

3

FABRY–PÉROT RESONATORS, BEAMS, AND RADIATION PRESSURE

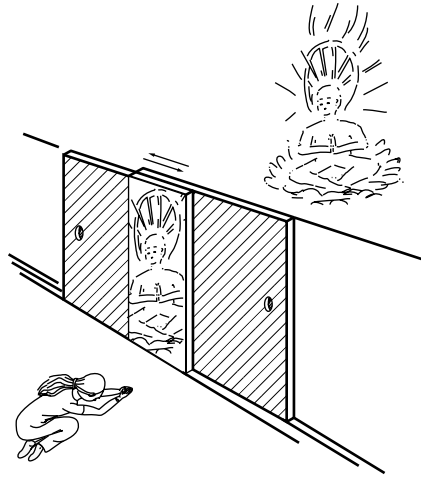
The two topics on Fabry–Pérot resonators and the properties of Gaussian beams complement each other because the Gaussian distribution function describes light beams not only in free space but also inside certain kinds of Fabry–Pérot resonators.

This chapter also covers Bessel beams, which are considered long-range nondiffracting beams, and radiation pressure, which is applied to light tweezers as well as laser cooling.

3.1 FABRY–PÉROT RESONATORS

Fabry–Pérot resonators selectively transmit or reflect a particular wavelength of light and are used for various applications, such as spectroscopy, stabilization of laser oscillation, and interference filters. Along with grating spectroscopy, Michelson interferometry, and Fourier transform spectroscopy, Fabry–Pérot spectroscopy is one of the most important means of analyzing a light spectrum.

Suppose that an individual plate is very opaque: let's say the reflectance R is 99.9%, and almost no light gets through. If two of these plates are placed precisely parallel to each other, something unexpected happens. The plates become transparent with almost nearly 100% transmittance at a particular wavelength, which is determined by the spacing between the reflecting plates, and at all other wavelengths, the plates will become even more reflective than before. The reason for this behavior is due to the large number of multiple reflections of the light inside the resonator. The plates are assumed to be nonabsorbing and are carefully aligned parallel to each other. The light bounces back and forth between the plates a large number of times, with only a small percentage of light being transmitted through the plate at each bounce. The larger the value of R is, the less light gets through the plate per bounce, and the greater the number of bounces. The transmitted light from each pass through the resonator interferes with the transmitted light from other passes. For the special case that the transmitted components from the multiple passes are in phase, the total amount of



Magic opaque screen. Opaque screens become transparent where they overlap.

transmitted light through the resonator becomes nearly 100%. This is known as the *resonance condition*. For the transmitted components to be in phase, the round-trip phase difference must be an even multiple of π radians. Wavelengths that satisfy this condition are called *resonant wavelengths*.

At resonance, the value of the peak transmittance is always nearly 100% regardless of the value of R . Although R does not affect the peak transmittance, the value of R does affect the sharpness of the resonance. The larger the value of R is, the greater the number of multiple reflections, and the more stringent the in-phase condition becomes. This phenomenon has made the Fabry-Pérot resonator quite popular as a device for selecting a particular wavelength of light. Thus, the Fabry-Pérot resonator transmits or reflects very selectively at particular wavelengths of light.

The most important application of the Fabry-Pérot resonator is as a tool for analyzing light spectra. It is also used inside gas as well as semiconductor lasers to host the action of lasing at a specified wavelength or as a component in a circuit for stabilizing the wavelength of the laser oscillation. Moreover, the principle of the Fabry-Pérot resonator is applied to designing such devices as an interference filter that can pick out an extremely narrow spectrum of light or a dichroic mirror whose optimum transmission and reflection take place at two specified wavelengths.

The Fabry-Pérot resonator is sometimes called the Fabry-Pérot etalon or simply etalon. By convention, resonators with a fixed spacing between reflectors are called etalons, and resonators with a variable spacing between reflectors are called interferometers.

3.1.1 Operating Principle of the Fabry-Pérot Resonator

The basic description of the operating principle of the Fabry-Pérot resonator is given in this section, followed by a more comprehensive analysis in the next section.

Figure 3.1a shows a schematic of a Fabry-Pérot resonator. It consists of left and right reflecting plates sandwiching a center medium. For simplicity, each reflecting plate has refractive index n_1 , and it is assumed that n_1 is larger than the refractive

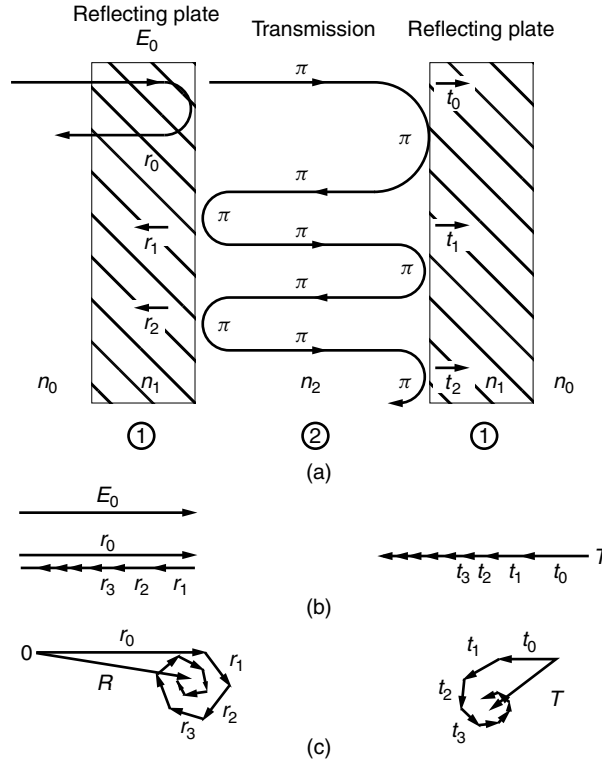


Figure 3.1 Principle of the Fabry-Pérot resonator. (a) Geometry of Fabry-Pérot resonator ($n_1 > n_2$). (b) Phasors at resonance. (c) Phasors off resonance.

index n_2 of the center medium, namely,

$$n_1 > n_2$$

Light with amplitude E_0 is incident from the left to the right and is normal to the reflecting plates. It is further assumed that the spacing between the reflecting plates is adjusted to be an integral multiple of a half-wavelength; namely, the Fabry-Pérot cavity is set at resonance.

