5

Coherence and Fringe Localization

T.D. Milster and N.A. Beaudry

A basic functional form for an electromagnetic disturbance from a laser can be described as a simple plane wave traveling in the **z** direction

$$E(z, t) = E_0 e^{j(kz - \omega t)} , \tag{5.1}$$

where the disturbance is known over all time and at every location in space. In Eq. (5.1), $k = 2\pi/\lambda$ is the propagation constant, ω is the radian frequency and t is time. Many theoretical formulations in physical optics assume this ideal form of a monochromatic electromagnetic wave. In fact, much insight is gained and many phenomena are adequately described using ideal waves. While an ideal wave is a close approximation to a very good laser beam, it is *never* a completely true representation of nature. Real wave fields exhibit phases, frequencies and wavelengths that vary randomly with respect to time. The random nature may be a very small fraction of the average frequency, as in a high-quality laser beam, or it may be much more significant, like the field emitted from a tungsten light bulb. Due to this random nature, the statistical description of light plays an extremely important role in determining the behavior of many optical systems.

While it is clear that some form of statistical description is necessary, in most cases a complete statistical description is not required, and a second order model is sufficient[5.1]. The second order average of the statistical properties of an optical field is often referred to as a *coherence function*, and the field of study known as *coherence theory* is simply the study of the coherence function. An extremely large amount of information is derived from these second order

statistics, where the most useful piece of information is the ability of the optical field to interfere.

As a simple example, consider the interference pattern created when a screen is placed in an optical field $E(\boldsymbol{r}, t)$, as shown in Fig. 5.1. The screen has two small pinholes that transmit light, but it is otherwise opaque. The pinholes are spaced by a distance equal to d. Cartesian vectors $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ describe the positions and distances of the pinholes from a point on the source. Pinholes P1 and P2 are small enough so that they act as point sources on transmission, and an interference pattern is produced on the observation screen along y_o. The interference pattern consists of bright and dark fringes, where the irradiance of the total field is the measurable quantity. The optical path length between P1 and y_o is vt_1, where v is wave velocity in the observation space and t_1 is the time required for the wave to travel from P1 to y_o. Similarly, the optical path length between P2 and y_o is vt_2. The characteristics of the fringe pattern on the observation screen depends on the type of source used to create field $E(\boldsymbol{r}, t)$.

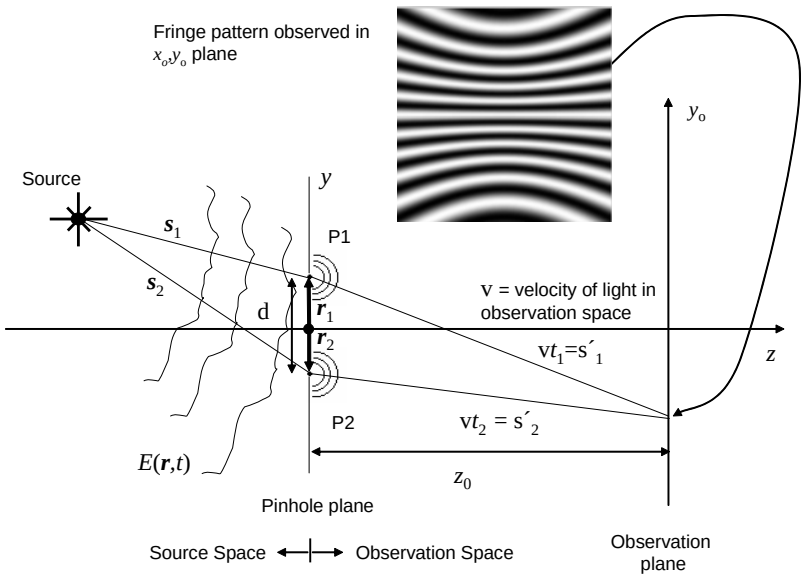


Figure 5.1. A fringe pattern is produced when a light source illuminates two pinholes in an opaque screen. Properties of the fringe pattern depend on the type and location of the source.

The most dramatic characteristic of the fringe pattern that changes as a result of the type of source is the fringe contrast. The fringe contrast is measured locally, and it is given by

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} , \tag{5.2}$$

where I_{max} and I_{min} are the maximum and minimum irradiances of the local fringe modulation *V*, as shown in Fig. 5.2. A high-contrast fringe pattern has $V \approx 1$, where I_{min} is much smaller than I_{max}. Fringes are clearly observed for $V > 0.5$, while it is difficult to discern the bright and dark portions of a fringe for $V < 0.2$. Note that $V = 0$ does not imply zero local irradiance, but it does imply that the irradiance in the region is void of modulation along y_o due to path difference $OPD' = v(t_1 - t_2)$.

If the source of the optical field $E(\boldsymbol{r}, t)$ is a high quality laser beam, the two pinholes in Fig. 5.1 produce high-contrast fringes along y_o. The analysis of this problem is identical to the analysis of the fields radiating from two coherent point sources, where a hyperboloidal fringe field is generated at the observation plane. If the laser beam illuminates the pinholes symmetrically (lengths $s_1 = s_2$ in Fig. 5.1), the fringes are hyperboloids that simplify to straight, equally spaced fringes along the central region of y_o when $d \ll z_0$, as shown in Fig. 5.2(A).

If the source is extended but nearly monochromatic, like light from a highly filtered light bulb, the fringe pattern exhibits a significant reduction in contrast, as shown in Fig 5.2(B). The contrast is inversely proportional to the source size. That is, as the source size increases, contrast reduces. Notice that the fringe contrast is not a function of y_o. The behavior of contrast reduction as a function of source size is a result of the effect known as *spatial coherence*.

If the source used to create field $E(\boldsymbol{r}, t)$ is a centered polychromatic point source, like a distant arc lamp, fringe contrast varies as a function of y_o as shown in Fig. 5.2(C). Notice that fringe contrast is higher near $y_o = 0$ and decreases as y_o increases. Behavior of fringe contrast as a function of the wavelength distribution is the result of the effect known as *temporal coherence.*

In this chapter, the characteristics of both temporal and spatial coherence are explored. Also, derivations for basic coherence formulas are provided, which usually result in simple Fourier transform relationships. The chapter is divided into three sections. Section 5.1 is an introduction to coherence theory with a simplified development. Section 5.2 discusses the topic of *fringe localization,* which concerns the spatial location of high-visibility fringes in different interferometers. Section 5.3 presents an advanced mathematical development at coherence theory. For further reading, suggested references are [5.2,5.3, and 5.4].

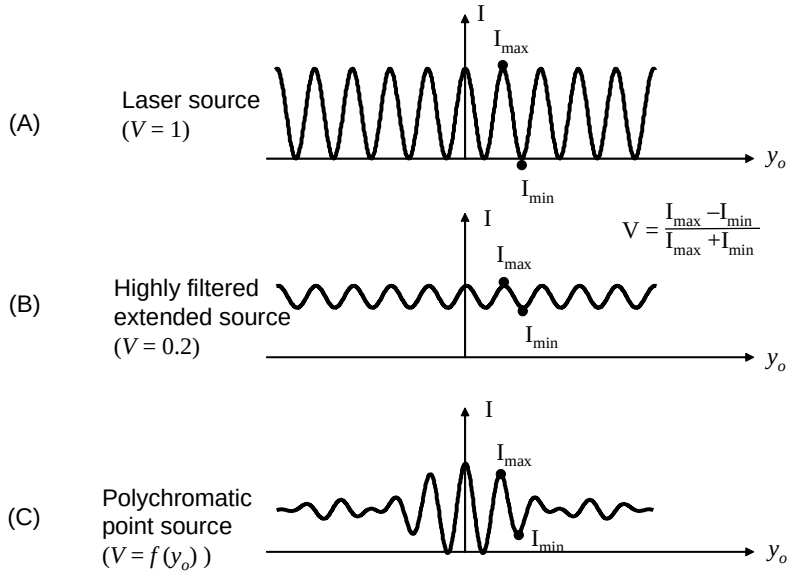


Figure 5.2. Line profiles of the fringe pattern irradiance along the y_o direction for different types of sources. A) Coherent laser source with visibility $V = 1$; B) Highly filtered/extended (narrow wavelength) source with $V = 0.2$; and C) Polychromatic point source, where V is a function of y_o. Profile B is an example of *spatial coherence*, where the visibility depends on the geometry of the source, but not the observation-space location. Profile C, where V is a function of observation-space coordinate y_o, is an example of *temporal coherence* effects.

5.1 Basic coherence

5.1.1 The mutual coherence function

Mathematically, the instantaneous electric field arriving at some position y_o in the observation plane is given by

$$E_{total}(\boldsymbol{r}_1, \boldsymbol{r}_2, t, t_1, t_2) = \sqrt{K_1}E(\boldsymbol{r}_1, t + t_1) + \sqrt{K_2}E(\boldsymbol{r}_2, t + t_2) \ , \qquad (5.3)$$

where K_1 and K_2 are constants that describe the fraction of energy transmitted by each pinhole to the observation screen, and t_1 and t_2 are transit times required for the optical wave to travel from the pinholes to the observation point. Irradiance of the pattern on the observation screen is proportional to

$$I(\boldsymbol{r}_1, \boldsymbol{r}_2; t_1, t_2) \propto K_1 \langle |E(\boldsymbol{r}_1, t + t_1)|^2 \rangle_t + K_2 \langle |E(\boldsymbol{r}_2, t + t_2)|^2 \rangle_t + \sqrt{K_1 K_2} Re[\langle E(\boldsymbol{r}_1, t + t_1) E^*(\boldsymbol{r}_2, t + t_2) \rangle_t] \quad , \tag{5.4}$$

where $\langle\ \rangle_t$ denotes a time average. The subscript t that represents a time average is suppressed throughout the remainder of the chapter and plain angle brackets $\langle\ \rangle$ imply a time average. Notice that the first two terms in Eq. (5.4) are simply irradiances at the observation point from each pinhole separately. The third term modifies the total irradiance and is called the *cross term* or *interference term*. If we assume that the constants K_1 and K_2 do not vary appreciably over the observation region, it is the cross term that, when at negative maximum, determines I_{min}. When maximum positive, the cross term determines I_{max}. Therefore, the cross term is a primary factor in determining the fringe visibility.

The primary quantitative expression in the cross term of Eq. (5.4) is the *mutual coherence function*. Expressed in its most general scalar form, the mutual coherence function is given by

$$\Gamma(\boldsymbol{r}_1, \boldsymbol{r}_2, t_1, t_2) = \langle E(\boldsymbol{r}_1, t + t_1) E^*(\boldsymbol{r}_2, t + t_2) \rangle \ . \tag{5.5}$$

The mutual coherence function is a statistical property that, expressed in words, is the temporal correlation of the electric field at two positions in space with respect to time. In the example of Fig. 5.1, Cartesian position vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ correspond to the pinhole locations.

Consider the case of *stationary fields*. That is, stationary implies that the statistics of the optical field are not affected by a shift in the time origin.[5.5,5.6] Stationary neglects any transient response that occurs when the source is turned on and assumes the average power of the source is not fluctuating on time scales comparable to the temporal period of the optical field or the measurement time. In this case, statistics of the optical field only depend on a time difference. That is, $t + t_1$ and $t + t_2$ in Eq. (5.5) are replaced by t and $t + \Delta t$, respectively. The mutual coherence function becomes

$$\Gamma(\boldsymbol{r}_1, \boldsymbol{r}_2, \Delta t) = \langle E(\boldsymbol{r}_1, t) E^*(\boldsymbol{r}_1, t + \Delta t) \rangle \ . \tag{5.6}$$

The time difference Δt corresponds to the *optical path difference* $OPD' = v\Delta t$ in the observation space of the pinhole interferometer. In Fig. 5.1, the OPD' is a function of y_o.[1] The interferometer in Fig. 5.1 is called a Young's double pinhole interferometer (YDPI). It is used throughout this chapter to refine concepts of coherence.

---

1. In this case, the prime on variable OPD' indicates that it is a function of the observation region that is part of the *observation space*. This prime notation will be used to differentiate similar variables in both the source space and the observation space. The source-space OPD is desctibed in section 5.1.5.

