaussian function is again a Gaussian function, and as this function differs from zero for all values of its argument ($-\infty < \xi < \infty$) it does not obey the second condition (25b). Thus the equality sign in (32) never applies. However, when the Gaussian function is centred on a frequency which is large compared to its r.m.s. width, the contribution to $\bar{\nu}$ and $\Delta\nu$ from the negative frequency range is negligible and it is evident that, for the high-frequency spectra encountered in optics, the value of the product $\Delta\tau\Delta\nu$ cannot differ appreciably from that which corresponds to the full Gaussian curve. Thus the inequality in (32) may be replaced by the order of magnitude sign, i.e.

$$\Delta\tau\Delta\nu \sim \frac{1}{4\pi}. \tag{34}$$

The definition of the coherence time just given is appropriate when the two interfering beams are obtained from a single beam by division at a point P. The definition can be extended to situations where the two interfering beams are derived by division at two points P_1 and P_2, for example, in Young's interference experiment. The generalization is straightforward. We have to employ the mutual coherence function $\Gamma_{12}(\tau) = \Gamma(P_1, P_2, \tau)$ in place of the self-coherence function $\Gamma(\tau) = \Gamma(P, P, \tau)$ and the mutual spectral density $G_{12}(\nu)$ in place of the ordinary spectral density $S(\nu)$. The only difference arises from the fact that $G_{12}(\nu)$ is now complex and that $\Gamma_{12}(\tau)$ is no longer necessarily an even function of τ and consequently $\bar{\tau}$ is not necessarily zero. The appropriate definitions are:

$$(\Delta\tau_{12})^2 = \frac{\displaystyle\int_{-\infty}^{\infty} (\tau - \bar{\tau}_{12})^2 |\Gamma_{12}(\tau)|^2 \, \mathrm{d}\tau}{\displaystyle\int_{-\infty}^{\infty} |\Gamma_{12}(\tau)|^2 \, \mathrm{d}\tau}, \qquad \bar{\tau}_{12} = \frac{\displaystyle\int_{-\infty}^{\infty} \tau |\Gamma_{12}(\tau)|^2 \, \mathrm{d}\tau}{\displaystyle\int_{-\infty}^{\infty} |\Gamma_{12}(\tau)|^2 \, \mathrm{d}\tau}, \tag{35}$$

$$(\Delta\nu_{12})^2 = \frac{\displaystyle\int_{0}^{\infty} (\nu - \bar{\nu}_{12})^2 |G_{12}(\nu)|^2 \, \mathrm{d}\nu}{\displaystyle\int_{0}^{\infty} |G_{12}(\nu)|^2 \, \mathrm{d}\nu}, \qquad \bar{\nu}_{12} = \frac{\displaystyle\int_{0}^{\infty} \nu |G_{12}(\nu)|^2 \, \mathrm{d}\nu}{\displaystyle\int_{0}^{\infty} |G_{12}(\nu)|^2 \, \mathrm{d}\nu}. \tag{36}$$

$\Delta\tau_{12}$ may be called the *mutual coherence time* and $\Delta\nu_{12}$ the *mutual effective spectral width* of the light at P_1 and P_2. By an obvious modification of the argument given in connection with $\Gamma_{11}(\tau)$ it follows that these quantities satisfy the reciprocity inequality

$$(\Delta\tau_{12})(\Delta\nu_{12}) > \frac{1}{4\pi}. \tag{37}$$

Finally, for quasi-monochromatic light we now have the following relations as generalizations of (33):

$$(\Delta\tau_{12})^2 = \frac{\displaystyle\int_{-\infty}^{\infty}(\tau - \overline{\tau}_{12})^2\mathcal{V}_{12}^2(\tau)\mathrm{d}\tau}{\displaystyle\int_{-\infty}^{\infty}\mathcal{V}_{12}^2(\tau)\mathrm{d}\tau}, \qquad \overline{\tau}_{12} = \frac{\displaystyle\int_{-\infty}^{\infty}\tau\mathcal{V}_{12}^2(\tau)\mathrm{d}\tau}{\displaystyle\int_{-\infty}^{\infty}\mathcal{V}_{12}^2(\tau)\mathrm{d}\tau}. \tag{38}$$

Here $\mathcal{V}_{12}(\tau)$ represents the visibility of the fringes formed by the light from P_1 and P_2, when the interfering beams have the same intensity.

10.9 Polarization properties of quasi-monochromatic light

In the preceding sections of this chapter we have treated the light disturbance as a scalar quantity. We shall now briefly consider some of the vectorial properties of a quasi-monochromatic light wave.

We have learned in §1.4 that strictly monochromatic light is always *polarized*, i.e. that with increasing time the end point of the electric (and also of the magnetic) vector at each point in space moves periodically around an ellipse, which may, of course, reduce in special cases to a circle or a straight line. We have also encountered *unpolarized* light. In this case the end point may be assumed to move quite irregularly, and the light shows no preferential directional properties when resolved in different directions at right angles to the direction of propagation. Like complete coherence and complete incoherence these two cases represent two extremes. In general the variation of the field vectors is neither completely regular, nor completely irregular, and we may say that the light is *partially polarized*. Such light arises usually from unpolarized light by reflection (see §1.5.3) or scattering (see §14.5.2). In this section we shall investigate the main properties of a partially polarized light wave. We shall see that its observable effects depend on the intensities of any two mutually orthogonal components of the electric vector at right angles to the direction of propagation, and on the correlation which exists between them.