The analysis will concentrate on the reflection and transmission at the boundaries between n_1 and n_2 . The reflection at the interface between the reflecting plate (n_1) and air (n_0) is not accounted for. First, consider the incident light undergoing reflection at the interface of the left reflecting plate and the center medium. The reflection coefficient r_{12} from Eq. (2.18) in Chapter 2 is

$$r_{12} = \frac{n_1 - n_2}{n_1 + n_2} \quad (3.1)$$

and r_{12} is a positive number. The phase of the reflected light $E_0 r_{12}$ is designated as 0 radians. Let this first reflected light component be called r_0 ,

$$r_0 = E_0 r_{12} \quad (3.2)$$

The transmission coefficient at the same boundary from Eq. (2.19) is

$$t_{12} = \frac{2n_1}{n_1 + n_2} \quad (3.3)$$

and t_{12} is also a positive number. The light incident into the center medium from the left reflecting plate is $E_0 t_{12}$, and its phase is designated as 0 radians.

The transmission coefficient t_{21} transmitting from the center medium into the right reflecting plate is

$$t_{21} = \frac{2n_2}{n_1 + n_2} \quad (3.4)$$

and t_{21} is a positive number. The light component t_0 emergent from the Fabry-Pérot resonator is

$$t_0 = E_0 t_{12} t_{21} e^{j(2m+1)\pi} \quad (3.5)$$

where m is an integer and the last factor accounts for the phase associated with transmission across the Fabry-Pérot resonator at the resonance condition, which is an odd multiple of π radians. Thus, the phase of the first transmitted component t_0 is π radians.

Next, the light component that is reflected back by the right reflector toward the left reflector is considered. The reflection coefficient r_{21} at the right reflecting plate is

$$r_{21} = \frac{n_2 - n_1}{n_1 + n_2} \quad (3.6)$$

and r_{21} is a negative number. As a complex number, r_{21} can be expressed as

$$r_{21} = |r_{21}| e^{j\pi} \quad (3.7)$$

Thus, there is a phase change of π associated with this reflection.

In simple terms, for each one-way trip through the center medium, the light acquires a phase of $(2m + 1)\pi$ radians, which is equivalent to π radians. There is also a phase change of π radians for each reflection occurring when the light in the center medium hits the reflector boundary, as illustrated in Fig. 3.1a.

The transmitted light t_1 is

$$t_1 = E_0 t_{12} e^{j(2m+1)\pi} r_{21} e^{j(2m+1)\pi} r_{21} e^{j(2m+1)\pi} t_{21}$$

which reduces to

$$t_1 = E_0 t_{12} r_{21}^2 t_{21} e^{j\pi}$$

Components of the transmitted light, t_0, t_1, t_2, \dots , all have the same π -radian phase and add constructively. The transmission through the cavity reaches a maximum at resonance.

On the other hand, the situation of the components r_0, r_2, r_3, \dots of the light reflected from the cavity is slightly different. The magnitude of the first reflected light component, which has never entered the cavity, is $E_0 r_{12}$ and is much larger than all other reflected components. Table 3.1 summarizes the amplitudes of the various components shown in Fig. 3.1a. The absolute value signs in r_{ij} are suppressed.

Table 3.1 Phase and amplitude of reflected and transmitted components

Reflected Components	Phase	Amplitude	Transmitted Components	Phase	Amplitude
r_0	0	$E_0 r_{12}$	t_0	π	$E_0 t_{12} t_{21}$
r_1	π	$E_0 t_{12} r_{21} t_{21}$	t_1	π	$E_0 t_{12} r_{21}^2 t_{21}$
r_2	π	$E_0 t_{12} r_{21}^3 t_{21}$	t_2	π	$E_0 t_{12} r_{21}^4 t_{21}$
r_n	π	$E_0 t_{12} r_{21}^{2n-1} t_{21}$	t_n	π	$E_0 t_{12} r_{21}^{2n} t_{21}$

The amplitude of the n th component decreases as n increases. The magnitude of the first reflected light r_0 has a phase of 0 radians, while all other smaller components r_1, r_2, r_3, \dots , which eventually come out of the cavity after multiple reflections inside the cavity, all have π -radian phase. The largest component and the accumulation of all the smaller components cancel each other, and the resultant reflected light reduces to a minimum at the resonance of the cavity. It is a common mistake to forget about r_0 , and to incorrectly conclude that the reflection also reaches a maximum at resonance due to the accumulation of r_1, r_2, r_3, \dots . The phasors at resonance are shown in Fig. 3.1b.

Finally, the case of off resonance is considered. Off resonance, the phase delay due to the round trip inside the cavity is no longer exactly an integral number of 2π radians. The components of the signal transmitted through the cavity are no longer in phase and the phasors representing them curl up as shown in Fig. 3.1c. The magnitude of the resultant phasors of the transmitted light from the cavity is small.

The phasors of the r_1, r_2, r_3, \dots light components reflected from inside the cavity also curl up, and the resultant phasor can no longer cancel the large phasor r_0 of the first reflected component that did not enter the cavity. Thus, the reflectance of the cavity reaches a large value when the cavity is not at resonance, as shown by the phasor R on the left side of Fig. 3.1c. In short, the transmitted light decreases off resonance while the reflected light increases.

In the next section, the case when the angle of incidence is not normal to the reflectors is described.