The general YDPI model can be applied to many classes of interferometers where the fringe contrast is of interest. That is, every interferometer exhibits a source space that produces field $E(\mathbf{r},t)$. In a *wavefront division interferometer*, like the Lloyd's mirror interfermeter shown in Fig. 5.3, $E(\mathbf{r},t)$ is sampled at two points $(\mathbf{r_1},\mathbf{r_2})$, and the contrast is measured as a function of OPD'. In an *amplitude division interferometer*, like the Twyman-Green interferemeter shown in Fig, 5.4, images of $E(\mathbf{r},t)$ are combined with an induced OPD', like that generated from reflection off a test mirror or transmission through a lens. Each point in the source field $E(\mathbf{r},t)$ interferes with a time-shifted version of itself, where the time shift Δt, and, hence, OPD', is a function of $\mathbf{r}$. The measurement of OPD' as a function of position y_o is of interest. (An example of a Twyman-Green interferometer is presented in Fig 5.4.)

A Lloyd's mirror wavefront-division interferomter uses diverging light from a point source and it's image reflectd from a flat mirror. In this case, two different portions of wavefront E(r,t) combined to form the interference pattern determined by $OPD' = S'_1 = S'_2$.

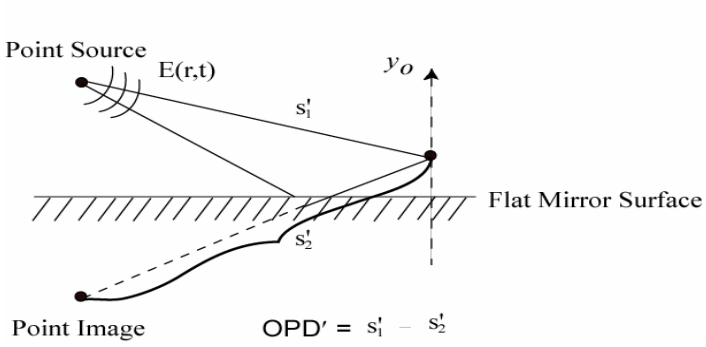
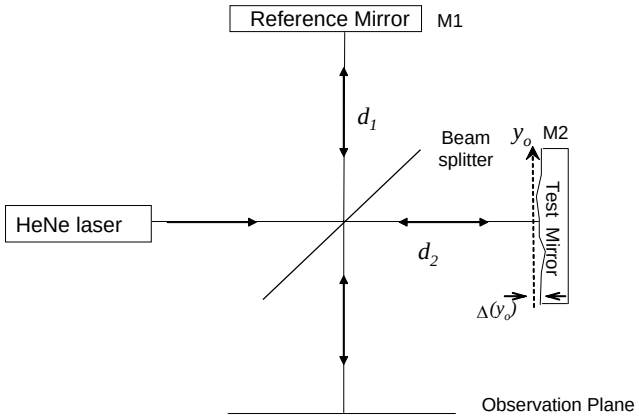


Figure 5.3. A Lloyd's mirror wavefront-division interferometer uses diverging light from a point source and its image reflected from a flat mirror. In this case, two different portions of wavefront E(r,t) combined to form the interference pattern determined by $OPD' = S_1' = S_2'$.



$$OPD' = 2(d_1 - d_2) - 2\Delta(y_o)$$

Figure 5.4. A Twyman-Green amplitude division interferometer uses a collimated laser beam that is divided by a beam splitter. Each split beam reflects off a mirror and recombines agter passing through the splitter (beam 1) and reflecting from the splitter (beam 2). The two beams form an interference pattern in the observation plane that is determined from OPD'.

5.1.2 The two-wavelength point source

The next step up from an ideal monochromatic wave is a source that contains two wavelengths. Consider an implementation of the Fig. 5.1 two-pinhole screen that is shown in Fig. 5.5 with an on-axis point source (lengths $s_1 = s_2$). The source emits spherical waves with wavelengths λ_a and λ_b. Since the point source is exactly halfway between the two pinholes, light emitted from each pinhole is exactly in phase with light emitted from the second pinhole. Pinholes are separated by distance *d*.

Light from the source arriving at each pinhole exhibits a high frequency modulation term and envelope modulation as described in Chapter 2. Media on both sides of the pinhole screen are non-dispersive, so the high-frequency and modulation components exhibit the same velocity, *v.* Therefore, behaviors of waves emitted by P1 and P2 and observed at y_o have fixed temporal relationships. If $K_1 = K_2 = 1$ and amplitudes of each component wave have the value *A*, wave components at the observation point y_o are

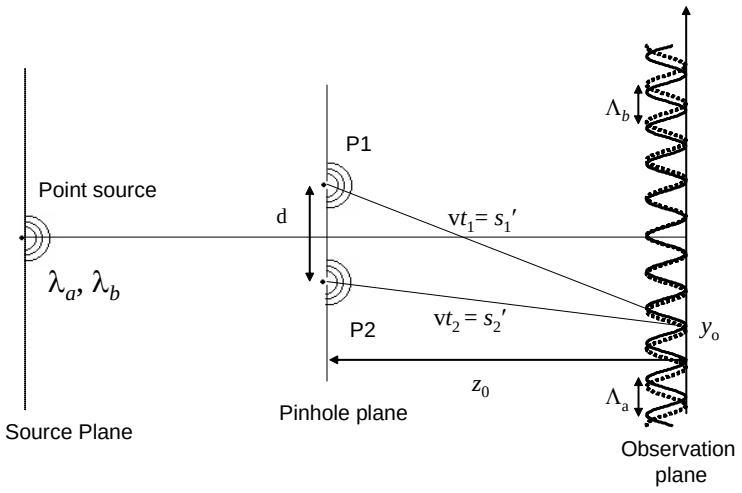


Figure 5.5. An on-axis dual wavelength source in a YDPI produces two independent fringe patterns in the observation plane with different periods Λ_a and Λ_b. The irradiances of the independent fringe patterns are added to produce the total irradiance.

$$
\begin{aligned}
E_{1a} &= A\cos(k_a s_1'(-\omega_a t) + \varphi_a)\,, \\
E_{1b} &= A\cos(k_b s_1' - \omega_b t + \varphi_b)\,, \\
E_{2a} &= A\cos(k_a s_1' - \omega_a(t + \Delta t) + \varphi_a)\,, \\
&\textit{and} \\
E_{2b} &= A\cos(k_b s_1' - \omega_b(t + \Delta t) + \varphi_b)\,,
\end{aligned}
\tag{5.7}
$$

where the first two terms result from P1 and the second two terms result from P2. Subscripts *a* and *b* refer to the first and second wavelengths, respectively. When the total field from the addition of components in Eq. (5.7) is squared, the result is

Addition of field components: (5.8)

$$|E_{total}|^2 = 2A^2 + 2A^2\cos(\omega_a \Delta t) \tag{5.8a}$$

$$+ 2A^2 + 2A^2\cos(\omega_b \Delta t) \tag{5.8b}$$

$$+ 2A^2\cos(2k_m s_1' - 2\omega_m t + 2\beta) \tag{5.8c}$$

$$+ 2A^2\cos(2k_m s_1' - 2\omega_m t + 2\beta + \omega_b \Delta t) \tag{5.8d}$$

$$+ 2A^2\cos(2k_m s_1' - 2\omega_m t + 2\beta + \omega_a \Delta t) \tag{5.8e}$$

$$+ 2A^2 \cos(2k_m s_1' - 2\omega_m t + 2\beta + \omega_m \Delta t) \ , \qquad (5.8f)$$

where terms (5.8a) and (5.8b) simply result from the two wavelengths independently, and terms (5.8c) through (5.8f) are inter-modulation beat terms. The *m* subscripts in Eq. (5.8) imply modulation frequency ω_m or modulation wave number k_m (that is, $\omega_m = (\omega_a - \omega_b)/2$ and $k_m = (k_a - k_b)/2$). The constant β is the difference in the phase terms of the original optical fields, that is, $\beta = (\varphi_a - \varphi_b)/2$. Notice that arguments of the cosines in Eqs. (5.8a) and (5.8b) do not depend on absolute time *t*. They depend only on wavelength, insomuch as wavelength is related to optical frequency by $\lambda = 2\pi c/\omega$, and Δt, which is a function of y_o. Conversely, inter-modulation phase terms in Eqs. (5.8c) through (5.8f) are functions of absolute time *t*.

In order to find irradiance in the observation plane, a time average is performed over Eq. (5.8). A general relationship for averaging the sum of variables is

$$\langle C_1 + C_2 + C_3 + \ldots \rangle = \langle C_1 \rangle + \langle C_2 \rangle + \langle C_3 \rangle + \ldots \ . \qquad (5.9)$$

Application of Eq. (5.9) to Eq. (5.8) results in the first two terms Eqs. (5.8a) and (5.8f) remaining unchanged, because they are not functions of absolute time *t*. For example, averaging Eq. (5.8a) yields

$$\langle 2A^2 \rangle + \langle 2A^2 \cos(\omega_a \Delta t) \rangle = 2A^2 + 2A^2 \cos(\omega_a \Delta t) \ . \qquad (5.10)$$

A different result is obtained when time averaging is performed on the remaining terms in Eq. (5.8). Since the cosine arguments are functions of absolute time, integration over the full temporal period ($n\lambda_{eq}/c$) of the modulation yields a zero net result. For example, averaging Eq. (5.8c) yields

$$2A^2 \langle \cos(2k_m s'_1 - 2\omega_m t + 2\beta) \rangle = 0 \ . \qquad (5.11)$$

Therefore, the result of time averaging Eq. (5.8) is

$$\begin{aligned} \langle |E_{total}|^2 \rangle &= 2A^2 + 2A^2 \cos(\omega_a \Delta t) + 2A^2 + 2A^2 \cos(\omega_b \Delta t) \\ &= 4A^2 [2 + \cos(\omega_a \Delta t) + \cos(\omega_b \Delta t)] \end{aligned} \ . \qquad (5.12)$$

That is, the irradiance pattern in the observation plane is effectively the *sum of irradiance patterns from individual wavelengths.*

Physically, the time average represents the integration time of the detector used to measure the energy at observation point y_o. If the integration time $T \gg \pi/\omega_m$, inter-modulation terms in Eq. (5.8) average to zero. For example, if λ_a = 550 nm and λ_b = 570 nm, $T \gg 5.2 \times 10^{-14}$ seconds in order to satisfy the integration-time requirement. Even several orders of magnitude (say, T=10^{-10} seconds) beyond this value is well below the temporal response characterics of all but the most sophisticated physical detectors, so it is usually a safe assumption to

approximate the total irradiance by adding individual irradiance patterns produced from individual wavelengths.

In the limit where $\lambda_a \to \lambda_b$, $\omega_a \approx \omega_b = \omega$, Eq. (5.8) reduces (before averaging) to

$$\left|E_{total}\right|^2 = 4A^2(1 + \cos\beta)(1 + \cos\omega\Delta t) \ . \tag{5.13}$$

The leading factor $(1 + \cos\beta)$ defines coherent addition of the wavelengths based on their phase difference. No phase shift with $\beta = 0$ produces a coherent sum, which is equivalent to the result obtained by simply increasing the source amplitude by a factor of two.

If the wavelengths are derived from physically distinct phenomena, like the beams from two laser cavities, the phase β is random in time. Time averaging of Eq. (5.13) with a zero-near β results in

$$\begin{aligned} \langle\left|E_{total}\right|^2\rangle &= 4A^2(1 + \langle\cos\beta\rangle)(1 + \cos\omega\Delta t) \\ &= 4A^2(1 + \cos\omega\Delta t) \end{aligned} \ , \tag{5.14}$$

because the time average of $\langle\cos\beta\rangle = 0$. Again, the total irradiance pattern of Eq. (5.8) is simply the addition of two irradiance patterns produced separately from the individual wavelengths, which both have equal amplitudes in this case. This argument can be expanded to apply to any two wavelengths in the spectrum of a polychromatic source.[1]

In the remaining sections of this chapter, total observation plane irradiance is derived by summing irradiance patterns resulting from each wavelength of the source. Distance Λ along y_o between adjacent fringe maxima corresponds to a change of one wavelength of optical path difference $OPD' = v\Delta t$. Equation (5.12) is written in these terms as

$$\begin{aligned} \langle\left|E_{total}\right|^2\rangle &= 2A^2[2 + \cos(k_a OPD') + \cos(k_b OPD')] \\ &\approx 2A^2\left[2 + \cos\left(\frac{2\pi}{\bar{\lambda}}OPD'\right)\cos\left(\frac{\pi}{\lambda_{eq}}OPD'\right)\right] \end{aligned} \ , \tag{5.15}$$

where $\lambda_{eq} = \lambda_a\lambda_b/\left|\lambda_a - \lambda_b\right|$ and $\bar{\lambda} = (\lambda_a + \lambda_b)/2$. Note that the cosine terms on the top right side of Eq. (5.15) have slightly different spatial periods Λ_a and Λ_b for the same range of OPD', due to the different wavelengths. The cosine terms on the bottom right side of Eq. (5.15) indicates a slow modulation corresponding to λeq. The higer-frequency fringes occur with period $OPD' = \bar{\lambda}$.