*10.9.1 The coherency matrix of a quasi-monochromatic plane wave**

Consider a quasi-monochromatic light wave of mean frequency $\overline{\nu}$ propagated in the positive z direction. Let

$$E_x(t) = a_1(t)\mathrm{e}^{\mathrm{i}[\phi_1(t)-2\pi\overline{\nu}t]}, \qquad E_y(t) = a_2(t)\mathrm{e}^{\mathrm{i}[\phi_2(t)-2\pi\overline{\nu}t]} \tag{1}$$

represent the components, at a point O, of the electric vector in two mutually orthogonal directions at right angles to the direction of propagation. We again use the complex representation discussed in §10.2, in which E_x and E_y are the 'analytic signals' associated with the true (real) components $E_x^{(r)} = a_1(t)\cos[\phi_1(t) - 2\pi\overline{\nu}t]$, $E_y^{(r)} = a_2(t)\cos[\phi_2(t) - 2\pi\overline{\nu}t]$. If the light were strictly monochromatic, a_1, a_2, ϕ_1 and ϕ_2 would be constants. For a quasi-monochromatic wave these quantities depend also on the time t, but, as we have seen, they change only by small relative amounts in

* The analysis given in §10.9.1 and §10.9.2 is based on investigations of E. Wolf, *Nuovo Cimento*, **13** (1959), 1165. Some further developments are described in a paper by G. B. Parrent and P. Roman, *ibid.*, **15** (1960), 370.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

any time interval that is small compared to the coherence time, i.e. that is small compared to the reciprocal of the effective spectral width $\Delta\nu$ of the light.

Suppose that the y-component is subjected to a retardation ε with respect to the x-component (this can be done, for example, by means of one of the compensators described in §15.4.2), and consider the intensity $I(\theta, \varepsilon)$ of the light vibrations in the direction which makes an angle θ with the positive x-direction (Fig. 10.20). This intensity would be observed by sending the light through a polarizer (§15.4.1) with the appropriate orientation.

The component of the electric vector in the θ direction, after the retardation ε has been introduced, is

$$E(t;\,\theta,\,\varepsilon) = E_x\cos\theta + E_y\,\mathrm{e}^{\mathrm{i}\varepsilon}\sin\theta, \tag{2}$$

so that

$$\begin{aligned} I(\theta,\,\varepsilon) &= \langle E(t;\,\theta,\,\varepsilon)E^{\star}(t;\,\theta,\,\varepsilon)\rangle \\ &= J_{xx}\cos^2\theta + J_{yy}\sin^2\theta + J_{xy}\mathrm{e}^{-\mathrm{i}\varepsilon}\cos\theta\sin\theta + J_{yx}\mathrm{e}^{\mathrm{i}\varepsilon}\sin\theta\cos\theta, \end{aligned} \tag{3}$$

where $J_{xx}, \ldots$ are the elements of the matrix

$$\boldsymbol{J} = \begin{bmatrix} \langle E_xE_x^{\star}\rangle & \langle E_xE_y^{\star}\rangle \\ \langle E_yE_x^{\star}\rangle & \langle E_yE_y^{\star}\rangle \end{bmatrix} = \begin{bmatrix} \langle {a_1}^2\rangle & \langle a_1a_2\mathrm{e}^{\mathrm{i}(\phi_1-\phi_2)}\rangle \\ \langle a_1a_2\mathrm{e}^{-\mathrm{i}(\phi_1-\phi_2)}\rangle & \langle {a_2}^2\rangle \end{bmatrix}. \tag{4}$$

The diagonal elements of $\boldsymbol{J}$ are real and are seen to represent the intensities of the components in the x and y directions. Hence the trace $\mathrm{Tr}\,\boldsymbol{J}$ of the matrix, i.e. the sum of its diagonal elements, is equal to the total intensity of the light,

$$\mathrm{Tr}\,\boldsymbol{J} = J_{xx} + J_{yy} = \langle E_xE_x^{\star}\rangle + \langle E_yE_y^{\star}\rangle. \tag{5}$$

The nondiagonal elements are in general complex, but they are conjugates of each other. (A matrix such as this, which satisfies the relation $J_{ji} = J_{ij}^{\star}$ for all i and j is said to be a *Hermitian* matrix.)

We shall normalize the mixed term J_{xy} in a similar way as before [see §10.4 (9b)], by setting

$$j_{xy} = |j_{xy}|\mathrm{e}^{\mathrm{i}\beta_{xy}} = \frac{J_{xy}}{\sqrt{J_{xx}}\sqrt{J_{yy}}}. \tag{6}$$

Then it follows by Schwarz' inequality, in the same way as in connection with §10.3 (17), that

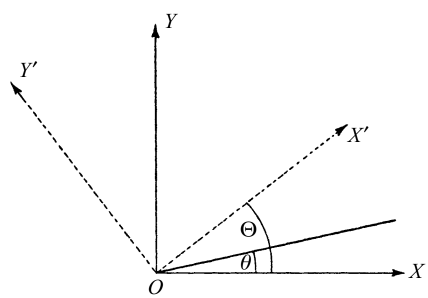


Fig. 10.20 Illustrating notation.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$|j_{xy}| \leqslant 1. \tag{7}$$

This complex correlation factor j_{xy} has a similar significance as the complex degree of coherence j_{12} introduced in §10.4.1. It is a measure of the correlation between the components of the electric vector in the x and y directions. The absolute value $|j_{xy}|$ is a measure of their 'degree of coherence,' and its phase β_{xy} is a measure of their 'effective phase difference.' We call $\boldsymbol{J}$ *the coherency matrix* of the light wave. Since J_{xx} and J_{yy} cannot be negative, (6) and (7) imply that the associated determinant is nonnegative, i.e. that

$$|\boldsymbol{J}| = J_{xx}J_{yy} - J_{xy}J_{yx} \geqslant 0. \tag{8}$$

If we use the relation $J_{yx} = J^{\star}_{xy}$, and denote by $\mathcal{R}$ the real part, (3) becomes

$$\begin{aligned} I(\theta, \varepsilon) &= J_{xx}\cos^2\theta + J_{yy}\sin^2\theta + 2\cos\theta\sin\theta\,\mathcal{R}(J_{xy}\mathrm{e}^{-\mathrm{i}\varepsilon}) \\ &= J_{xx}\cos^2\theta + J_{yy}\sin^2\theta + 2\sqrt{J_{xx}}\sqrt{J_{yy}}\cos\theta\sin\theta|j_{xy}|\cos(\beta_{xy} - \varepsilon), \end{aligned} \tag{9}$$

where we substituted from (6) on going from the first to the second line. If we set $J_{xx}\cos^2\theta = I^{(1)}$, $J_{yy}\sin^2\theta = I^{(2)}$, the last formula becomes identical with the basic interference law §10.4 (11) of quasi-monochromatic wave fields.