3.1.2 Transmittance and Reflectance of the Fabry-Pérot Resonator with an Arbitrary Angle of Incidence

Referring to the geometry of Fig. 3.2, the general expressions for the transmittance and reflectance of a Fabry-Pérot resonator [1–3] will be derived. These quantities are obtained by summing an array of light beams produced by multiple reflections from a pair of reflectors and reaching point P of the focus of a convex lens. Let the refractive index of the medium of the Fabry-Pérot etalon be n_2 , the refractive index of the external medium be n_1 , and the spacing between the reflectors be d . There is no reflective film deposited on the etalon, and the reflectors are the discontinuities of the refractive indices between n_1 and n_2 . The cases for $n_1 > n_2$ and $n_2 > n_1$ will be treated concurrently. As a matter of fact, the final results are the same for both cases. Also, let the angle of incidence from the normal to the reflector surface be θ_i and the internal angle inside the reflectors be θ . Snell's law gives the relationship between θ and θ_i as

$$n_1 \sin \theta_i = n_2 \sin \theta \quad (3.8)$$

The emergent angle from the Fabry-Pérot resonator is again θ_i .

assumed, then 2π disappears. Either way, 2π does not influence the subsequent result. Removing the common path and noting that $\overline{AB} = \overline{BC}$, ϕ becomes the optical path difference between \overline{ABC} and \overline{AD}

$$\phi = k(2n_2\overline{AB} - n_1\overline{AD}) \quad (3.10)$$

$$\overline{AB} = \frac{d}{\cos \theta}$$

$$\overline{AD} = \overline{AC} \sin \theta_i \quad (3.11)$$

$$\overline{AC} = 2d \tan \theta \quad (3.12)$$

$$\phi = k \left(\frac{2n_2d}{\cos \theta} - 2n_1d \tan \theta \sin \theta_i \right) \quad (3.13)$$

where θ and θ_i are related by Snell's law. Equation (3.8) and ϕ finally become

$$\phi = 2n_2dk \cos \theta \quad (3.14)$$

Next, the array of transmitted beams will be summed to obtain the resultant of the light reaching point P through the Fabry-Pérot resonator. The phases of the multiple reflected beams t_1, t_2, \dots, t_n are always compared to the phase of the first beam t_0 . In order to reach P , optical paths $\overline{C'P}$ and $\overline{D'P}$ are identical and the phase of t_2 at point C' is compared to that at D' on t_0 . The media outside and inside the resonator will be denoted by subscripts 1 and 2, respectively. The following notations are used:

- E_0 Amplitude of the incident light
- t_{12} Amplitude transmission coefficient from medium ① to ②
- t_{21} Amplitude transmission coefficient from medium ② to ①
- r_{12} Amplitude reflection coefficient when the light is incident from medium ① toward medium ②
- r_{21} Amplitude reflection coefficient from medium ② toward medium ①
- t_i Amplitude of the transmitted light component after the i th reflection from the top boundary
- r_i Amplitude of the reflected light component after the $(i + 1)$ th reflection from the top boundary
- Δ Phase associated with the optical path length from point O to A
- R Reflectance of the reflector, defined as $r_{21}r_{21}$
- T Transmittance of the reflector, defined as $t_{12}t_{21}$
- ϕ Phase difference due to the optical path difference between \overline{ABC} and \overline{AD} given by Eq. (3.13)

The elements of the transmitted beam are

$$\begin{aligned} t_0 &= E_0 t_{12} e^{j\Delta} t_{21} \\ t_1 &= E_0 t_{12} e^{j\Delta} r_{21} r_{21} e^{j\phi} t_{21} \\ t_2 &= E_0 t_{12} e^{j\Delta} r_{21}^2 r_{21}^2 e^{j2\phi} t_{21} \\ t_3 &= E_0 t_{12} e^{j\Delta} r_{21}^3 r_{21}^3 e^{j3\phi} t_{21} \\ t_n &= E_0 t_{12} e^{j\Delta} r_{21}^{2n} e^{jn\phi} t_{21} \end{aligned} \quad (3.15)$$

The sum E_t of the transmitted light is

$$E_t = E_0 e^{j\Delta} T (1 + R e^{j\phi} + R^2 e^{j2\phi} + R^3 e^{j3\phi} + \dots) \quad (3.16)$$

where

$$R = r_{21} r_{21} \quad (3.17a)$$

$$T = t_{12} t_{21} \quad (3.17b)$$

$$E_t = E_0 e^{j\Delta} \frac{T}{1 - R e^{j\phi}} \quad (3.18)$$

If the reflector is lossy and its loss A is defined as

$$T + R + A = 1 \quad (3.19)$$

then Eq. (3.18) can be written as

$$E_t = E_0 e^{j\Delta} \frac{1 - R - A}{1 - R e^{j\phi}} \quad (3.20)$$

Similarly, the amplitude of the light reflected from the Fabry-Pérot resonator will now be calculated. The phase of the reflected beam r_i is always referred to that of beam r_0 . The reflected light components are

$$\begin{aligned} r_0 &= E_0 r_{12} \\ r_1 &= E_0 t_{12} r_{21} e^{j\phi} t_{21} \\ r_2 &= E_0 t_{12} r_{21}^3 e^{j2\phi} t_{21} \\ r_3 &= E_0 t_{12} r_{21}^5 e^{j3\phi} t_{21} \\ r_n &= E_0 t_{12} r_{21}^{2n-1} e^{jn\phi} t_{21} \end{aligned} \quad (3.21)$$

where ϕ is the same as before since

$$\overline{OE} = \overline{AD} \quad (3.22)$$

The total sum E_r of the light reflected toward the source is

$$\begin{aligned} E_r &= E_0 [r_{12} + T r_{21} e^{j\phi} (1 + r_{21}^2 e^{j\phi} + r_{21}^4 e^{j2\phi} + \dots)] \\ &= E_0 \left(r_{12} + \frac{T r_{21} e^{j\phi}}{1 - R e^{j\phi}} \right) \end{aligned} \quad (3.23)$$

From Eqs. (3.1) and (3.6) for normal incidence, it is easy to see that

$$r_{12} = -r_{21} \quad (3.24)$$

which also holds true for either case of $n_1 > n_2$ or $n_2 < n_1$.

Equations (3.17a) and (3.24) give

$$\left. \begin{aligned} r_{12} &= \sqrt{R} \\ r_{21} &= -\sqrt{R} \end{aligned} \right\} \quad (3.25)$$

where r_{12} is a positive number and r_{21} is a negative number for $n_1 > n_2$, and the signs are reversed for $n_1 < n_2$.