The central region of the observation screen along y_o, which is important in the development of coherence functions in following sections, is defined as

1. In fact, it is possible to phase two laser beams such that the beat frequency is only 1 Hz. However, this laboratory experiment is very complicated and is not representative of naturally occurring sources.

$y_o \ll z_0$, with the constraint that $d \ll z_0$. As shown in Fig. 5.6, $OPD' = s'_1 - s'_2 \approx d\sin\theta \approx dy_o/z_0$, where lengths $s'_1 = vt_1$ and $s'_2 = vt_2$ define distances from the observation point to P1 and P2, respectively. Under t hese conditions, OPD' is a linear function of position along y_o. The period between bright fringes is $\Lambda = z_0\lambda/d$. The cosine terms in Eq. (5.15) produce straight and equally spaced fringes in the central region for wavelengths λ_a and λ_b with periods Λ_a and Λ_b, respectively, as shown in Fig. 5.5.

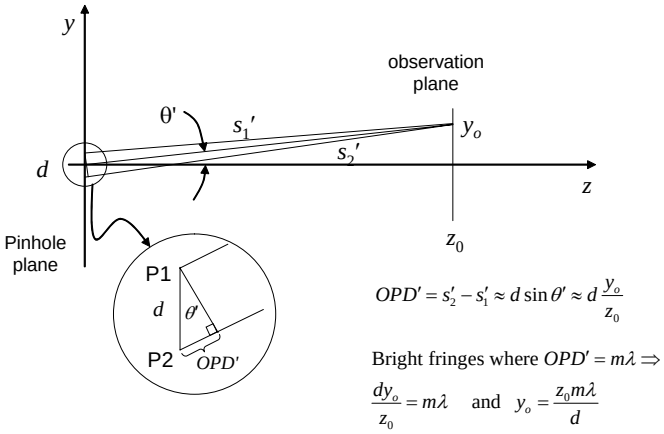


Figure 5.6. When observation is made with $|y_o|$ small, $d << z_o$ and $zo >> \lambda$ the optical path difference in observation space OPD' can be approximated by $OPD' \approx dy_o/z_o$. Fringes are straight and equally spaced.

Figure 5.7 shows a gray-scale representation of fringes in the $x_o y_o$ observation plane for a two-wavelength source. The observation range is restricted near $y_o = 0$. Notice that the fringe modulation washes out periodically along y_o. Figure 5.8 shows line profiles of the component fringes, the combined fringe pattern and the fringe visibility curve as functions of $OPD' = dy_o/z_o$. In the exact center of the pattern where $y_o = 0$, $\Delta t = 0$ and $OPD' = 0$. The cosine terms are in phase, and $V = 1$. As y_o increases away from zero, fringes become out of phase, and visibility reduces. Fringes are π out of phase and $V = 0$ when $m\lambda_a = (n + 1/2)\lambda_b = OPD'$, where m and n are integers. The relationship between m and n at this OPD' is

$$\frac{\Delta\lambda}{\lambda} = \frac{n - m + 1/2}{m}. \tag{5.16}$$

The example in Figs. 5.7 and 5.8 exhibits $m = 5$ and $n = 4$. Notice that the visibility is a smooth envelope that decays to zero at the point where $y_o = z_0\lambda_{eq}/d$. This distance between the maximum fringe contrast at $y_o = 0$ and the first zero of fringe visibility is the *coherence length*. Coherence length can be defined as a physical distance y_o or in terms of OPD'. Beyond the point of minimum visibility, fringes become more in phase, until $V = 1$ again. Notice that the combined pattern in Fig.5.8(b) has a periodic variation in visibility, between V = 0 and V = 1, but the average irradiance is constant. The high-frequency fringes occur with period $\bar{\lambda}$ in OPD' and change phase between modulation lobes. The visibility curve in Fig.5.8(c) is periodic and corresponds to the modulation term in Eq. (5.15). The coherence length is λ_{eq}, which is approximately $\bar{\lambda}^2/\Delta\lambda$, in OPD' units.

As the wavelength distribution of the source widens, the range of good visibility along the observation plane decreases. For example, Fig. 5.9 shows the result of adding a third wavelength, and Fig. 5.10 shows the result for five wavelengths.

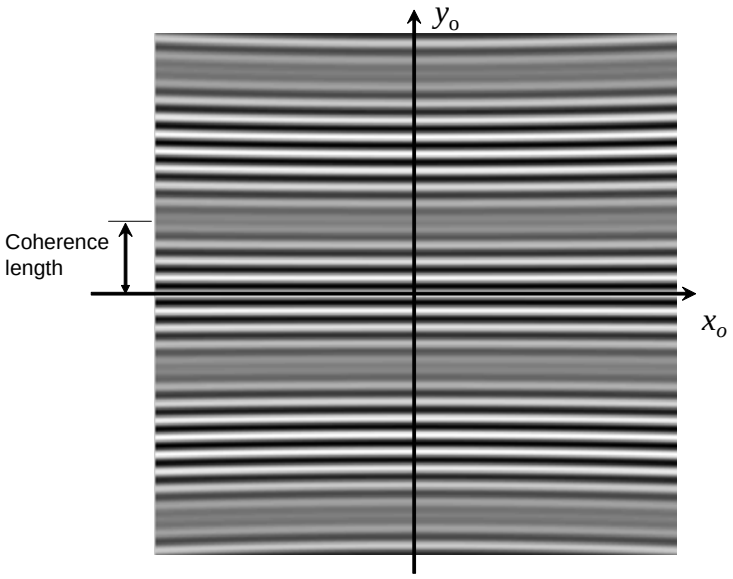


Figure 5.7. Fringes in the YDPI observation plane for small y_o with a dual-wavelength on-axis soure. Fringes are straight and equally spaced . Regions of low *V* are easily observed. Notice that low *V* regions contain significiant irradiance, but they do not contain clear modulation. The coherence length is defined as the distance from the central fringe with maximum *V* to the first region of minimum *V.*

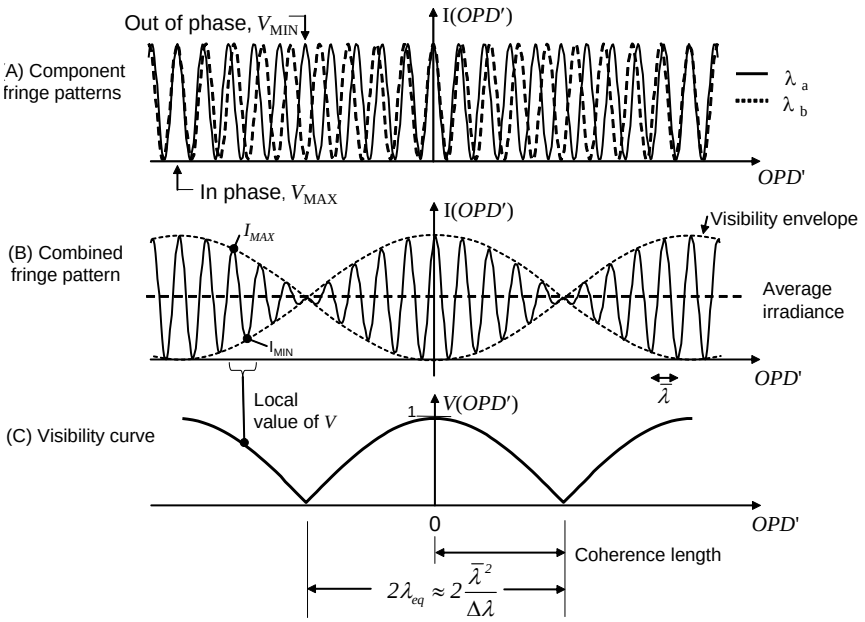


Figure 5.8. Fringe pattern analysis for a YDPI two-wavelength source. A) Line profiles of the component fringes; B) The total irradiance, which is the linear sum of the component fringe irradiance. Notice that the modulation envelope of this pattern is the visibility function, and the high frequency fringes occur with period $\bar{\lambda}$ in units of OPD'; C) The visibility as a function of OPD' shows a periodic variation. The coherence length is $\lambda_{eq} = \bar{\lambda}^2/\Delta\lambda$.

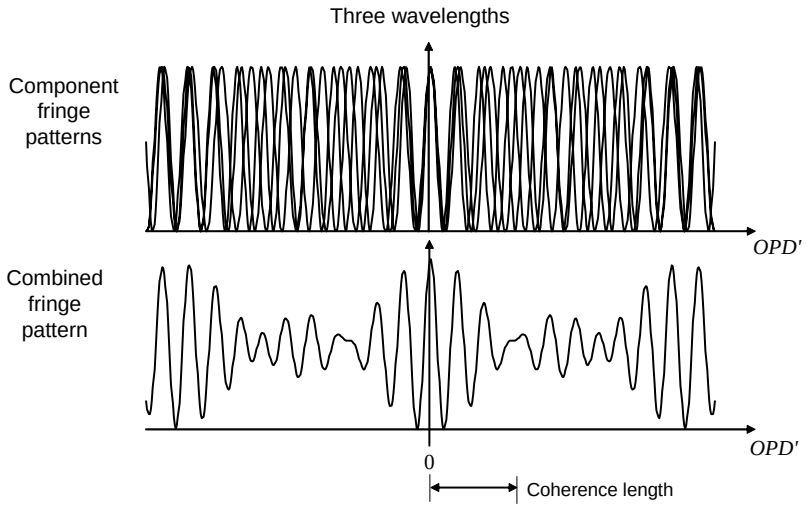


Figure 5.9. Fringes in the YDPI observation plane for small y_o with a dual-wavelength on-axis soure. Fringes are straight and equally spaced. Regions of low *V* are easily observed, notice that low *V* regions contain significiant irradiance, but they do not contain clear modulation. The oherence length is defined as the distance from the central fringe with maximum *V* to the first region of minimum *V.*

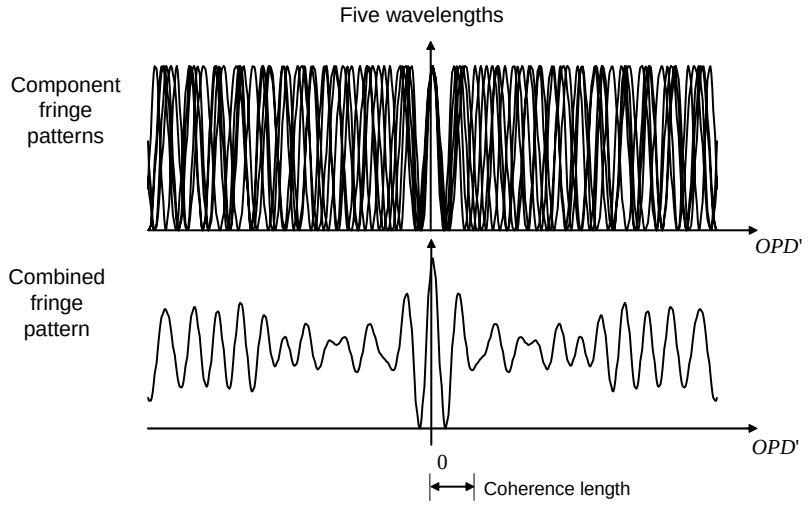


Figure 5.10. Fringe pattern analysis similar to Fig 5.9 except five wavelengths are used. Notice the reduction in coherence length due to increased bandwidth.