Like the coherence functions which we considered earlier, the elements of the coherency matrix of a given wave may be determined by means of relatively simple experiments. This may be done in many different ways. One only needs to measure the intensity for several different values of θ (orientation of the polarizer) and ε (delay introduced by the compensator), and solve the corresponding relations obtained from (3). Let $\{\theta, \varepsilon\}$ denote the measurement corresponding to a particular pair, θ, ε. A convenient set of measurements is the following:*

$$\{0°, 0\}, \{45°, 0\}, \{90°, 0\}, \{135°, 0\}, \left\{45°, \frac{\pi}{2}\right\}, \left\{135°, \frac{\pi}{2}\right\}. \tag{10}$$

It follows from (3) that, in terms of the intensities determined from these six measurements, the elements of the coherency matrix are given by

$$\left.\begin{aligned} J_{xx} &= I(0°, 0), \\ J_{yy} &= I(90°, 0), \\ J_{xy} &= \tfrac{1}{2}\{I(45°, 0) - I(135°, 0)\} + \tfrac{1}{2}\mathrm{i}\left\{I\left(45°, \frac{\pi}{2}\right) - I\left(135°, \frac{\pi}{2}\right)\right\}, \\ J_{yx} &= \tfrac{1}{2}\{I(45°, 0) - I(135°, 0)\} - \tfrac{1}{2}\mathrm{i}\left\{I\left(45°, \frac{\pi}{2}\right) - I\left(135°, \frac{\pi}{2}\right)\right\}. \end{aligned}\right\} \tag{11}$$

We see that only a polarizer is needed to determine J_{xx}, J_{yy} and the real part of J_{xy} (or J_{yx}). J_{xx} and J_{yy} may be determined from measurements with a polarizer oriented so as to transmit the components in the azimuths $\theta = 0$ and $\theta = 90°$ respectively. The real part of J_{xy} may be determined from measurements with a polarizer oriented so that it first transmits the component in the azimuth $\theta = 45°$ and then the component in the azimuth $\theta = 135°$. To determine the imaginary part of J_{xy} (or J_{yx}) we also need,

* In fact, only four measurements are required to determine the four Stokes parameters, but the set represented by (10) is experimentally convenient.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

according to the last two relations in (11), a compensator which introduces a phase difference of a quarter period (e.g. a quarter-wave plate, see §15.4.2) between the x- and y-components; the polarizer is again used, oriented so as to pass first the component in the azimuth $\theta = 45°$ and then the component in the azimuth $\theta = 135°$. We shall learn in §15.4.2 that the last two measurements are those made to detect right-handed and left-handed circular polarization.

It is evident from (9) that two beams of light which have the same coherency matrix are equivalent in the sense that they will yield the same (time averaged) intensity in similar experiments with a polarizer and a compensator.*

Let us now examine how the observed intensity $I(\theta, \varepsilon)$ changes for a given wave when one of the arguments (θ or ε) is kept fixed, whilst the other varies. Suppose first that we keep θ fixed and change ε. We see from (9) that the intensity varies sinusoidally between the values

$$\left.\begin{aligned} I_{\max(\varepsilon)} &= J_{xx}\cos^2\theta + J_{yy}\sin^2\theta + 2|J_{xy}|\sin\theta\cos\theta \\ \text{and}\qquad & \\ I_{\min(\varepsilon)} &= J_{xx}\cos^2\theta + J_{yy}\sin^2\theta - 2|J_{xy}|\sin\theta\cos\theta. \end{aligned}\right\} \tag{12}$$

Hence

$$\frac{I_{\max(\varepsilon)} - I_{\min(\varepsilon)}}{I_{\max(\varepsilon)} + I_{\min(\varepsilon)}} = \frac{|J_{xy}|\sin 2\theta}{J_{xx}\cos^2\theta + J_{yy}\sin^2\theta}. \tag{13}$$

Eq. (13) indicates an alternative way of determining the absolute value of J_{xy} (and hence also of $|j_{xy}|$); it shows that this quantity may be obtained from measurements of J_{xx}, J_{yy}, $I_{\max(\varepsilon)}$ and $I_{\min(\varepsilon)}$; the phase of J_{xy} may be obtained from measurement of the value of ε at which the maxima or minima occur. For, according to (9),

$$\left.\begin{aligned} I &= I_{\max(\varepsilon)}, \quad \text{when} \quad \varepsilon = \beta_{xy} \pm 2m\pi \qquad (m = 0, 1, 2, \ldots), \\ I &= I_{\min(\varepsilon)}, \quad \text{when} \quad \varepsilon = \beta_{xy} \pm (2m+1)\pi \quad (m = 0, 1, 2, \ldots). \end{aligned}\right\} \tag{14}$$

To see how the intensity changes when ε is fixed and θ is varied, it is convenient to re-write (9) in a somewhat different form. Simple calculation gives

$$I(\theta, \varepsilon) = \tfrac{1}{2}(J_{xx} + J_{yy}) + R\cos(2\theta - \alpha), \tag{15}$$

where

$$\begin{aligned} R &= \tfrac{1}{2}\sqrt{(J_{xx} - J_{yy})^2 + 4J_{xy}J_{yx}\cos^2(\beta_{xy} - \varepsilon)}, \\ \tan\alpha &= \frac{2|J_{xy}|\cos(\beta_{xy} - \varepsilon)}{J_{xx} - J_{yy}}. \end{aligned} \tag{16}$$

* This statement is true only within the approximation of the quasi-monochromatic theory, for it is only within the range of validity of this theory that the expression (9) for the intensity holds. The two beams may behave quite differently when the phase delay ε introduced between the two orthogonal components is not negligible in comparison with the coherence length, measured in units of the mean wavelength $\bar{\lambda}$. For a fuller description of the observable properties of a beam it is necessary to introduce more general coherency matrices which characterize the correlations between the components at different times and also at different points. See E. Wolf, *Nuovo Cimento*, **12** (1954), 884; also his contribution in *Proc. Symp. Astronom. Optics*, ed. Z. Kopal (Amsterdam, North-Holland Publishing Co., 1956), p. 177; P. Roman and E. Wolf, *Nuovo Cimento*, **17** (1960), 462, 477; P. Roman, *ibid.*, **20** (1961), 759; *ibid.*, **22** (1961), 1005.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