Using Eqs. (3.19), (3.23), and (3.25), the expression for E_r is

$$E_r = \pm E_0 \sqrt{R} \frac{1 - (1 - A)e^{j\phi}}{1 - Re^{j\phi}} \quad (3.26)$$

The upper sign is for $n_1 > n_2$, and the lower for $n_1 < n_2$.

Next, the intensity of the transmitted light I_t will be calculated from Eq. (3.20) using the relationship $I_t = E_t E_t^*$:

$$I_t = E_0^2 \frac{(1 - R - A)^2}{(1 - Re^{j\phi})(1 - Re^{-j\phi})} \quad (3.27)$$

$$= I_0 \frac{(1 - R - A)^2}{(1 - R)^2 + 4R \sin^2(\phi/2)} \quad (3.28)$$

Similarly, the intensity of the reflected light is

$$I_r = I_0 R \frac{A^2 + 4(1 - A) \sin^2(\phi/2)}{(1 - R)^2 + 4R \sin^2(\phi/2)} \quad (3.29)$$

Equations (3.28) and (3.29) hold true for either $n_1 > n_2$ or $n_1 < n_2$.

Let's look more closely at the transmitted power. By dividing both denominator and numerator by $(1 - R)^2$, Eq. (3.28) can be rewritten as

$$I_t = I_0 \left(1 - \frac{A}{1 - R} \right)^2 \frac{1}{1 + M \sin^2(\phi/2)} \quad (3.30)$$

where

$$M = \frac{4R}{(1 - R)^2} \quad (3.31)$$

and ϕ is given by Eq. (3.14).

The transmitted power I_t is periodic with respect to ϕ and every time ϕ approaches

$$\phi_m = 2m\pi \quad (3.32)$$

the value of $M \sin^2(\phi/2)$ in Eq. (3.30) becomes zero and I_t reaches a peak value (resonance). When $\phi_m = 2\pi(m + \frac{1}{2})$, I_t reaches a valley (antiresonance). The response curve is shown in Fig. 3.3. The larger the value of M is, the greater the change of

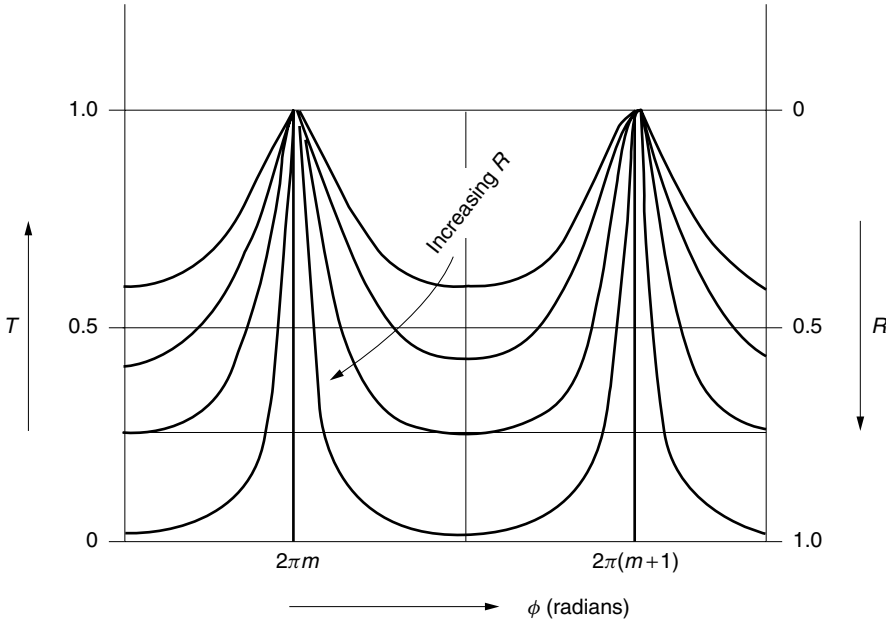


Figure 3.3 Transmittance T and reflectance R of the Fabry-Pérot resonator with respect to ϕ .

$M \sin^2(\phi/2)$ with respect to ϕ , and the sharper the resonance peaks become. For some applications, such as in a monochromator or a laser cavity, the sharper peaks are more desirable, whereas broader peaks are more desirable for other applications. In an optical filter, the bandwidth is manipulated by the value of M .

In monochromator applications, a large value of R is needed, but a sufficient amount of reflection cannot be realized solely by the difference in the indices of refraction n_1 and n_2 . A thin metal film is deposited on the inner surfaces of the reflector plates in order to reach the needed value of R . The presence of the metal film can be accounted for with phase $e^{j\gamma}$, loss A , and the new value of R . In the case of the metal film, the phase change of Eq. (3.7) between r_{12} and r_{21} is no longer π radians but has a phase of γ . The energy lost in the metal film is accounted for by A . With these modifications, the above analysis holds true for practical purposes.

Important facts about the Fabry-Pérot resonator are that if $A = 0$, regardless of the value of R , transmission is 100% at resonance, as seen from Eq. (3.30), and moreover, the values of ϕ_m do not depend on the value of R . These two facts are the very reasons why the Fabry-Pérot resonator is so useful as a device for spectroscopy.

The results obtained so far have been derived from a scalar field approach, meaning the direction of the light polarization has not been taken into consideration. As was shown in the previous chapter, the amplitude and phase of the reflection coefficient r depend on the direction of polarization. Fortunately, however, the value of ϕ_m is independent of R and the use of the scalar field approach is justifiable.

On the other hand, if the refractive index n_2 of the medium inside the Fabry-Pérot cavity depends on the direction of light polarization (namely, n_2 is birefringent), then

ϕ_m is polarization dependent, and the vector field approach has to be used [1]. If the medium inside the Fabry-Pérot cavity is a Faraday rotator (see Chapter 5), such as iron garnet, then the multiple reflections taking place inside the Fabry-Pérot resonator enhance the rotation [4]. The measured result becomes dependent on the state of light polarization.