5.1.3 The power spectrum of the source

The discrete wavelength character of the optical sources considered in Section 5.1.2 is not found in nature. Thermal sources, like tungsten filaments, emit a continuous distribution of wavelengths that varies as a function of the filament temperature. Gas discharge lamps can provide a narrow spectrum of radiation limited by the Lorentzian linewidth of atomic transitions, but separate molecules in the gas emit energy independently at frequencies over the linewidth. Even with laser sources, a continuous description of the wavelength spectrum is required. Stimulated emission, like that found in laser cavities, can exhibit a narrow spectrum of less than one part in 10^{13}, but mechanical instabilities in the laser cavity and other sources of instability result in slight variations of phase $\beta(t)$ of the output. The phase variations cause tiny modulations of the wavelength, which can influence interference measurements. Fortunately, all of these wavelength characteristics can be described in a straightforward formalism called the *power spectrum*, which is the more common name for the wavelength spectrum.

Consider an optical wave that contains a collection of frequency components with differential amplitudes given by $a(\nu)$ and phases given by $\phi(\nu)$. The total electric field is

$$E(t) = \int_0^\infty a(\nu) e^{j\varphi(\nu)} e^{j2\pi\nu t} d\nu . \tag{5.17}$$

Notice that Eq.(5.17) restricts the temporal frequency of the field to positive frequencies. This restriction is a consequence of relating the real-valued electromagnetic disturbance to a complex quantity, which is known as the *analytic signal*. (More detail concerning the analytic signal is provided in Section 5.3.) Given $E(t)$, $a(\nu)e^{j\varphi(\nu)}$ is found by

$$a(\nu)e^{j\varphi(\nu)} = \int_{-\infty}^{\infty} E(t)e^{j2\pi\nu t}dt. \tag{5.18}$$

Irradiance I of the optical field represented by Eq. (5.17) is

$$I \propto E(t)E^*(t) = \int_0^{\infty} a^2(\nu)d\nu = \int_0^{\infty} G(\nu)d\nu. \tag{5.19}$$

The quantity $a^2(\nu)d\nu$ in Eq. (5.19) is proportional to the contribution to the irradiance from the frequency range $(\nu, \nu + d\nu)$. $G(\nu)$ is the *spectral density* of the light vibrations, and $G(\nu) = a^2(\nu)$ is referred to as the *power spectrum* of the light field. A typical power spectrum is shown in Fig. 5.11. The frequency range is limited to width $\Delta\nu$ around mean frequency $\bar{\nu}$. A typical power spectrum of optical source, with mean frequency $\bar{\nu}$ and bandwidth $\Delta\nu$.

Section 5.1.2 establishes that, for the purpose of calculating irradiance patterns, the patterns from two separate wavelengths are independent. Extrapolation of this result to a continuous wavelength distribution yields that each spectral density component of the power spectrum is also independent. Therefore, the calculation of irradiance patterns for interference phenomena is accomplished by a linear sum of the patterns from each spectral density component.

Real measurements of the power spectrum are taken over non-infinite time intervals. For example, if $E_T(t)$ is the analytic signal taken over time interval $|t| < T$ and $a_T(\nu)e^{j\varphi_r(\nu)}$ are the spectral components,

$$E_T(t) = \int_0^{\infty} a_T(\nu)e^{j\varphi_T(\nu)}e^{j2\pi\nu t}d\nu. \tag{5.20}$$

With the assumption of stationary and ergodic statistics,[1]

$$I = \int_0^{\infty} \lim_{T\to\infty} \frac{\overline{a_T^2(\nu)}}{T} d\nu, \tag{5.21}$$

1. Stationarity is defined in Section 5.1.1, and *ergodic statics* imply that an ensemble average is equal to the correspondig time average involving a typical member of the ensemble.

where the bar denotes an ensemble average over many samples. Therefore, the *measured* power spectrum is actually

$$G(\nu) = \lim_{T\to\infty} \frac{\overline{a_T^2(\nu)}}{T}. \tag{5.22}$$

5.1.4 Basic temporal coherence

Functional relationships that describe the form of the fringe and the change of fringe contrast with $OPD' = v\Delta t$ are derived with a simple argument. Consider the YDPI of Fig. 5.5 in air with a power spectrum given by $a^2(\nu)$. Spectral component $a^2(\nu_0)$ of the source power spectrum produces an interference pattern given by

$$I(\nu_0;\Delta t) \approx a^2(\nu_0)\left[K_1(y_0) + K_2(y_0) + 2\sqrt{K_1(y_0)K_2(y_0)}\cos\left(2\pi\nu_0\frac{OPD'}{c}\right)\right], \tag{5.23}$$

where $K_1(y_0)$ and $K_2(y_0)$ are diffraction constants that depend on pinhole size and $\Delta t = OPD'/c$, where c is the speed of light in a vacuum (3×10^8 m/s). $K_1(y_0)$ and $K_2(y_0)$ vary slowly with wavelength. For this development, assume that K_1 and K_2 are simple constants that do not vary with wavelength and are not a function of observation-plane position. Pinholes are extremely small, so that diffraction effects from the finite pinhole size can be neglected.

Irradiance patterns from individual source wavelengths are now combined to obtain the total fringe pattern. If Eq. (5.23) is integrated over the bandwidth of the source,

$$\begin{aligned} I(\Delta t) &\approx (K_1 + K_2)\int_\infty a^2(\nu)d\nu + 2\sqrt{K_1K_1}\int_\infty a^2(\nu)(2\pi\nu_0\Delta t)d\nu \\ &= (K_1 + K_2)I_W + 2\sqrt{K_1K_1}Re\left\{\int_\infty a^2(\nu)(2\pi\nu_0\Delta t)d\nu\right\} \\ &= (K_1 + K_2)I_W + 2\sqrt{K_1K_1}I_W Re\left\{\frac{\Im\{a^2(\nu)\}\Delta t}{I_W}\right\} \\ &= (K_1 + K_2)I_W + 2\sqrt{K_1K_1}I_W m_{12}(2\pi\nu_0\Delta t)\cos\left|(\varphi_{12}(\Delta t))\right| \end{aligned}, \tag{5.24}$$

where

$$I_W = \int_\infty a^2(\nu)d\nu \tag{5.25}$$

is the integral of the source power spectrum,

$$m_{12}(\Delta t) = \left| \frac{\Im\{a^2(\nu)\}_{\Delta t}}{I_W} \right| \tag{5.26}$$

is the normalized Fourier transform of the source power spectrum, and

$$\varphi_{12}(\Delta t) = arg[\Im\{a^2(\nu)\}_{\Delta t}] \tag{5.27}$$

is the phase of the Fourier transform. Fringe contrasting, i.e. visibility is

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} = \frac{2\sqrt{K_1 K_2}}{K_1 + K_2} m_{12}(\Delta t) \quad . \tag{5.28}$$

Usually,

$$a^2(\nu) = f(\nu - \bar{\nu}) , \tag{5.29}$$

where $\bar{\nu}$ is the mean frequency of the power spectrum. In this case,

$$\varphi_{12}(\Delta t) = 2\pi\bar{\nu}\Delta t + \beta_{12}(\Delta t) \qquad , \tag{5.30}$$

where

$$\beta_{12}(\Delta t) = arg[\Im\{f(\nu)\}_{\Delta t}] . \tag{5.31}$$

Equations (5.24) through (5.31) imply that integration of cosine fringe patterns from individual wavelengths in the power spectrum yields an aggregate fringe period $OPD' = c/\bar{\nu} = \bar{\lambda}$ and a visibility that depends on the Fourier transform of the power spectrum of the source. Fringe shift β_{12} is observed if the power spectrum is asymmetric about the mean frequency $\bar{\nu}$.

Example 5.1: Simple rectangular power spectrum

Consider a source with a power spectrum given by a simple rectangular function, with a mean frequency $\bar{\nu}$ and a bandwidth $\Delta\nu$. Mathematically, the rectangular distribution is described by

$$a^2(\nu) = \frac{1}{\Delta\nu} \mathrm{rect}\left(\frac{\nu - \bar{\nu}}{\Delta\nu} \right) . \tag{5.32}$$

Calculation of the Fourier transform of Eq. (5.32) yields

$$\Im\{a^2(\nu)\}_{\Delta t} = \mathrm{sinc}(\Delta\nu\Delta t) e^{j2\pi\bar{\nu}\beta t} \quad . \tag{5.33}$$

From Eq. (5.25), we calculate that $I_W = 1$. Using Eqs. (5.26) and (5.27) with the result from Eq. (5.33), m_{12} and φ_{12} are found to be

$$m_{12}(\Delta t) = \left| \mathrm{sinc}(\Delta\nu\Delta t) \right| , \tag{5.34}$$

$$\varphi_{12}(\Delta t) = 2\pi\bar{\nu}\Delta t + \arg[\mathrm{sinc}(\Delta\nu\Delta t)]\,. \tag{5.35}$$

Substitution of Eqs. (5.34) and (5.35) into Eq. (5.24) and recalling that for a two pinhole interferometer $OPD' = c\Delta t \approx dy_0/z_0$ yield the following interference pattern in the observation plane

$$I(y') = \left[1 + \mathrm{sinc}\left(\frac{d\Delta\nu y_o}{cz_o}\right)\cos\left(2\pi\frac{d\bar{\nu}y_o}{cz_o}\right)\right] \quad , \tag{5.36}$$

where the cosine term creates fringes with period $\Lambda = cz_0/d\bar{\nu}$ and the sinc term is the modulation in visibility. A gray-scale picture of the fringe pattern is shown in Fig. 5.12, and plots of the irradiance and visibility along the y-axis are shown in Fig. 5.13. Note that the sinc function in Eq. (5.36) can be either positive or negative, depending on the value of OPD'. Since m_{12} is the absolute value of the sinc function, this phase information is not considered to be part of m_{12}. Instead, it is considered as part of β_{12}.

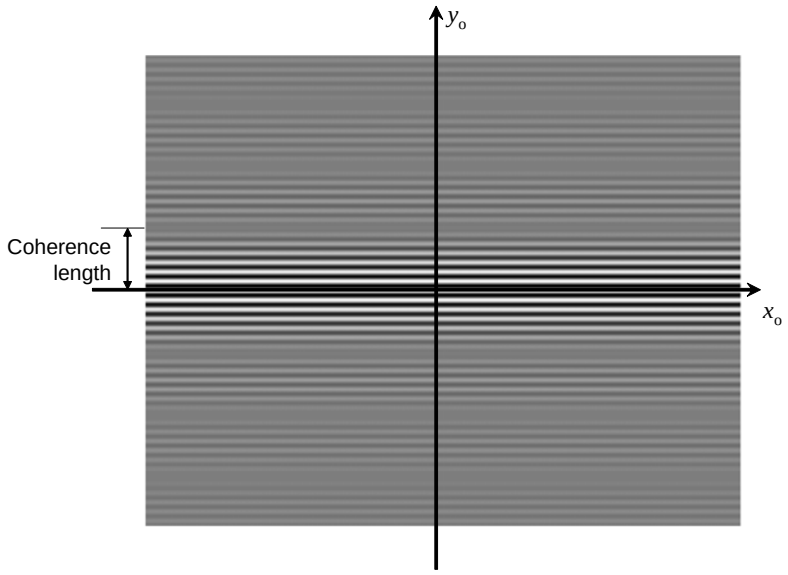


Figure 5.11. Fringes in the YDPI observation plane for the rectangular power spectrum of Example 5.1.