We see from (15) that, as θ changes, the intensity again varies sinusoidally; its extremes are

$$\left.\begin{aligned} I_{\max(\theta)} &= \tfrac{1}{2}(J_{xx} + J_{yy}) + R, \\ I_{\min(\theta)} &= \tfrac{1}{2}(J_{xx} + J_{yy}) - R. \end{aligned}\right\} \tag{17}$$

On the right of (17) only R depends on ε. It takes its largest value when $\cos^2(\beta_{xy} - \varepsilon) = 1$, i.e. when ε has one of the values given by (14), and is then equal to

$$\begin{aligned} R_{\max(\varepsilon)} &= \tfrac{1}{2}\sqrt{(J_{xx} - J_{yy})^2 + 4J_{xy}J_{yx}} \\ &= \tfrac{1}{2}(J_{xx} + J_{yy})\sqrt{1 - \frac{4|\boldsymbol{J}|}{(J_{xx} + J_{yy})^2}}, \end{aligned} \tag{18}$$

where $|\boldsymbol{J}|$ is the determinant (8) of the coherency matrix. It follows that the absolute maxima and minima (with respect to both θ and ε) of the intensity are

$$\left.\begin{aligned} I_{\max(\theta,\varepsilon)} &= \tfrac{1}{2}(J_{xx} + J_{yy})\left[1 + \sqrt{1 - \frac{4|\boldsymbol{J}|}{(J_{xx} + J_{yy})^2}}\right], \\ I_{\min(\theta,\varepsilon)} &= \tfrac{1}{2}(J_{xx} + J_{yy})\left[1 - \sqrt{1 - \frac{4|\boldsymbol{J}|}{(J_{xx} + J_{yy})^2}}\right]. \end{aligned}\right\} \tag{19}$$

Hence

$$\frac{I_{\max(\theta,\varepsilon)} - I_{\min(\theta,\varepsilon)}}{I_{\max(\theta,\varepsilon)} + I_{\min(\theta,\varepsilon)}} = \sqrt{1 - \frac{4|\boldsymbol{J}|}{(J_{xx} + J_{yy})^2}}. \tag{20}$$

We shall see later that this quantity has a simple physical meaning.

So far we have referred the electric vibrations to arbitrary but fixed rectangular axes OX, OY. We shall now consider how the coherency matrix transforms when a new set of axes is chosen. Suppose that we take new rectangular axes OX', OY', again in the plane perpendicular to the direction of propagation, such that OX' makes an angle Θ with OX (see Fig. 10.20). In terms of E_x, E_y, the components of the electric vector referred to the new axes are

$$\left.\begin{aligned} E_{x'} &= \quad E_x \cos\Theta + E_y \sin\Theta, \\ E_{y'} &= -E_x \sin\Theta + E_y \cos\Theta. \end{aligned}\right\} \tag{21}$$

The elements of the transformed coherency matrix $\boldsymbol{J}'$ are

$$J_{k'l'} = \langle E_{k'} E^{\star}_{l'} \rangle, \tag{22}$$

where k' and l' each take on the values x' and y'. From (21) and (22) it follows that

$$\boldsymbol{J}' = \begin{bmatrix} J_{xx}c^2 + J_{yy}s^2 + (J_{xy} + J_{yx})cs & (J_{yy} - J_{xx})cs + J_{xy}c^2 - J_{yx}s^2 \\ (J_{yy} - J_{xx})cs + J_{yx}c^2 - J_{xy}s^2 & J_{xx}s^2 + J_{yy}c^2 - (J_{xy} + J_{yx})cs \end{bmatrix}, \tag{23}$$

where

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$c = \cos\Theta, \qquad s = \sin\Theta. \tag{24}$$

It is seen that the trace of the matrix is invariant under rotation of the axes. A straightforward calculation shows that its determinant is also invariant under this transformation. Both these results also follow from well-known theorems of matrix algebra.

We shall now consider the form of the coherency matrix for some cases of particular interest.

(a) Completely unpolarized light (natural light)

Light which is most frequently encountered in nature has the property that the intensity of its components in any direction perpendicular to the direction of propagation is the same; and, moreover, this intensity is not affected by any previous retardation of one of the rectangular components relative to the other into which the light may have been resolved. In other words

$$I(\theta, \varepsilon) = \text{constant} \tag{25}$$

for all values of θ and ε. Such light may be said to be *completely unpolarized* and is often also called *natural light*.

It is evident from (9) that $I(\theta, \varepsilon)$ is independent of ε and θ if and only if

$$j_{xy} = 0 \quad \text{and} \quad J_{xx} = J_{yy}. \tag{26a}$$

The first condition implies that E_x and E_y are mutually incoherent. According to (6) and the relation $J_{yx} = J^{\star}_{xy}$, (26a) may also be written as

$$J_{xy} = J_{yx} = 0, \qquad J_{xx} = J_{yy} \tag{26b}$$

and it follows that the coherency matrix of natural light of intensity $J_{xx} + J_{yy} = I_0$ is

$$\tfrac{1}{2}I_0\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{27}$$

(b) Completely polarized light

Suppose first that the light is strictly monochromatic. Then the amplitudes a_1 and a_2, and the phase factors ϕ_1 and ϕ_2 in (1) do not depend on the time, and the coherency matrix has the form

$$\begin{bmatrix} a_1^2 & a_1 a_2 e^{i\delta} \\ a_1 a_2 e^{-i\delta} & a_2^2 \end{bmatrix}, \tag{28}$$

where*

$$\delta = \phi_1 - \phi_2. \tag{29}$$

* For the purpose of later applications of some results of §1.4 we note that ϕ_1 and ϕ_2 correspond to $-\delta_1$ and $-\delta_2$ of §1.4.2, so that (29) is consistent with the earlier definition §1.4 (16), viz. $\delta = \delta_2 - \delta_1$.