As seen from Eq. (3.14), there are several physical means of sweeping the value of ϕ , such as by changing the (1) spacing d between the reflectors, (2) the internal angle θ , (3) the refractive index of the medium inside the cavity, and (4) the wavelength of the incident light.

Fabry-Pérot devices are best categorized by the means used for sweeping the value of ϕ . These categories, in direct correspondence to the above list of sweeping methods, are (1) the scanning Fabry-Pérot spectrometer, (2) the Fabry-Pérot etalon, (3) the liquid crystal filter, and (4) the laser frequency stabilizer.

Each of the above will be described in more detail in the following sections.

3.2 THE SCANNING FABRY-PÉROT SPECTROMETER

A scanning Fabry-Pérot spectrometer consists of these basic parts: a scanning Fabry-Pérot resonator, a voltage generator, a photodetector, and a display scope. A photograph of a scanning Fabry-Pérot resonator is shown in Fig. 3.4. The resonator has a piezotransducer (PZT), which is driven by a sawtooth voltage generator. As the mirror of the resonator is displaced by the PZT, the transmitted light is detected and displayed on the scope. The block diagram of the scanning Fabry-Pérot spectrometer is shown in Fig. 3.5.

The scanning Fabry-Pérot resonator consists of one stationary and one movable plate. The stationary plate is equipped with a parallelism adjustment assembly onto

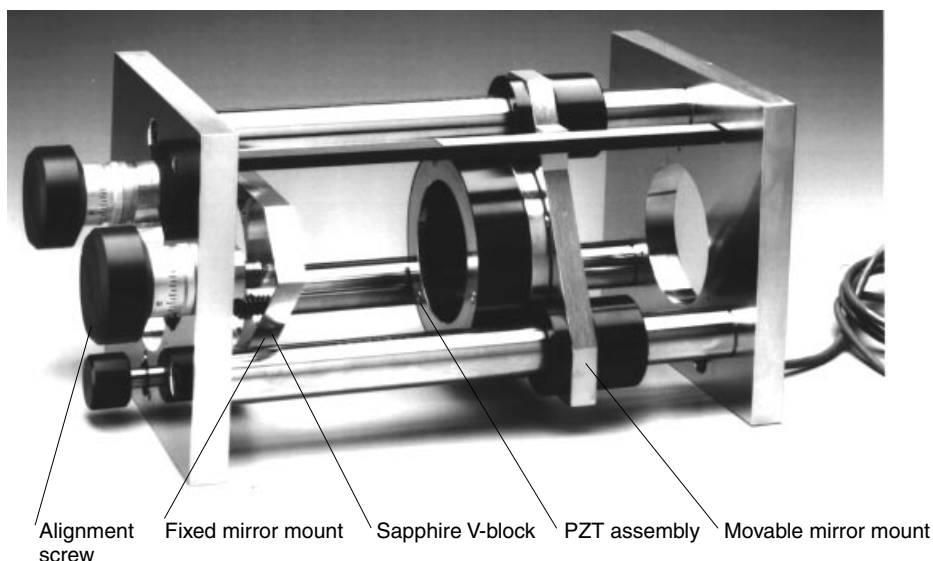


Figure 3.4 Structure of a scanning Fabry-Pérot resonator. (Courtesy of Burleigh Instruments Inc.)

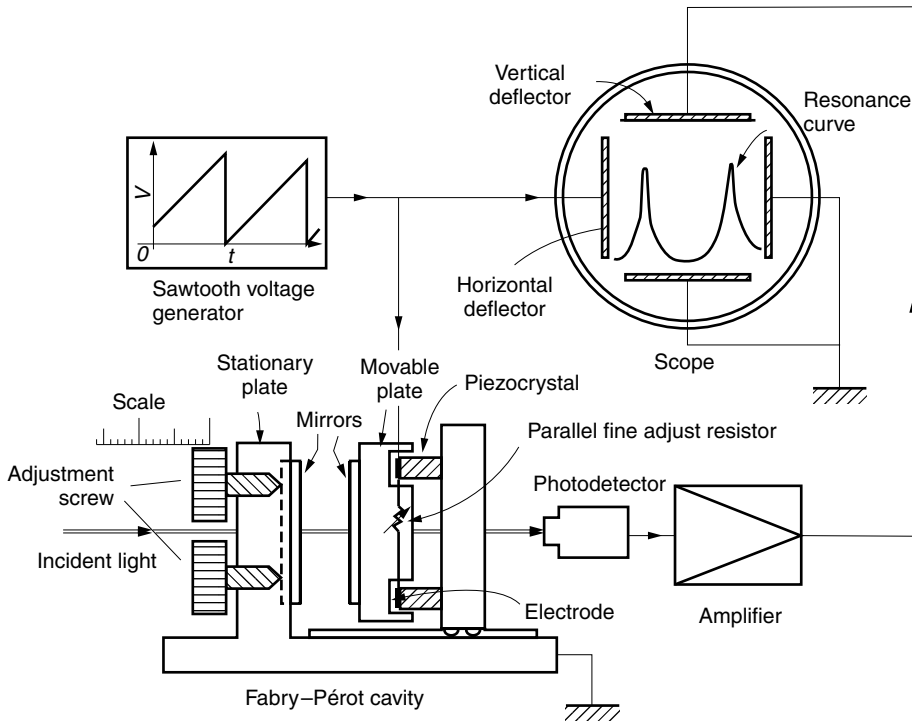


Figure 3.5 Block diagram of the scanning Fabry-Pérot spectrometer.

which one of the two mirrors is fastened. The movable plate can be coarsely set at a distance typically between $50\ \mu\text{m}$ and $15\ \text{cm}$ from the stationary plate. A PZT assembly, which holds the other mirror, is installed on the moving plate. Thus, the spacing between the mirrors can be coarsely set between $50\ \mu\text{m}$ to $15\ \text{cm}$ and the spacing can be finely scanned for about $2\text{--}3\ \mu\text{m}$ by the PZT from the coarsely set position.

The output from the scanning Fabry-Pérot resonator is detected by either a photomultiplier or a photodiode and amplified. The amplified output signal is displayed on an oscilloscope. The output from the sawtooth voltage generator is fed to both the PZT and the horizontal deflection plate of the oscilloscope. Thus, the light output from the scanning Fabry-Pérot device is displayed with respect to displacement of the movable mirror.