In Example 5.1, visibility reduces as Δt (and, thus OPD′) increases from zero, as evidenced by the form of the sinc function in Eq. (5.36). When visibility

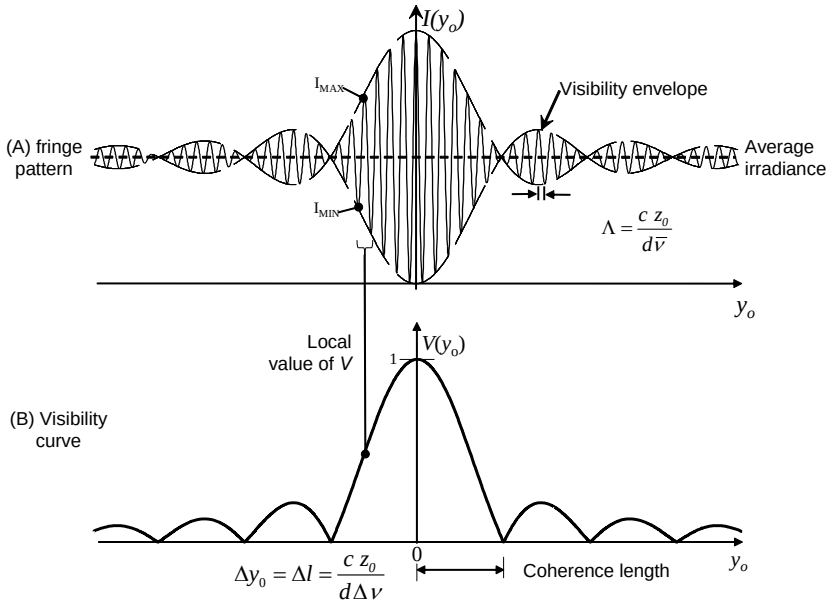


Figure 5.12. Line profiles of the fringe pattern in Fig 5.11, where the envelope of the fringe pattern is a sinc function.

drops to zero, there is no fringe contrast. y_0 distance between maximum fringe visibility at $OPD' = 0$ and the first zero in visibility is called the *coherence length*. For example, in Eq. (5.36), coherence length $\Delta l = cz_0/d\Delta\nu$ *Coherence time*, which corresponds to coherence length, is given by $\Delta t_c = 1/\Delta\nu$. In a practical interferometer, like a Twyman-Green interferometer, coherence length is the maximum difference in path length between the two arms of the interferometer that result in high contrast fringes. Many interferometers exhibit similar behavior, with the general result that coherence time is approximately the inverse of the width of the power spectrum. In practice, coherence length may not be defined where visibility reaches an absolute zero, but, rather, where visibility reaches some minimum value, like $V = 0.2$.

Example 5.2: Twyman-Green interferometer with a narrow Gaussian power spectrum.

Consider the power spectrum of a frequency-stabilized HeNe laser ($\bar{\lambda}$ = 632.8 nm) with a Gaussian power spectrum and $\Delta\nu$ = 300 kHz. The application is for a Twyman-Green interferometer, where the flatness Δy_o of a mirror is to be tested,

as shown in Fig. 5.4. What is the maximum distance that can be set for the mirror separation if the minimum resolvable contrast is $V = 0.2$?

By the peculiar property of a Gaussian function, the Fourier transformation of the Gaussian power spectrum is also in the form of a Gaussian function. If we express the power spectrum $a^2(\nu) = \text{gaus}[(\nu - \bar{\nu})/\Delta\nu]$ when $\text{gaus}(\nu) = \exp(-\pi\nu^2)$, the calculation of its Fourier transformation yields

$$\left|\Im\{a^2(\nu)\}_{\Delta t}\right| = \left|\Delta\nu\,\text{gaus}(\Delta\nu\Delta t)e^{-2\pi\bar{\nu}\Delta t}\right|. \tag{5.37}$$

From (5.26), the normalized Fourier transform of the source power spectrum is

$$m_{12}(\Delta t) = \left|\text{gaus}(\Delta\nu\Delta t)\right|, \tag{5.38}$$

after normalization with I_W and taking the absolute value. The fringe contrast V is proportional to the absolute of the normalized Fourier transform. That is,

$$V = \left|\text{gaus}(\Delta\nu\Delta t)\right|. \tag{5.39}$$

After replacing OPD' by twice the mirror separation $2|d_1 - d_2|$ (because light in a Twyman-Green interferometer passes an arm twice) and manipulating the contrast to be larger than 0.2, we obtain the maximum mirror separation:

$$\left|d_1 - d_2\right|_{max} = \frac{c}{2\Delta\nu}\sqrt{\frac{\ln 5}{\pi}} \cong 358\ m. \tag{5.40}$$

For $V > 0.5$, where the fringes can be clearly observed, the maximum mirror separation is 235 m. Due to the long coherence length originating from the narrow spectrum width of the HeNe laser, the fringe visibility does not change much over the small OPD' changes associated with the mirror imperfections.

Example 5.3: Laser diode power spectrum

Consider the power spectrum shown in Fig. 5.13 of a typical laser diode with a mean wavelength $\bar{\lambda} = 649.2$ nm ($\bar{\nu} = c/\lambda = 4.621\times10^{14}$ Hz), where parameters of the diode laser include $n_{cavity} = 3.5$, L = 200 μm, $\Delta\nu = 3c/2n_{cavity}L = 6.4\times10^{11}$ Hz and $\Delta\nu_m = 25$ MHz. $\Delta\nu_m$ is the bandwidth of an individual laser mode. A mathematical description of the power spectrum is

$$a^2(\nu) \approx A_0\,\text{rect}\left(\frac{\nu - \bar{\nu}}{\Delta\nu}\right)\left[\text{comb}\left[\frac{\nu}{c/2n_{cavity}L}\right]\text{wgaus}\left(\frac{\nu}{\Delta\nu_m}\right)\right], \tag{5.41}$$

where (*) denotes one-dimentional convolution.

Hence ,

$$m_{12}(\Delta t) \propto \frac{\left|\Im\{a^2(v)\}_{\Delta t}\right|}{Iw} = \left|\mathrm{sinc}(\Delta t \Delta v)\left[\mathrm{comb}\left(\frac{c\Delta t}{2n_{cavity}L}\right)\mathrm{gaus}(\Delta \nu_m \Delta t)\right]\right|. \quad (5.42)$$

Figure 5.14 shows the visibility curve versus OPD'. Figure 5.15 shows the irradiance versus $OPD' = c\Delta t$.

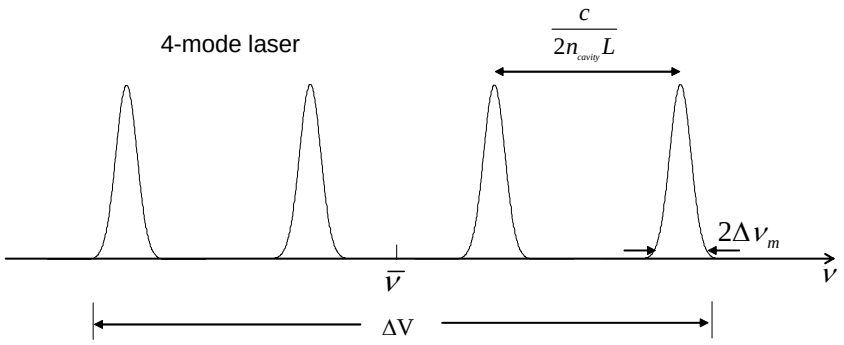


Figure 5.13. Power spectrum typical of laser diodes.

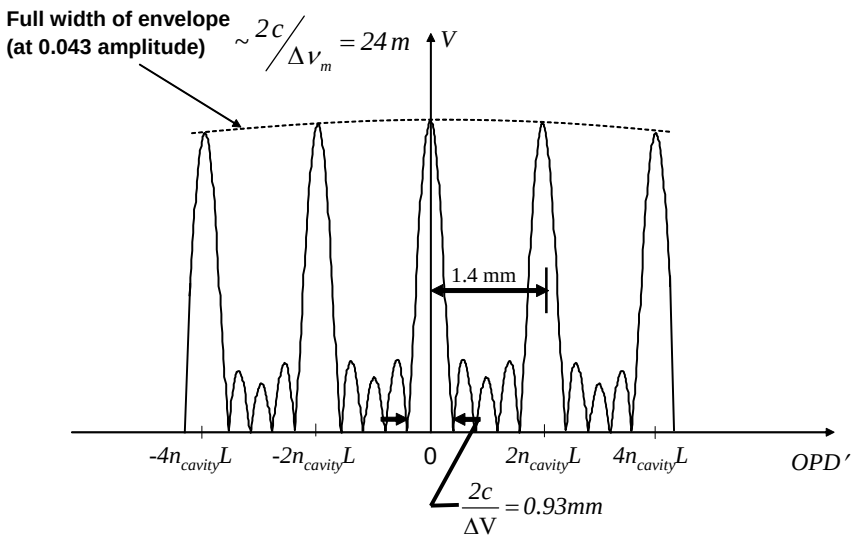


Figure 5.14. Visibility function associated with Fig. 5.13 and the parameters of Example 5.3.

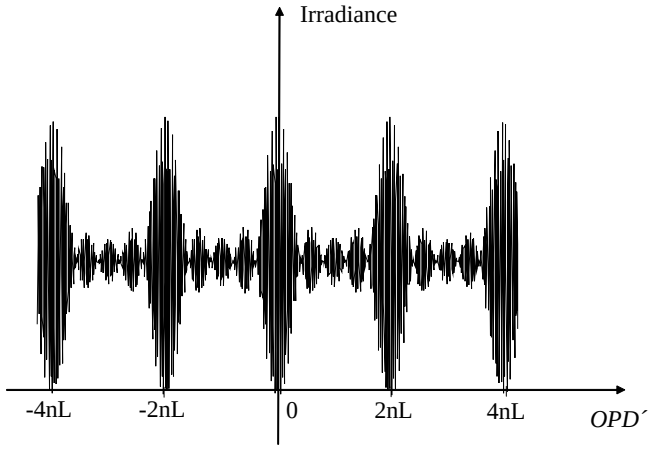


Figure 5.15. Profile of the fringe pattern observed for Example 5.3.

If a single-mode version of this laser is used in a Twyman-Green interferometer, the visibility only depends on $\Delta\nu_m$, which leads to a 12 m coherence length. However, the multiple-mode nature of the power spectrum in Fig. 5.13 dramatically shortens the coherence length to under 0.5 mm. The visibility returns periodically at $OPD' = 2mn_{cavity}L$, where m is an integer.

Example 5.4: Fourier Transform Spectroscopy

Often it is necessary to measure the power spectrum of a source. A detailed description of the power spectrum can yield information about the chemical composition of a star, energy level data of atomic transitions, atmospheric pollution and other applications. The device often used to measure the power spectrum is a *Fourier Transform Spectrometer.* A simple Fourier Transform Spectrometer is shown in Fig. 5.16. It is basically a Twyman-Green Interferometer, in which the path length of one arm is adjusted with a movable scan mirror. d_1 and d_2 are the lengths of each arm, respectively. Thus, the optical path difference (OPD') is

$$OPD' = 2(d_2 - d_1) \ . \tag{5.43}$$

When observing a distant point source, the fringe visibility is recorded as a function of OPD'. For example, Fig. 5.17 shows a hypothetical fringe pattern and visibility measured by the detector that is characteristic of astronomical observations. The spectral power density is found from the inverse Fourier transform of Eq. (5.26), or

$$a^2(\nu) = I_W \Im^{-1}\{V(\Delta t)\}_\nu \quad , \tag{5.44}$$

where V is the visibility. The spectral power density calculated from the data in Fig. 5.17 is shown in Fig. 5.18. For more information on Fourier Transform Spectroscopy, the reader is directed to the References 5.8 and 5.9.

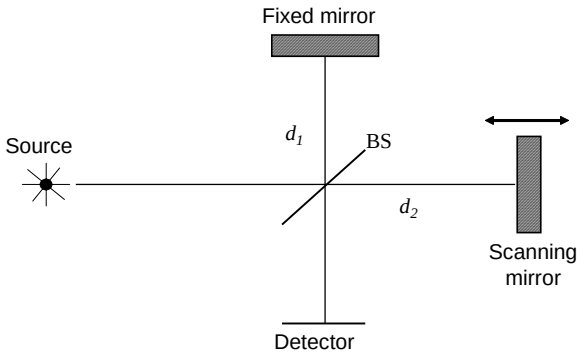


Figure 5.16. Tymann-Green interferometer used as a Fourier Transform Spectrometer.