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

We see that in this case

$$|\boldsymbol{J}| = J_{xx}J_{yy} - J_{xy}J_{yx} = 0, \tag{30}$$

i.e. the determinant of the coherency matrix is zero. The complex degree of coherence of the components E_x and E_y now is

$$j_{xy} = \frac{J_{xy}}{\sqrt{J_{xx}}\sqrt{J_{yy}}} = \mathrm{e}^{\mathrm{i}\delta}, \tag{31}$$

i.e. its absolute value is unity (complete coherence) and its phase is equal to the difference between the phases of the two components.

In the special case when the light is *linearly polarized* we have [see §1.4 (33)], $\delta = m\pi$ ($m = 0, \pm 1, \pm 2, \ldots$). Hence the coherency matrix of linearly polarized light is

$$\begin{bmatrix} a_1^2 & (-1)^m a_1 a_2 \\ (-1)^m a_1 a_2 & a_2^2 \end{bmatrix}. \tag{32}$$

The electric vector vibrates in the direction given by $E_y/E_x = (-1)^m a_2/a_1$. In particular, the matrices

$$I\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad I\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \tag{33}$$

each represent linearly polarized light of intensity I, with the electric vector in the x direction ($a_2 = 0$) and the y direction ($a_1 = 0$) respectively; and the matrices

$$\tfrac{1}{2}I\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad \tfrac{1}{2}I\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{34}$$

each represent linearly-polarized light of intensity I and with the electric vector in directions making angles 45° and 135° with the x direction respectively ($a_1 = a_2$, $m = 0$ and $a_1 = a_2$, $m = 1$).

For *circularly-polarized light* we have [see §1.4 (35), §1.4 (36)], $a_1 = a_2$, $\delta = m\pi/2$ ($m = \pm 1, \pm 3, \ldots$), so that the coherency matrix is

$$\tfrac{1}{2}I\begin{bmatrix} 1 & \pm\mathrm{i} \\ \mp\mathrm{i} & 1 \end{bmatrix}, \tag{35}$$

where I is the intensity of the light. By §1.4 (38) and §1.4 (40) the upper or lower sign is taken according to whether the polarization is right- or left-handed.

Condition (30) may also be satisfied when the light is not monochromatic. For if a_1, a_2, ϕ_1 and ϕ_2 depend on time in such a way that the *ratio of the amplitudes* and the *difference in the phases* are time-independent, so that

$$\frac{a_2(t)}{a_1(t)} = q, \qquad \delta = \phi_1(t) - \phi_2(t) = \chi, \tag{36}$$

where q and χ are constants, then

$$\begin{aligned} J_{xx} &= \langle a_1^2\rangle, & J_{xy} &= q\langle a_1^2\rangle \mathrm{e}^{\mathrm{i}\chi}, \\ J_{yx} &= q\langle a_1^2\rangle \mathrm{e}^{-\mathrm{i}\chi}, & J_{yy} &= q^2\langle a_1^2\rangle, \end{aligned} \tag{37}$$

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

and the condition (30) holds. The coherency matrix with the elements (37) is the same as that of strictly monochromatic light with components

$$E_x = \sqrt{\langle a_1^2 \rangle} \mathrm{e}^{\mathrm{i}(\alpha - 2\pi\bar{\nu}t)}, \qquad E_y = q\sqrt{\langle a_1^2 \rangle} \mathrm{e}^{\mathrm{i}(-\chi + \alpha - 2\pi\bar{\nu}t)}, \tag{38}$$

where α is any real constant. It follows that in experiments involving a polarizer and a compensator, the quasi-monochromatic wave which obeys the conditions (36) will behave in exactly the same way as the strictly monochromatic and hence completely polarized wave (38). (It is assumed, of course, that the phase difference introduced by the compensator is small compared to the coherence length of the light, measured in units of the mean wavelength.) Condition (30) may, therefore, be said to characterize a *completely polarized* light wave.

10.9.2 Some equivalent representations. The degree of polarization of a light wave

If several *independent* light waves which are propagated in the same direction are superposed, the coherency matrix of the resulting wave is equal to the sum of the coherency matrices of the individual waves. To prove this result let $E_x^{(n)}$, $E_y^{(n)}$ $(n = 1, 2, \ldots, N)$ be the components of the electric vectors (in the usual complex representation) of the individual waves. The components of the electric vector of the resulting wave are

$$E_x = \sum_{n=1}^{N} E_x^{(n)}, \qquad E_y = \sum_{n=1}^{N} E_y^{(n)}, \tag{39}$$

so that the elements of its coherency matrix are given by

$$\begin{aligned} J_{kl} = \langle E_k E_l^\star \rangle &= \sum_{n=1}^{N} \sum_{m=1}^{N} \langle E_k^{(n)} E_l^{(m)\star} \rangle \\ &= \sum_{n} \langle E_k^{(n)} E_l^{(n)\star} \rangle + \sum_{n \neq m} \langle E_k^{(n)} E_l^{(m)\star} \rangle. \end{aligned} \tag{40}$$

Since the waves are assumed to be independent, each term under the last summation sign is zero, and it follows that

$$J_{kl} = \sum_{n} J_{kl}^{(n)}, \tag{41}$$

where $J_{kl}^{(n)} = \langle E_k^{(n)} E_l^{(n)\star} \rangle$ are the elements of the coherency matrix of the nth wave. Eq. (41) shows that the coherency matrix of the combined wave is equal to the sum of the coherency matrices of all the separate waves.

Conversely any wave may be regarded as the sum of independent waves, which evidently may be chosen in many different ways. One particular choice is of special significance and will now be briefly considered.

We will show that any quasi-monochromatic light wave may be regarded as the sum of a completely unpolarized and a completely polarized wave, which are independent of each other, and that this representation is unique.