The performance of the Fabry-Pérot resonator is characterized quantitatively by two factors: the finesse and the free spectral range. The finesse relates to the sharpness of the resonance, and the free spectral range is approximately equal to the separation between neighboring resonances. This separation can be expressed in terms of wavelength, frequency, or wavenumber (reciprocal of wavelength).

3.2.1 Scanning by the Reflector Spacing

In the case of normal incidence, $\theta = 0^\circ$, into an air-filled Fabry-Pérot cavity with refractive index $n_2 = 1$, the condition for the m th order resonance is, from Eqs. (3.14)

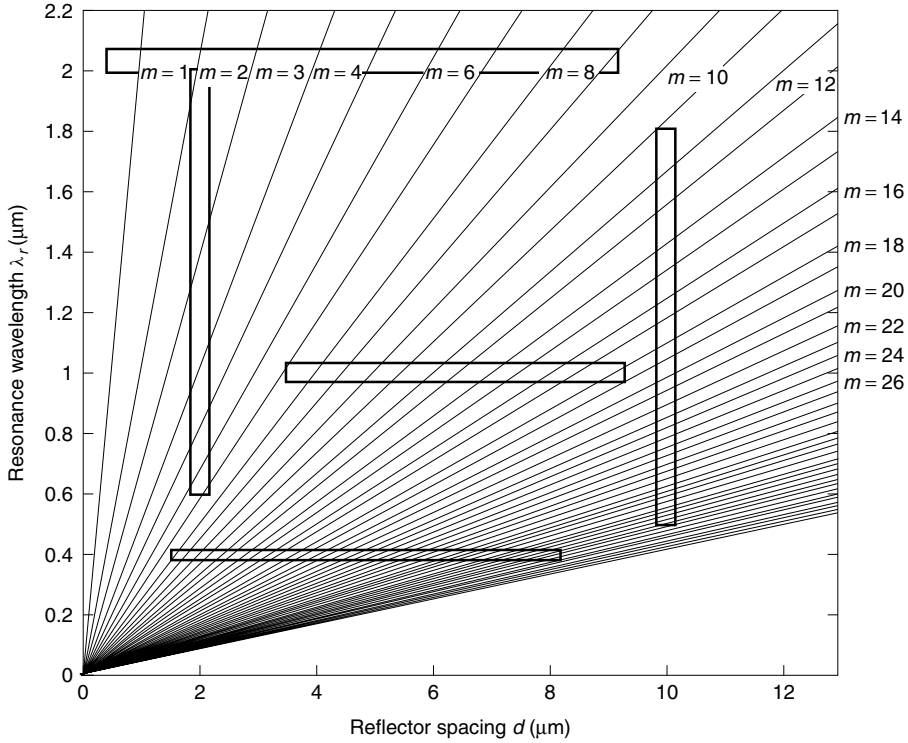


Figure 3.6 Resonance wavelength with respect to the reflector spacing with the mode number m as a parameter: $(\lambda_r/2)m = d$.

and (3.32),

$$\frac{\lambda_r}{2}m = d \quad (3.33)$$

The performance of the Fabry-Pérot resonator is controlled by this equation. Figure 3.6 is a plot of Eq. (3.33) and shows the resonance wavelength λ_r as a function of reflector spacing with the mode number m as a parameter. Equation (3.33) looks quite simple, but there are three variables, and interpretation of Fig. 3.6 needs further explanation. A small section near the origin of Fig. 3.6 has been expanded for detailed illustration in Fig. 3.7, which is described next.

Referring to Fig. 3.6, the intersections between a horizontal line and the resonance curves are equally spaced. For example, if a horizontal line is drawn at $\lambda = 0.4 \mu\text{m}$, the intersections are $0.2 \mu\text{m}$ apart. If a horizontal line is drawn higher on the graph, say, at $1.0 \mu\text{m}$, the intersections are still equally spaced but are now $0.5 \mu\text{m}$ apart. This is why peaks corresponding to a specific wavelength appear equally spaced in the output display of a scanning Fabry-Pérot interferometer. The intersections between the vertical line and the resonance curves are not equally spaced. The spacing between intersections decreases gradually as the resonance wavelength decreases. In the shorter resonance wavelength regions, the spacings are approximated as uniform for most practical purposes.

3.2.1.1 Fabry-Pérot Resonator with a Fixed Resonator Spacing (Etalon)

Figure 3.7a shows the graph with the reflector spacing fixed at $d = 2 \mu\text{m}$. The resonance takes place at every mode number. The resonance wavelengths for the first three mode numbers are $\lambda_r = 4 \mu\text{m}$ for $m = 1$, $\lambda_r = 2 \mu\text{m}$ for $m = 2$, and $\lambda_r = 1.33 \mu\text{m}$ for $m = 3$. The graph on the left shows the output spectra when light with a continuum spectrum is incident on the resonator. The light field distribution in the resonator for each mode number is shown on the right. The mode of the light spectrum from a Fabry-Pérot cavity-type laser is a good example of this case.

The spacing $\Delta\lambda_r$ between the resonator wavelengths for a large m is

$$\Delta\lambda_r = \frac{2d}{m} - \frac{2d}{m+1} \div \frac{2d}{m^2} = \frac{\lambda^2}{2d} \quad (3.34)$$

3.2.1.2 Monochromatic Incident Light with Scanned Reflector Spacing

Figure 3.7b shows the graphs associated with a single spectrum input light with wavelength $\lambda = 2.0 \mu\text{m}$ and angle of incidence $\theta_i = 0^\circ$. The reflector spacing is swept from 0 to $3 \mu\text{m}$. When the spacing reaches $d = 1 \mu\text{m}$, the resonance wavelength $\lambda_r = 2 \mu\text{m}$ of the resonator matches the incident light wavelength. The light is transmitted through the resonator and the first output peak appears. The output is shown in the bottom graph of Fig. 3.7b. The first peak is associated with the $m = 1$ mode. As d is swept further, the $m = 2$ mode resonance wavelength matches that of the incident light and the second peak appears at $d = 2 \mu\text{m}$. The same will be repeated as d is scanned at every $\lambda/2$ or $1 \mu\text{m}$.