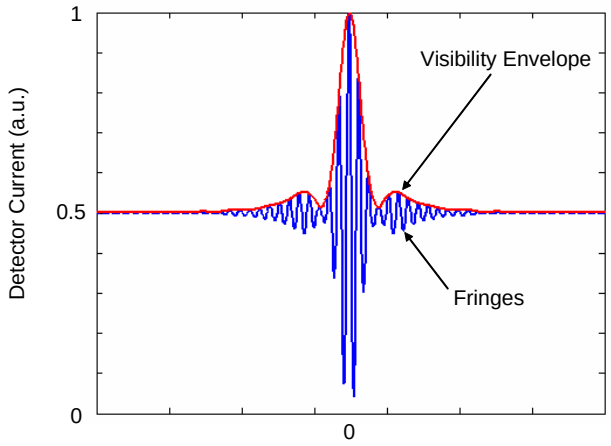


Figure 5.17. Fringe profile used in Example 5.4, which describes Fourier Transfprm Spectroscopy.

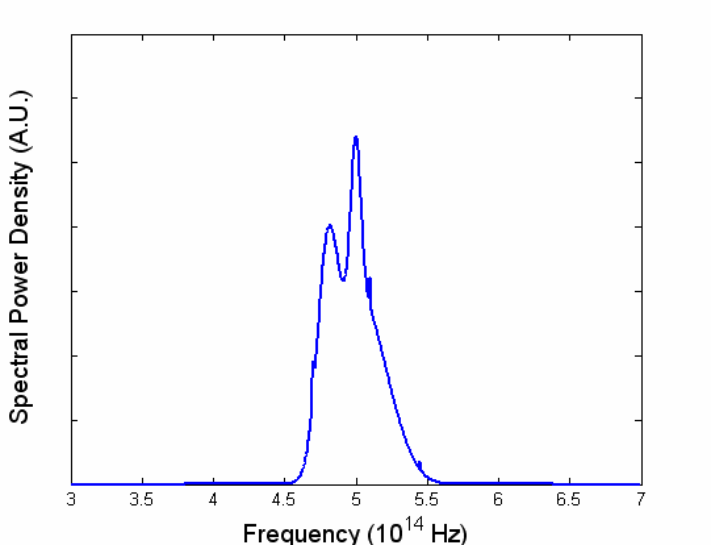


Figure 5.18. Power spectral density calculated by the inverse transfrom of data displayed in Fig 5.17.

5.1.5 Basic Spatial Coherence

Section 5.1.4 assumes that the electric field reaching the observation point from P2 is simply a delayed copy of the electric field reaching the observation point from P1. This condition occurs only if the two electric fields originate from precisely the same position in space, *i.e.* a perfect point source. However, if the electric field originates from a source of finite size, there are additional spatial fluctuations in the electric field.

For example, consider the effects of two independent point sources located at y_{s1} and y_{s2}, which are used as separate sources in a Young's double pinhole experiment, as shown in Fig. 5.19. The sources are quasimonochromatic. That is, they have nearly the same wavelength, but the temporal beat frequency resulting from the combination of the light from the two sources is beyond the temporal bandwidth of the detector used to measure the aggregate interference pattern. Because the sources are independent, the interference pattern from one source adds in irradiance with the pattern from the other source. This behavior is similar to the result found in Section 5.1.2, where the irradiance patterns from different wavelengths are added independently, except that now the irradiance patterns resulting from displaced source points are added.

The on-axis source at y_{s1} generates a centered cosine fringe pattern on the observation plane. The off-axis source at y_{s2} generates a shifted fringe pattern due to the additional phase difference

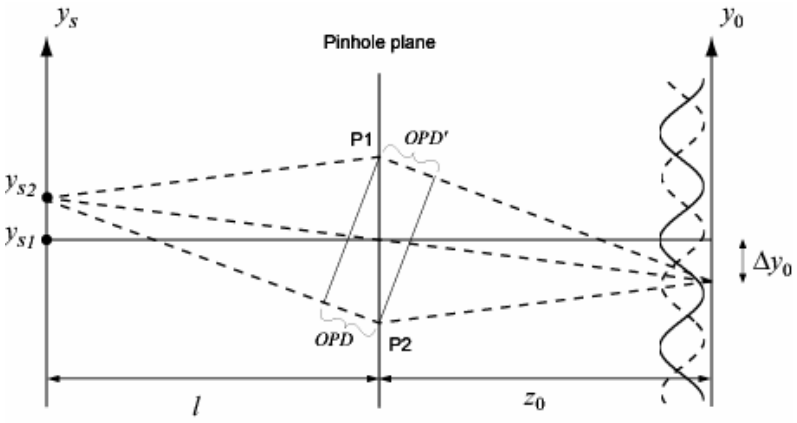


Figure 5.19. The effect of adding a second quasimonochromatic point source at y_{s2} is to create a second fringe pattern in the observation plane that is shifted by Δy_0 .

$$OPD = \frac{dy_{s2}}{l}, \tag{5.45}$$

where d is the spacing of the pinholes. The shift Δy_0 of the fringe pattern caused by this phase difference along the y_0-axis is given explicitly by

$$\Delta y_0 = \frac{z_0 y_{s2}}{l} \quad , \tag{5.46}$$

which is independent of source wavelength and pinhole separation. In fact, the shift is simply found by applying similar triangles to the geometry of Fig. 5.19.

The component fringe patterns and the combined fringe pattern are shown in Fig. 5.20 for two source points, where $\Delta y_0 = \Lambda / 5$ and Λ is the fringe period. Notice that the combined pattern has reduced visibility, and the visibility is not a function of $OPD' = d\Delta y_0 / z_0$. Also, the center of the combined pattern is slightly shifted off axis. If more point sources are added to the source distribution, this trend continues. For example, in Fig. 5.21, the combined fringe pattern from three equally-spaced point sources exhibits significantly degraded visibility. Like with two point sources, the visibility is not a function of OPD' . If the source distribution is wide enough such that the fringe centers fill one fringe period, the result can be $V = 0$, as shown in Fig. 5.22. Although $V = 0$ in Fig. 5.22, there is significant irradiance across the observation region.

Figures 5.20 through 5.22 illustrate basic properties of spatial coherence. Firstly, *visibility is not a function of location in the observation space*. Secondly, *visibility is inversely proportional to the source size*. These properties differentiate spatial coherence from temporal coherence.

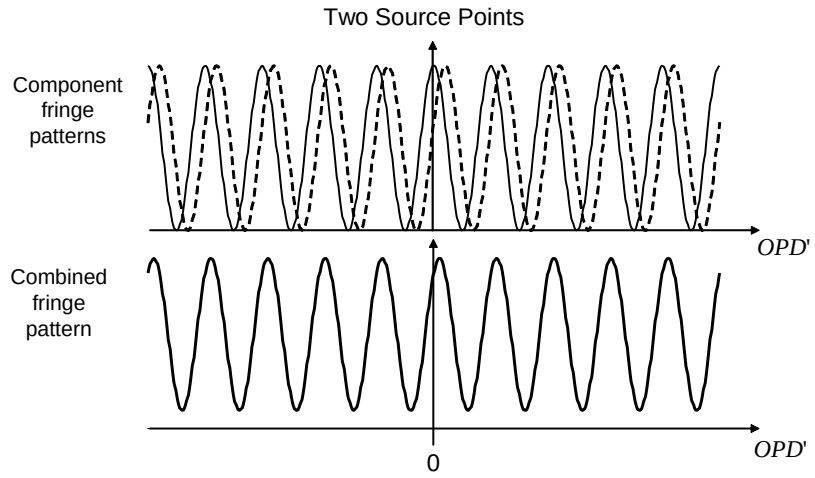


Figure 5.20. The component fringe patterns from an on-axis source point and an off-axis source point is shown, where $\Delta y_0 \approx \Lambda / 5$. The total irradiance is the sum of the two fringe patterns. Notice that visibility is reduced uniformly across the observation plane, where $OPD' = d\Delta y_0 / z_0$. The center of the combined pattern is slightly shifted, but it retains the same period as the component fringes.

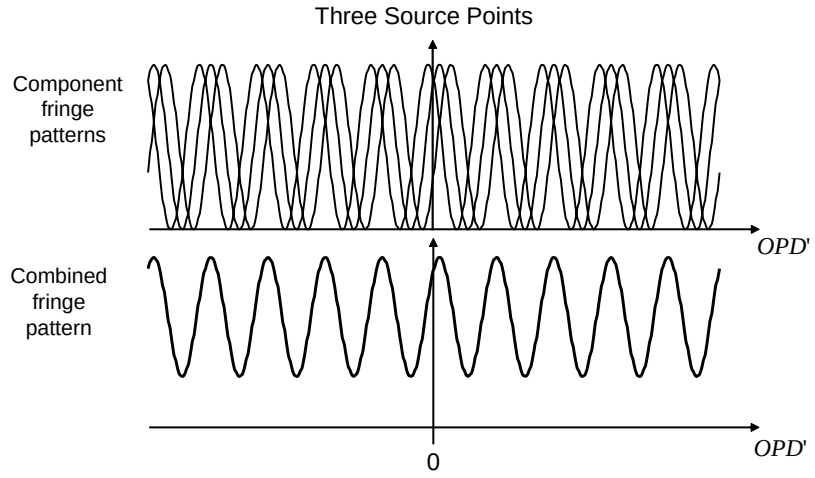


Figure 5.21. A third source point is added, which further reduces contrast.

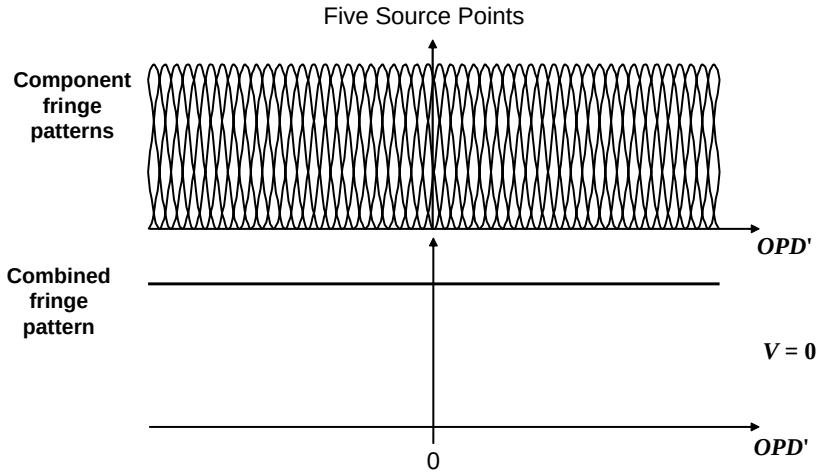


Figure 5.22. When five equally-spaced source points are in the source plane, this particular geometry produces no fringe modulation in the observation plane.

Now the idea of adding irradiance patterns from individual source points is extended to a continuous source distribution. The fringe pattern for a point source with radiance $a^2(y_s)$ in the source plane is

$$I(OPD';y_s) = K_1 a^2(y_s) + K_2 a^2(y_s) + \sqrt{K_1 K_2}\, a^2(y_s) \cos\left[k\left(OPD' - \frac{dy_s}{l}\right)\right] \quad (5.47)$$

where $k = 2\pi/\bar{\lambda}$, and $\bar{\lambda}$ is the average wavelength of the quasimonochromatic source. Integration over the source spatial distribution yields

$$\begin{aligned} I(OPD') &= \int_{source} \left\{ [K_1 + K_2] a^2(y_s) + \sqrt{K_1 K_2}\, a^2(y_s) \cos\left[k\left(OPD' - \frac{dy_s}{l}\right)\right] \right\} dy_s \\ &= (K_1 + K_2) I_L + 2\sqrt{K_1 K_2}\, Re\left\{ e^{jkOPD'} \int_{source} a^2(y_s) e^{-jk\theta_A y_s} dy_s \right\} \\ &= (K_1 + K_2) I_L + 2\sqrt{K_1 K_2}\, Re\left\{ e^{jkOPD'} \frac{\Im\{a^2(\nu)\}_{\frac{\theta_A}{\bar{\lambda}}}}{I_L} \right\} \\ &= (K_1 + K_2) I_L + 2\sqrt{K_1 K_2}\, I_L \mu_{12} \cos(kOPD' + \beta_{12}) \end{aligned} \quad (5.48)$$

where $\theta_A = d/l$ is the angular subtense of the pinholes as seen from the source,

$$\mu_{12}\left(\frac{\theta_A}{\bar{\lambda}}\right) = \left|\frac{\Im\{a^2(\nu)\}_{\frac{\theta_A}{\bar{\lambda}}}}{I_L}\right|, \tag{5.49}$$

$$I_L = \int_{source} a^2(y_s)dy_s \tag{5.50}$$

and

$$\beta_{12}\left(\frac{\theta_A}{\bar{\lambda}}\right) = arg\left\{\Im\{a^2(\nu)\}_{\frac{\theta_A}{\bar{\lambda}}}\right\}. \tag{5.51}$$

Note that the fringe visibility,

$$V = \frac{2\sqrt{K_1K_2}}{K_1 + K_2}\mu_{12}\left(\frac{\theta_A}{\bar{\lambda}}\right) \tag{5.52}$$

is not a function of position OPD' in the observation space. Although Eqs. (5.52) and (5.28) are similar; Eq. (5.52) does not depend on OPD'. Fringe spacing $\Lambda = z_0\bar{\lambda}/d$ in the observation plane is identical to the spacing of fringes in the interference pattern resulting from an individual point in the source.