To establish this result it is only necessary to show that any coherency matrix $\boldsymbol{J}$ can be uniquely expressed in the form

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

$$\boldsymbol{J} = \boldsymbol{J}^{(1)} + \boldsymbol{J}^{(2)}, \tag{42}$$

where, in accordance with (27) and (30),

$$\boldsymbol{J}^{(1)} = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}, \qquad \boldsymbol{J}^{(2)} = \begin{bmatrix} B & D \\ D^\star & C \end{bmatrix}, \tag{43}$$

with $A \geqslant 0$, $B \geqslant 0$, $C \geqslant 0$ and

$$BC - DD^\star = 0. \tag{44}$$

If J_{xx}, J_{xy}, ... are the elements of the coherency matrix which characterizes the original wave, we must have, according to (42) and (43),

$$\left.\begin{aligned} A + B = J_{xx}, \qquad D = J_{xy}, \\ D^\star = J_{yx}, \quad A + C = J_{yy}. \end{aligned}\right\} \tag{45}$$

On substituting from (45) to (44) we obtain the following equation for A:

$$(J_{xx} - A)(J_{yy} - A) - J_{xy}J_{yx} = 0; \tag{46}$$

thus A is a characteristic root (eigenvalue) of the coherency matrix $\boldsymbol{J}$. The two roots of (46) are

$$A = \tfrac{1}{2}(J_{xx} + J_{yy}) \pm \tfrac{1}{2}\sqrt{(J_{xx} + J_{yy})^2 - 4|\boldsymbol{J}|}, \tag{47}$$

where, as before, $|\boldsymbol{J}|$ is the determinant (8). Since $J_{yx} = J_{xy}^\star$, the product $J_{xy}J_{yx}$ is nonnegative and it follows from (8) that

$$|\boldsymbol{J}| \leqslant J_{xx}J_{yy} \leqslant \tfrac{1}{4}(J_{xx} + J_{yy})^2,$$

so that both the roots (47) are real and nonnegative. Consider first the solution with the negative sign in front of the square root. We then have

$$A = \tfrac{1}{2}(J_{xx} + J_{yy}) - \tfrac{1}{2}\sqrt{(J_{xx} + J_{yy})^2 - 4|\boldsymbol{J}|}, \tag{48}$$

$$\left.\begin{aligned} B = \tfrac{1}{2}(J_{xx} - J_{yy}) + \tfrac{1}{2}\sqrt{(J_{xx} + J_{yy})^2 - 4|\boldsymbol{J}|}, \qquad D = J_{xy}, \\ D^\star = J_{yx}, \qquad C = \tfrac{1}{2}(J_{yy} - J_{xx}) + \tfrac{1}{2}\sqrt{(J_{xx} + J_{yy})^2 - 4|\boldsymbol{J}|}. \end{aligned}\right\} \tag{49}$$

Now

$$\sqrt{(J_{xx} + J_{yy})^2 - 4|\boldsymbol{J}|} = \sqrt{(J_{xx} - J_{yy})^2 + 4J_{xy}J_{yx}} \geqslant |J_{xx} - J_{yy}|.$$

Hence B and C are also nonnegative, as required. The other root (47) (with the positive sign in front of the square root) leads to negative values of B and C and must therefore be rejected. We have thus obtained a unique decomposition of the required kind.

The total intensity of the wave is

$$I_{\text{tot}} = \operatorname{Tr} \boldsymbol{J} = J_{xx} + J_{yy}; \tag{50}$$

and the total intensity of the polarized part is

$$I_{\text{pol}} = \operatorname{Tr} \boldsymbol{J}^{(2)} = B + C = \sqrt{(J_{xx} + J_{yy})^2 - 4|\boldsymbol{J}|}. \tag{51}$$

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

The ratio of the intensity of the polarized portion to the total intensity is called the *degree of polarization* P of the wave; according to (50) and (51) it is given by

$$P = \frac{I_{\text{pol}}}{I_{\text{tot}}} = \sqrt{1 - \frac{4|\boldsymbol{J}|}{(J_{xx} + J_{yy})^2}}. \tag{52}$$

Since this expression involves only the two rotational invariants of the coherency matrix $\boldsymbol{J}$, the degree of polarization is independent of the particular choice of the axes OX, OY, as might have been expected. From (52) and the inequality preceding (48) it follows that

$$0 \leqslant P \leqslant 1. \tag{53}$$

When $P = 1$ there is no unpolarized component, so that the wave is then *completely polarized*. In this case $|\boldsymbol{J}| = 0$, so that $|j_{xy}| = 1$ and consequently E_x and E_y are mutually coherent. When $P = 0$ the polarized component is absent. The wave is then *completely unpolarized*. In this case $(J_{xx} + J_{yy})^2 = 4|\boldsymbol{J}|$, i.e.

$$(J_{xx} - J_{yy})^2 + 4J_{xy}J_{yx} = 0. \tag{54a}$$

Since $J_{yx} = J^\star_{xy}$ we have the sum of two squares equal to zero, and this is only possible if each vanishes separately, i.e. if

$$J_{xx} = J_{yy} \quad \text{and} \quad J_{xy} = J_{yx} = 0, \tag{54b}$$

in accordance with (26b). E_x and E_y are then mutually incoherent ($j_{xy} = 0$). In all other cases ($0 < P < 1$) we say that the light is *partially polarized*. Comparison of (52) and (20) shows that the quantity $(I_{\max(\theta,\varepsilon)} - I_{\min(\theta,\varepsilon)})/(I_{\max(\theta,\varepsilon)} + I_{\min(\theta,\varepsilon)})$ is precisely the degree of polarization P.

When E_x and E_y are mutually incoherent (but the light not necessarily natural), the expression for the degree of polarization takes a simple form. Since now $J_{xy} = J_{yx} = 0$, $|\boldsymbol{J}| = J_{xx}J_{yy}$ and (52) reduces to*

$$P = \left| \frac{J_{xx} - J_{yy}}{J_{xx} + J_{yy}} \right|. \tag{55}$$

This expression is in agreement with the formula §1.5 (42) employed in connection with polarization of natural light by reflection.