At the top of Fig. 3.7b, the light field E inside the Fabry-Pérot resonator is drawn for each mode. Every time the cavity length reaches an integral multiple of a half-wavelength, resonance takes place. As a matter of fact, from the interval between the resonance peaks, the wavelength of the incident light can be determined.

3.2.1.3 Free Spectral Range (FSR)

When the incident light has more than one wavelength, especially when the wavelengths are either too far apart or too close to each other, the proper selection of d , (hence m), becomes important.

Let us take the specific case of an incident light spectrum consisting of a main peak at $\lambda = 2.0 \mu\text{m}$. The wavelength of the auxiliary peak $\lambda_a = \lambda + \Delta\lambda$ is larger than the main peak by 20%. The scanning Fabry-Pérot resonator is used to analyze this compound signal. The response of $\lambda = 2.0 \mu\text{m}$ alone has already been shown in Fig. 3.7b. While peaks associated with λ appear at an interval of $\lambda/2$ as d is scanned, the peaks associated with $\lambda + \Delta\lambda$ appear at a longer interval of $(\lambda + \Delta\lambda)/2$. Two sets of peaks with different intervals are displayed as d is scanned. The response of both λ and λ_a is shown at the bottom of Fig. 3.7c as the solid and dashed lines, respectively. The optimum region of m (or d) for displaying the spectrum will be sought.

From the bottom graph in Fig. 3.7c, the separation of the peak of λ_a from that of λ becomes larger and larger as m (or d) is increased, and the resolution between the two spectral lines increases with m (or d). There is, however, a limit on the value of m (or d). As the value of m exceeds $m = 5$ or $d = 5 \mu\text{m}$, the shift of the peak of λ_a from that of λ becomes equal to or longer than the regular interval of $1 \mu\text{m}$ of the peaks of λ . The peak of λ_a , which belongs to $m = 5$, overlaps the peak of λ , which belongs to $m = 6$, and the peaks of λ and λ_a start to intermingle. The mode pattern at $m = 7$ restarts that of $m = 1$ and goes back to the case of the lowest resolution.

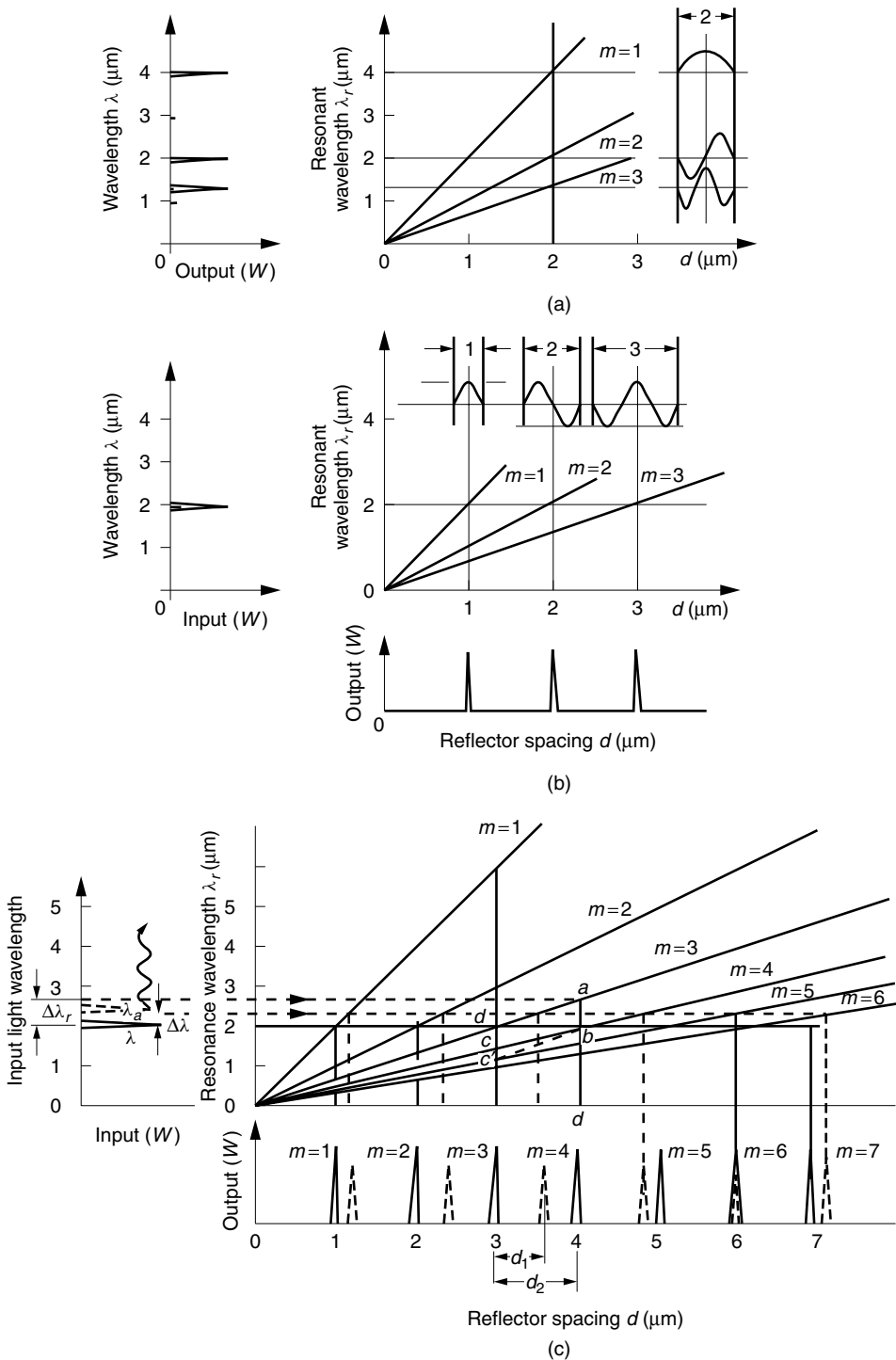


Figure 3.7 Expanded graph of Fig. 3.6. (a) Continuum spectrum light into a fixed reflector spacing. (b) Monochromatic light into scanned reflector spacing. (c) Illustration of how to select the reflector spacing.