5.1.6 Interpretation of the van Cittert-Zernike Theorem

Equations (5.49) and (5.52) are elements of the van-Cittert Zernike theorem, which can be stated:

If a quasimonochromatic source is a considerable distance from the aperture plane and pinhole separation is small, fringe visibility from an extended source is proportional to the Fourier transform of the source's spatial distribution. The transform variable is the angular separation of the measurement points divided by the wavelength.

Example 5.5: Spatial coherence from a one-dimensional line source

---

Consider a quasimonochromatic one-dimensional source as shown in Fig. 5.23, which is given by

$$a^2(y_s) = \text{rect}\left(\frac{y_s}{L}\right), \tag{5.53}$$

where, if $K_1 = K_2 = 1$,

$$V = \mu_{12}\left(\frac{\theta_A}{\bar{\lambda}}\right) = \left|\mathrm{sinc}\left(L\frac{\theta_A}{\bar{\lambda}}\right)\right|. \tag{5.54}$$

A plot of Eq. (5.54) is shown in Fig. 5.24 with $\theta_A = d/l$. A reference line at $Ld/l\bar{\lambda} = L\theta_A/\bar{\lambda} = 0.6$ is drawn to help illustrate variable dependencies. A complete understanding of the van Cittert-Zernike theorem is obtained by recognizing how visibility changes for each variable separately.

Increasing l or decreasing L has the effect of decreasing the maximum fringe shift and increasing visibility. The contrast is increased because the range of shifts in the component fringe patterns is reduced compared to the fringe period. (Remember, the fringe period does not depend on l or L.) Thus, decreasing the angular extent of the source as seen by the pinholes increases visibility. Increasing pinhole separation d or decreasing $\bar{\lambda}$ has the effect of reducing the fringe spacing, which decreases visibility because the range of shifts in the observation area does not change.

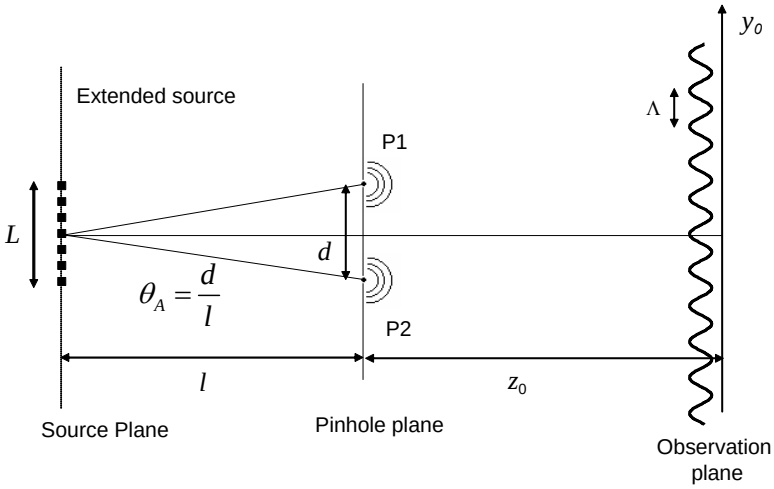


Figure 5.23. One-dimensional source distribution with length L and angular substance θ_A for Example 5.5.

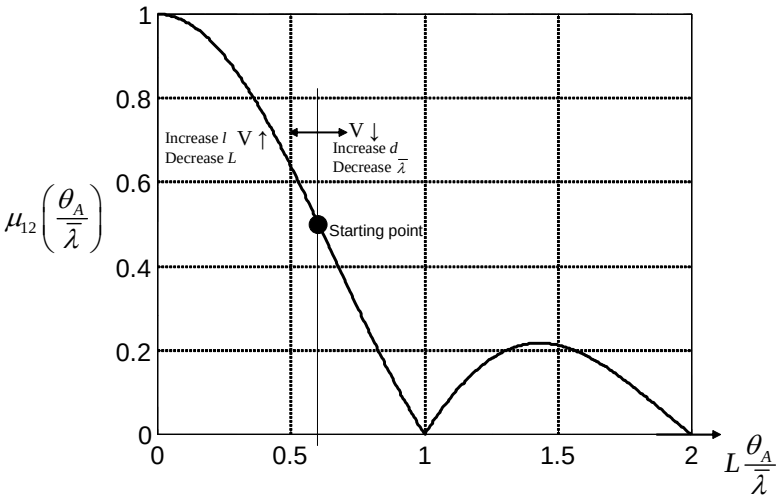


Figure 5.24. Visibility as a function of the normalized variables $L\frac{\theta_A}{\bar{\lambda}}$ for the one-dimensional extended source in Example 5.5. In order to increase visibility from the starting point, the soure-to-pinhole distane must be increased or the source-to-pinhole distance must be increased or the soure length L must be decreased increasing pinhole separation for decreasing wavelength $\bar{\lambda}$.

Example 5.6: Spatial coherence with a lateral shearing interferometer

Consider a lateral shearing interferometer, as shown in Fig 5.25. This type of interferometer is typically used to test collimation of laser beams and other optical systems. The shearing interferometer shifts the wavefront by the shear distance *S*, which is a function of the tilt angle φ, the thickness *T* and refractive index *n* of the plate. The incident wavefront must exhibit good spatial coherence over the shear distance in order to observe high contrast fringes.

If the extended quasimonochromatic source in Fig. 5.25 is collimated by a lens of focal length *f*, the *angular* distribution of the light exiting the lens determines the spatial coherence properties. That is, since the lateral shearing interferometer detects the derivative of the wavefront, tilt caused by off-axis point sources produces fringe shift. Shifts from all sources must be significantly less than a fringe period in order to observe high contrast fringes. The OPD between points interfering in the shift wavefront due to the source distribution.

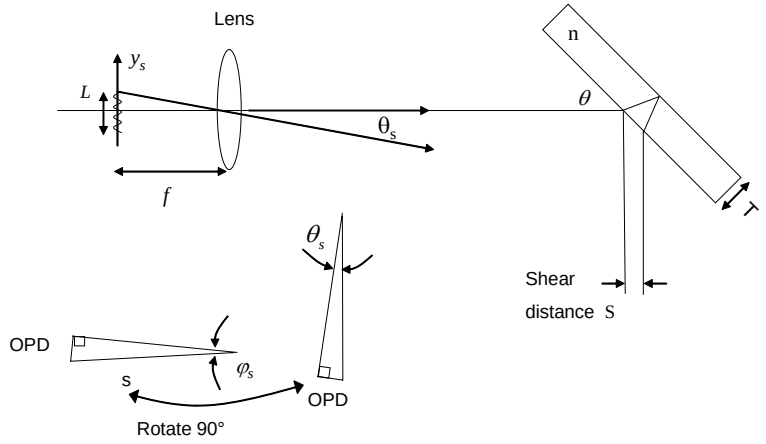


Figure 5.25. Spatial coherence for a lateral shearing interferometer (LSI).

$$OPD = S\sin\theta_s, \tag{5.55}$$

where

$$\sin\theta_s \approx \frac{y_s}{f}. \tag{5.56}$$

The fringe shift term in Eq. (5.47) is now Sy_s/f, so $\theta_A = d/l$ must be replaced with $\theta_A = S/f$.

For a rectangular source distribution,

$$\begin{aligned} \mu_{12}\left(\frac{\theta_A}{\bar{\lambda}}\right) &= \left|\operatorname{sinc}\left(L\frac{\theta_A}{\bar{\lambda}}\right)\right| \\ &= \left|\operatorname{sinc}\left(L\frac{S}{\bar{\lambda}f}\right)\right| \end{aligned}. \tag{5.57}$$

Therefore, the extent L of the incoherent quasimonochromatic source must be smaller than $f\bar{\lambda}/S$ to obtain good contrast. For example, if T = 12.5 mm, n = 1.5 and θ = 45°, S = 5.2 mm. If f = 100 mm, the source size that first produces a minimum in visibility is

$$L = \frac{f\bar{\lambda}}{S} = \frac{(100\times10^{-3}\ m)(550\times10^{-9}m)}{5.2\times10^{-3}m} = 1.1\times10^{-5}m = 11\mu m\ .$$

Example 5.7: Michelson Stellar Interferometer

The Michelson stellar interferometer is shown in Fig. 5.26. The fringe spacing on the observation screen depends on d. However, the visibility depends on *h*, the separation of the collection mirrors, because, like in Example 5.6, $OPD = h\sin\theta_s$. However, the source radiance is specified in terms of the angle as viewed from the center of the telescope, not as a physical distance in a plane. In this case, the development of Eq. (5.49) is modified to use $a^2(\theta_s)$ with $h/\bar{\lambda}$ as the transform variable. Thus, the observer maps the fringe visibility by changing *h*. The angular radiance distribution $a^2(\theta_s)$ is calculated by taking the inverse Fourier transform of measured visibility.

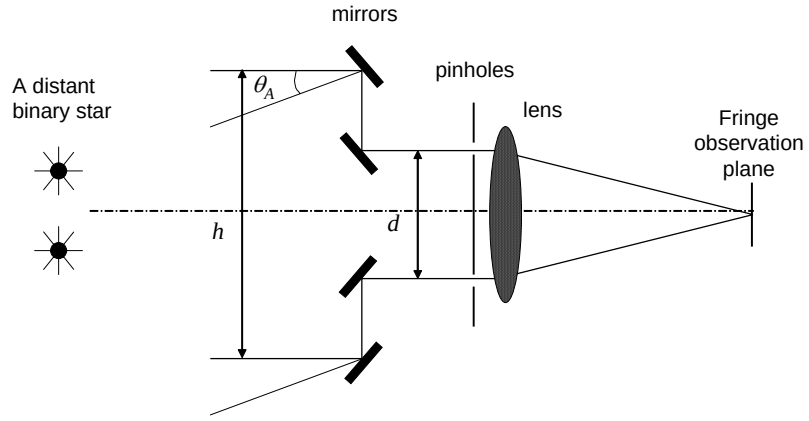


Figure 5.26. Michelson Stellar Interferometer for viewing distant stars..

One application of the Michelson stellar interferometer is used to measure the angular separation of binary stars. The radiance distribution is

$$a^2(\theta_s) = I_0[\delta(\theta_s + \Delta\theta) + \delta(\theta_s - \Delta\theta)] \ , \qquad (5.58)$$

where $\pm\Delta\theta$ are the angles between the optical axis and the ray coming from each star, and I_0 is the radiance of each star. Thus, the visibility is

$$V = \frac{\left|\Im^{-1}[a^2(\theta_s)]_{\frac{h}{\bar{\lambda}}}\right|}{2I_0} = \frac{\left|I_0\left(e^{j2\pi h\Delta\theta/\bar{\lambda}} + e^{-j2\pi h\Delta\theta/\bar{\lambda}}\right)\right|}{2I_0} = \left|\cos(2\pi h\Delta\theta/\bar{\lambda})\right| \quad . \tag{5.59}$$

Therefore, if the separation of mirrors and wavelength are determined, and the visibility is measured, the angular separation $\Delta\theta$ can be obtained. In the example above, if h is the mirror separation that produces the first zero in fringe visibility, $\Delta\theta$ is given by

$$\Delta\theta = \frac{\bar{\lambda}}{4h} \quad . \tag{5.60}$$

For example, consider when the first zero of fringe visibility occurs at h = 27 m with $\bar{\lambda}$ = 500 nm, $\Delta\theta = 4.6\times10^{-9}$ radians, and the angular separation of the binary stars is about 2 milli-arc-seconds (mas), where 1 mas = 4.6 nanoradians (10^{-9} radians). For comparison, the angular resolution of the human eye is about 12.5 arc-seconds or 58 microradians (10^{-6} radians). The Hale observatory has a resolution of about 25 mas, and the Naval Observatory Michelson Stellar Interferometer has a resolution of about 200 micro-arc-seconds.[1] In simple terms this capability is like resolving the width of a human hair at a distance of 25 miles. The optical schematic for the Mark III Stellar Interferometer at the Naval Observatory is shown in Fig. 5.27.