We note some useful representations of *natural light*. The coherency matrix (27) of natural light may always be expressed in the form

$$\tfrac{1}{2}I\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \tfrac{1}{2}I\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \tfrac{1}{2}I\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \tag{56}$$

* Since every Hermitian matrix may be made diagonal by a unitary transformation, and since the values of $|\boldsymbol{J}|$ and $\operatorname{Tr}\boldsymbol{J}$ remain invariant under this transformation, the degree of polarization may always be expressed in the form

$$P = \left| \frac{A_1 - A_2}{A_1 + A_2} \right|,$$

where A_1 and A_2 are the two eigenvalues [given by (47)]. However, the unitary transformation does not, in general, represent a real rotation of axes about the direction of the propagation of the wave.

It is of interest to note that the eigenvalues A_1 and A_2 are equal to the values $I_{\max(\theta,\varepsilon)}$ and $I_{\min(\theta,\varepsilon)}$ given by (19).

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

and this implies, according to (33), that a wave of natural light, of intensity I, is equivalent to two independent linearly polarized waves, each of intensity $\frac{1}{2}I$, with their electric vectors vibrating in two mutually perpendicular directions at right angles to the direction of propagation.

Another useful representation of natural light is

$$\tfrac{1}{2}I\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} = \tfrac{1}{4}I\begin{bmatrix}1 & +\mathrm{i}\\ -\mathrm{i} & 1\end{bmatrix} + \tfrac{1}{4}I\begin{bmatrix}1 & -\mathrm{i}\\ +\mathrm{i} & 1\end{bmatrix}, \tag{57}$$

and implies, according to (35), that a wave of natural light of intensity I is equivalent to two independent circularly polarized waves, one right-handed, the other left-handed, each of intensity $\frac{1}{2}I$.

Returning to the general case (partially polarized light), it is to be noted that unlike the degree of polarization, P, the degree of coherence $|j_{xy}|$ depends on the choice of the x and y directions. One may, however, readily see that $|j_{xy}|$ cannot exceed P. For if in (52) we write out the determinant $|\boldsymbol{J}|$ in full and use (6) we find that

$$1 - P^2 = \frac{J_{xx}J_{yy}}{[\frac{1}{2}(J_{xx}+J_{yy})]^2}[1 - |j_{xy}|^2]. \tag{58}$$

Since the geometric mean of any two positive numbers cannot exceed their arithmetic mean, it follows that $1 - P^2 \leqslant 1 - |j_{xy}|^2$, i.e.

$$P \geqslant |j_{xy}|. \tag{59}$$

The equality sign in (59) will hold if and only if $J_{xx} = J_{yy}$, i.e. if the (time averaged) intensities associated with the x and y directions are equal. We will show now that a pair of directions always exists for which this is the case.

If the x- and y-axes are rotated in their own plane, through an angle θ in the anticlockwise sense, J_{xx} and J_{yy} are transformed into $J_{x'x'}$ and $J_{y'y'}$ respectively, where, according to (23)

$$\left.\begin{aligned} J_{x'x'} &= J_{xx}\cos^2\theta + J_{yy}\sin^2\theta + (J_{xy}+J_{yx})\cos\theta\sin\theta,\\ J_{y'y'} &= J_{xx}\sin^2\theta + J_{yy}\cos^2\theta - (J_{xy}+J_{yx})\cos\theta\sin\theta.\end{aligned}\right\} \tag{60}$$

From (60) it follows that $J_{x'x'} = J_{y'y'}$, if the axes are rotated through the angle $\theta = \Theta$, where

$$\tan 2\Theta = \frac{J_{yy} - J_{xx}}{J_{xy} + J_{yx}}. \tag{61}$$

Since $J_{yx} = J_{xy}^{\star}$ and J_{xx} and J_{yy} are both real, (61) always has a real solution for Θ. Thus *there always exists a pair of mutually orthogonal directions for which the intensities are equal. For this pair of directions the degree of coherence $|j_{xy}|$ of the electric vibrations has its maximum value and this value is equal to the degree of polarization P of the wave.**

* The geometrical significance of this special pair of directions is discussed by E. Wolf in *Nuovo Cimento*, **13** (1959), 1180–1181.

10.9.3 The Stokes parameters of a quasi-monochromatic plane wave

We have seen that, in order to characterize a quasi-monochromatic plane wave, four real quantities are in general necessary, for example J_{xx}, J_{yy} and the real and imaginary parts of J_{xy} (or J_{yx}). In his investigations relating to partially polarized light, Stokes* introduced a somewhat different four-parameter representation which is closely related to the present one. We have already encountered it, in a restricted form, in connection with monochromatic light in §1.4.2. The general Stokes parameters are the four quantities

$$\left.\begin{aligned} s_0 &= \langle a_1^2 \rangle + \langle a_2^2 \rangle, \\ s_1 &= \langle a_1^2 \rangle - \langle a_2^2 \rangle, \\ s_2 &= 2\langle a_1 a_2 \cos\delta \rangle, \\ s_3 &= 2\langle a_1 a_2 \sin\delta \rangle, \end{aligned}\right\} \tag{62}$$

where, as before, a_1 and a_2 are the instantaneous amplitudes of the two orthogonal components E_x, E_y of the electric vector and $\delta = \phi_1 - \phi_2$ is their phase difference. When the light is monochromatic, a_1, a_2 and δ are independent of the time, and (62) reduces to the 'monochromatic Stokes parameters' defined in §1.4 (43).