5.1.7 The Concept of Coherence Area

The van Cittert-Zernike Theorem of Section 5.1.6 applied to a two-dimensional quasimonochromatic source distribution provides insight into the coherent nature of the light field. The concept of *coherence area* suggests that there is some area over which the light field behaves coherently at various distances away from the source. For example, the sketch in Fig. 5.28 shows the outline of a source distribution and the outline of it's associated Fourier transform in an aperture/pinhole plane. The aperture/pinhole plane could be the pinhole plane of a YDPI, the input aperture of an interferometer or the object plane of a microscope. The distribution of the Fourier transform, appropriately scaled from Eq. (5.49), determines the visibility of interference fringes for separations d of points in the transform plane. The coherence area is defined by *the region where points in the plane*

1. See http://www.mtwilson.edu/Tour/NRL/mark3.html for more information about the Naval Observatory's instrument.

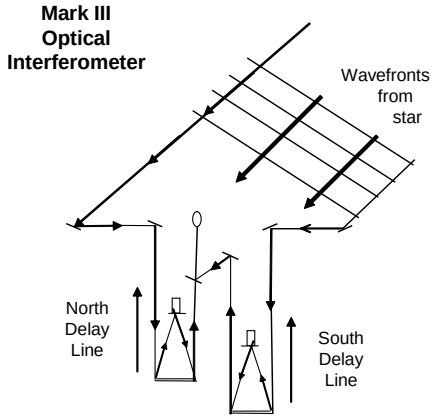


Figure 5.27. The Mark III stellar Interferometer at the Naval Observatory.

around the optical axis can interference with high visibility. The maximum extent of this region is the boundary determined by the minimum visibility acceptable to the optical system of interest. Usually, the zeros of the visibility function are the easiest limits to consider, but a value of $V = 0.2$ is often used for experimental data.

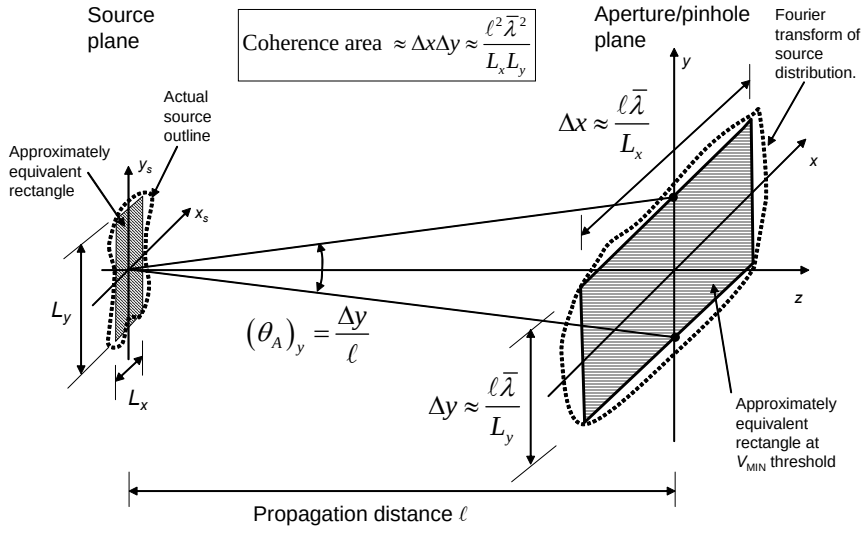


Figure 5.28. The concept of coherence area is an extension of the Van Cittert Zernike Theorem.

The source in Fig. 5.28 is longer in the y_S dimension than in the x_S dimension. The source distribution can be roughly approximated with a rectangle with sides of length $L_x L$ and $L_x L$. The Fourier transform of the rectangle is separable, and the zeros of the resulting sinc visibility function are used to define the coherence area. For example, the maximum width of for good coherence (zeros of *V*) in the y-dimension is $\Delta y = l\bar{\lambda}/L_y$. Note that, because of the rectangular shape in the source plane, the coherence area is also wider in one dimension than in the other. The wide dimension of the coherence area corresponds to the narrow dimension of the source. That is, *narrow source dimensions present a wide area coherence in the pinhole/aperture plane*. In the example of Fig. 5.28, the coherence area is

$$\text{Coherence area} = \Delta x \Delta y = \frac{l^2 \bar{\lambda}^2}{L_x L_y} . \tag{5.61}$$

Example 5.8: Coherence from a small, round distributed source.

A 1 mm pinhole is placed immediately in front of an extended, spatially incoherent source of average wavelength 500 nm. The light passed by the pinhole is to be used in a diffraction experiment, for which it is desired to illuminate a distant 3 mm diameter aperture coherently. Calculate the minimum distance between the pinhole source and the diffracting aperture for good coherence. State any assumptions being made.

Assuming that the source distribution is uniform over the pinhole, a one-dimensional approxmation is $a^2(y_s) = \text{rect}(y_s/A)$. The coherence at the aperture is obtained by applying the van Cittert-Zernike Theorem. The visibility is proportional to the Fourier transform of the source distribution that is,

$$V = \mu_{12}\left(\frac{\theta_A}{\bar{\lambda}}\right) = \left|\text{sinc}\left(L\frac{\theta_A}{\bar{\lambda}}\right)\right| = \left|\text{sinc}\left(A\frac{d}{l\bar{\lambda}}\right)\right| , \tag{5.62}$$

where *A* is the pinhole diameter, *d* is the diameter of the distant aperture and *l* is the distance from the pinhole to the aperture. Assume illumination is coherent until the visibility goes to zero. The minimum distance for good coherence is obtained where the visibility is zero, which is

$$l = A\frac{d}{\bar{\lambda}} = \frac{(1\times10^{-3})(3\times10^{-3})}{0.5\times10^{-6}} = 6\ m . \tag{5.63}$$

5.1.8 Combination of Temporal and Spatial Coherence

The complete description of coherence properties resulting from a multiple-wavelength, extended source is provided in Section 5.2. However, it is possible to understand the basic character of coherence from a heuristic argument. For

example, consider the off-axis polychromatic point source pictured in the YDPI of Fig. 5.29. With reference to Fig. 5.20, maximum fringe visibility is observed at the point on the observation screen where $OPD = OPD'$. The decrease in V away from this point is governed by the bandwidth $\Delta\nu$ of the source. Narrow $\Delta\nu$ implies good V over a wide distance along y_0. The fringe pattern associated with the single source point is the point response of the YDPI system. Shifting the source point simply shifts the point response along the observation screen. In the central region of the observation screen, the relationship is linear and shift invariant in y_0. The relationship is always linear and shift invariant in OPD'.

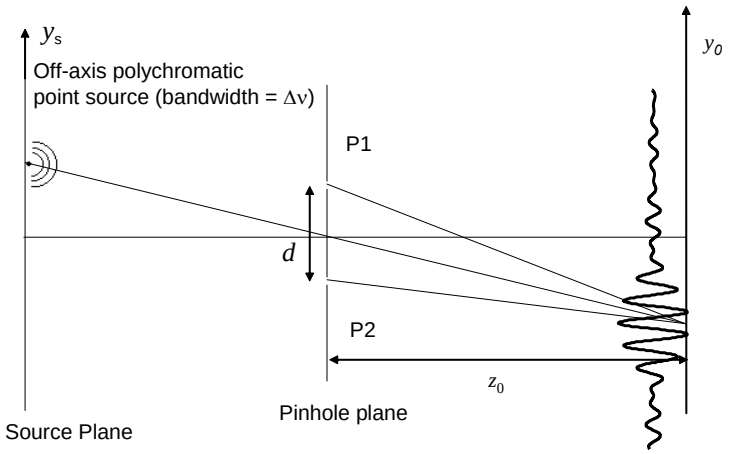


Figure 5.29. The point response of an off-axis polychromatic source.

If more than one source point illuminates the YDPI, the point response fringe patterns are summed to produce the total fringe pattern. In the limit of a continuous source distribution, the total fringe pattern is found by convolving the point response with the appropriately scaled source distribution.

For example, consider a source distribution that is a simple $\mathrm{rect}(y_s/L)$ centered about the optical axis. If L is small, the shifts of the point responses are less than the width of a fringe period. Therefore, the convolution results in a uniform loss factor in visibility across observation screen, but the basic shape of the point response remains. At the point on the screen where $OPD' = 0$, the loss in visibility is due only to the spatial extent of the source, and not the temporal bandwidth. As a very rough approximation, the total visibility is governed by the relationship

$$V \propto \mu_{12}\left(\frac{\theta_A}{\bar{\lambda}}\right) m_{12}\left(\frac{OPD^r}{c}\right) \ . \tag{5.64}$$

5.1.9 Terminology Used in Coherence Theory

Equation (5.4) combined with Eq. (5.6) can be written in the following form

$$I(y_0) = I_1(y_0) + I_2(y_0) + 2\sqrt{I_1 I_2}\,Re[\gamma_{12}(\tau)] \ , \tag{5.65}$$

where $I_1(y_0)$ and $I_2(y_0)$ irradiances at the observation point at y_0 from each pinhole independently, $\Gamma_{12}(\tau)$ is the *mutual coherence function* resulting from positions P_1 and P_2 in the light field and delayed by time $\tau = \Delta t$, and $\gamma_{12}(\tau)$ is the normalized *complex degree of coherence*, which is given by

$$\frac{K_1 K_2}{\sqrt{I_1 I_2}}\Gamma_{12}(\tau) \quad . \tag{5.66}$$

It is understood that $\tau = f(y_0)$. The complex degree of coherence is normalized, so that $0 \le |\gamma_{12}(\tau)| \le 1$. The term $\gamma_{12}(\tau)$ is called the degree of coherence. Equation (5.65) can also be written as

$$I(y_0) = I_1(y_0) + I_2(y_0) + 2\sqrt{I_1 I_2}\,|\gamma_{12}(\tau)|\cos[\alpha_{12}(\tau) - \delta] \ , \tag{5.67}$$

where

$$\alpha_{12}(\tau) = \arg[\gamma_{12}(\tau)] \ . \tag{5.68}$$

When working with temporal coherence problems, the complex degree of coherence is related to the terms in Section 5.1.4 by

$$m_{12}\left(\frac{OPD^r}{c}\right) = m_{12}(\tau) = |\gamma_{12}(\tau)| = \left|\frac{K_1 K_2}{\sqrt{I_1 I_2}}\Gamma_{12}(\tau)\right| \ , \tag{5.69}$$

and

$$\beta_{12}\left(\frac{OPD^r}{c}\right) = \beta_{12}(\tau) = \arg[\gamma_{12}(\tau)] \ . \tag{5.70}$$

When working with spatial coherence problems, the complex degree of coherence is related to the terms in Section 5.1.5 by

$$\mu_{12} = |\gamma_{12}(0)| = \left|\frac{K_1 K_2}{\sqrt{I_1 I_2}}\Gamma_{12}(o)\right| = \left|\frac{K_1 K_2}{\sqrt{I_1 I_2}}J_{12}\right| \ , \tag{5.71}$$

and

$$\beta_{12} = \arg[\gamma_{12}(0)] \ , \tag{5.72}$$

where

$$J_{12} = \Gamma_{12}(0) = \langle \tilde{E}_1(t)\tilde{E}_2^*(t)\rangle \tag{5.73}$$

is the *mutual intensity*. In this case, μ_{12} is the normalized mutual intensity. Notice that, if pinhole positions P_1 and P_2 are coincident, $J_{12} = \Gamma_{11}(0) = I_1$ is the *irradiance* at that point in the light field. Also, the *self coherence* of the light field is given by $\Gamma_{11}(\tau)$.