It follows from (62) and (4) that the Stokes parameters and the elements of the coherency matrix are related by the formulae

$$\left.\begin{aligned} s_0 &= J_{xx} + J_{yy}, \\ s_1 &= J_{xx} - J_{yy}, \\ s_2 &= J_{xy} + J_{yx}, \\ s_3 &= \mathrm{i}(J_{yx} - J_{xy}); \end{aligned}\right\} \quad (63\text{a}) \qquad \left.\begin{aligned} J_{xx} &= \tfrac{1}{2}(s_0 + s_1), \\ J_{yy} &= \tfrac{1}{2}(s_0 - s_1), \\ J_{xy} &= \tfrac{1}{2}(s_2 + \mathrm{i}s_3), \\ J_{yx} &= \tfrac{1}{2}(s_2 - \mathrm{i}s_3). \end{aligned}\right\} \quad (63\text{b})$$

Like the elements of the coherency matrix, the Stokes parameters of any quasi-monochromatic plane wave may be determined from simple experiments. If as before $I(\theta, \varepsilon)$ denotes the intensity of the light vibrations in the direction making an angle θ with OX, when the y-component is subjected to a retardation ε with respect to the x-component, then, according to (11) and the relations (63a),

* G. G. Stokes, *Trans. Cambr. Phil. Soc.*, **9** (1852), 399. Reprinted in his *Mathematical and Physical Papers*, Vol. III (Cambridge, Cambridge University Press, 1901), p. 233. See also P. Soleillet, *Ann. de Physique* (10), **12** (1929), 23; F. Perrin, *J. Chem. Phys.*, **10** (1942), 415; S. Chandrasekhar, *Radiative Transfer* (Oxford, Clarendon Press, 1950), §15; M. J. Walker, *Amer. J. Phys.*, **22** (1954), 170; E. Wolf, *Nuovo Cimento*, **12** (1954), 884; S. Pancharatnam, *Proc. Ind. Acad. Sci.*, A, **44** (1956), 398; *ibid.*, **57** (1963), 218, 231.

The Stokes parameters are also employed in quantum mechanical treatments of polarization of elementary particles. See U. Fano, *J. Opt. Soc. Amer.*, **39** (1949), 859; *ibid*, **41** (1951), 58; *Phys. Rev.*, **93** (1954), 121; D. L. Falkoff and J. E. MacDonald, *J. Opt. Soc. Amer.*, **41** (1951), 861; W. H. McMaster, *Amer. J. Phys.*, **22** (1954), 351; J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons* (Cambridge, MA, Addison-Wesley Publ. Co., 1955), §2.8. See also N. Wiener, *Acta Math.*, **55** (1930), §9, especially pp. 189–192.

$$\left.\begin{aligned} s_0 &= I(0°, 0) + I(90°, 0), \\ s_1 &= I(0°, 0) - I(90°, 0), \\ s_2 &= I(45°, 0) - I(135°, 0), \\ s_3 &= I\left(45°, \frac{\pi}{2}\right) - I\left(135°, \frac{\pi}{2}\right). \end{aligned}\right\} \tag{64}$$

The parameter s_0 evidently represents the total intensity. The parameter s_1 is equal to the excess in intensity of light transmitted by a polarizer which accepts linear polarization in the azimuth $\theta = 0°$, over the light transmitted by a polarizer which accepts linear polarization in the azimuth $\theta = 90°$. The parameter s_2 has a similar interpretation with respect to the azimuths $\theta = 45°$ and $\theta = 135°$. Finally, the parameter s_3 is equal to the excess in intensity of light transmitted by a device which accepts right-handed circular polarization, over that transmitted by a device which accepts left-handed circular polarization.

If we use the relations (63b) our previous results may be expressed in terms of the Stokes parameters, rather than in terms of the coherency matrix. In particular, the condition (8), viz. $J_{xx}J_{yy} - J_{xy}J_{yx} \geqslant 0$, becomes

$$s_0{}^2 \geqslant s_1^2 + s_2^2 + s_3^2. \tag{65}$$

For monochromatic light we have, according to (30), $J_{xx}J_{yy} - J_{xy}J_{yx} = 0$ and the equality sign then holds in (65), in agreement with §1.4 (44).

Let us now consider the decomposition of a given wave into an unpolarized and a polarized portion which are mutually independent, using the Stokes parameter representation. It follows from (41) and (63) that the Stokes parameters of a mixture of *independent* waves are sums of the respective Stokes parameters of the separate waves. From (27) and (63a) it follows that an unpolarized wave (wave of natural light) is characterized by $s_1 = s_2 = s_3 = 0$. Denoting by a single symbol $\boldsymbol{s}$ the four Stokes parameters s_0, s_1, s_2, s_3, the required decomposition of the wave characterized by $\boldsymbol{s}$ evidently is

$$\boldsymbol{s} = \boldsymbol{s}^{(1)} + \boldsymbol{s}^{(2)}, \tag{66}$$

where

$$\boldsymbol{s}^{(1)} = s_0 - \sqrt{s_1^2 + s_2^2 + s_3^2},\ 0,\ 0,\ 0, \tag{67a}$$

$$\boldsymbol{s}^{(2)} = \sqrt{s_1^2 + s_2^2 + s_3^2},\ s_1,\ s_2,\ s_3. \tag{67b}$$

$\boldsymbol{s}^{(1)}$ represents the unpolarized and $\boldsymbol{s}^{(2)}$ the polarized part. Hence, in terms of the Stokes parameters, the degree of polarization of the original wave is

$$P = \frac{I_{\text{pol}}}{I_{\text{tot}}} = \frac{\sqrt{s_1^2 + s_2^2 + s_3^2}}{s_0}, \tag{68}$$

as may also be verified by substituting from (63b) into (52). We may also easily write down expressions which give the ellipticity and the orientation of the polarization ellipse associated with the polarized part (67b). If, as in §1.4 (28),

$$\tan\chi = \mp b/a, \qquad (-\pi/4 < \chi \leqslant \pi/4)$$

https://doi.org/10.1017/9781108769914.013 Published online by Cambridge University Press

represents the ratio of the minor and the major axes and the sense in which the ellipse is described ($\chi \gtrless 0$ according as the polarization is right- or left-handed), then according to (67b) and §1.4 (45c),

$$\sin 2\chi = \frac{s_3}{\sqrt{s_1^2 + s_2^2 + s_3^2}}; \tag{69}$$

and the angle ψ $(0 \leqslant \psi < \pi)$, which the major axis makes with OX, is according to (67b) and §1.4 (46) given by

$$\tan 2\psi = \frac{s_2}{s_1}. \tag{70}$$

We see that the Stokes parameters, just like the coherency matrix, provide a useful tool for a systematic analysis of the state of polarization of a quasi-monochromatic wave.