In conclusion, the best choice of m is $4 \leq m \leq 5$, not only from the viewpoint of the highest resolution but also avoiding ambiguity.

Next, the question is asked in reverse. What is the maximum spread of the spectrum that can be determined by a Fabry-Pérot resonator without ambiguity for a given wavelength, resonator spacing, and scan range?

Let us say $d = 3 \mu\text{m}$ and $\lambda = 2 \mu\text{m}$ are given. This example calculates the maximum deviation $\Delta\lambda$ of the wavelength from λ that can be determined by the Fabry-Pérot resonator.

The order of mode m of the output peak appearing at $d = 3 \mu\text{m}$ is, from Eq. (3.33),

$$m = \frac{2d}{\lambda} = \frac{2(3)}{2} = 3$$

The output peak of the auxiliary wavelength, $\lambda + \Delta\lambda$, shifts from that of the third order mode of λ toward the fourth order mode of λ as $\Delta\lambda$ is increased. The value $\Delta\lambda$ for which the auxiliary output peak reaches that of the fourth order mode of λ is

$$m \frac{(\lambda + \Delta\lambda)}{2} = (m + 1) \frac{\lambda}{2} \quad (3.35)$$

With $m = 3$, the answer is

$$\Delta\lambda = 0.67 \mu\text{m}$$

The maximum width of the spectrum that can be determined without ambiguity is called the *free spectral range* (FSR). The FSR of the Fabry-Pérot resonator whose reflector spacing is $3 \mu\text{m}$ is $0.67 \mu\text{m}$ at the wavelength $\lambda = 2.0 \mu\text{m}$.

From Eqs. (3.33) and (3.35), the value of the FSR for the general case with spacing d , around the wavelength λ , is

$$\Delta\lambda_{\text{FSR}} = \frac{\lambda^2}{2d} \quad (3.36)$$

The FSR decreases with an increase in the reflector spacing.

Next, it will be shown that the free spectral range $\Delta\lambda_{\text{FSR}}$ is approximately equal to the spacing between the adjacent resonance wavelengths of a resonator with a fixed reflector spacing d and operated at λ . Referring to Fig. 3.7c, as $\Delta\lambda$ is increased from zero to $\Delta\lambda_r$, the intersection of the horizontal dotted line with \overline{da} moves from d to a . Line \overline{ab} represents the free spectral range $\Delta\lambda_{\text{FSR}}$; and line \overline{dc} represents the wavelength spacing between the $m = 3$ and $m = 4$ resonant wavelengths for a fixed value of $d = 3 \mu\text{m}$. From the parallelogram $abcd$, \overline{ab} can be approximated as \overline{cd} . Thus, the free spectral range is almost equal to the spacing between the resonant wavelengths for a fixed value of d . This approximation becomes better as m gets larger.

$$\Delta\lambda_{\text{FSR}} \doteq \Delta\lambda_r \quad (3.37)$$

It is sometimes useful to express the free spectral range in terms of the light carrier frequency. In terms of frequency, the free spectral range is

$$\Delta\nu_{\text{FSR}} = \frac{c}{\lambda} - \frac{c}{\lambda + \Delta\lambda_{\text{FSR}}} \doteq \frac{c}{\lambda^2} \Delta\lambda_{\text{FSR}} \quad (3.38)$$

With Eq. (3.36), Eq. (3.38) becomes

$$\Delta\nu_{\text{FSR}} = \frac{c}{2d} \quad (3.39)$$

Thus, it should be noted that $\Delta\nu_{\text{FSR}}$ is independent of the light frequency. The appropriate selection of the spacing of the mirrors, which determines the value of the FSR, is important for operating with the highest resolution, yet without ambiguity.

Example 3.1 Find the spacing d and the mode number of a scanning Fabry-Pérot spectrum analyzer that can display the spectrum of a superluminescent laser diode (SLD). The center wavelength of the SLD is $1.540 \mu\text{m}$ and the width of the spectrum is $\pm 10 \text{ nm}$ from the center wavelength.

Solution The free spectral range in wavelength has to be 20 nm and, from Eq. (3.36),

$$d = \frac{(1.54)^2}{2(0.02)} = 59.3 \mu\text{m}$$

The mode number of operation is

$$m = \frac{2d}{\lambda} = \frac{2(59.3)}{1.54} = 77 \quad \square$$

From the output display of a Fabry-Pérot resonator, such as shown at the bottom of Fig. 3.7c or in Fig. 3.8 of the next example, the input spectrum is to be determined. In the output display, the peak associated with λ_a is located at $(d_1/d_2)\lambda_{\text{FSR}}$ from one of the main peaks associated with λ , as shown at the bottom of Fig. 3.7c:

$$(\lambda_a - \lambda) = \frac{d_1}{d_2} \Delta\lambda_{\text{FSR}} \quad (3.40)$$

where $\Delta\lambda_{\text{FSR}}$ is the free spectral range given by Eq. (3.36). In terms of frequency, the equivalent of Eq. (3.40) is

$$(\nu - \nu_a) = \frac{d_1}{d_2} \Delta\nu_{\text{FSR}} \quad (3.41)$$

Example 3.2 shows the calculation for a specific case.

Example 3.2 Figure 3.8 shows the display of a scanning Fabry-Pérot resonator when a helium-neon laser beam is phase modulated by an MNA (2-methyl-4-nitroaniline) crystal. The wavelength is $\lambda = 0.6328 \mu\text{m}$ and the mirror spacing is swept around $d = 600 \mu\text{m}$.

- What is the scan length of the PZT from point a to point b in the display in Fig. 3.8?
- From Fig. 3.8, find the frequency of the phase modulation of the He-Ne laser light due to the MNA crystal. The phase modulation creates two side lobes centered at the carrier frequency.

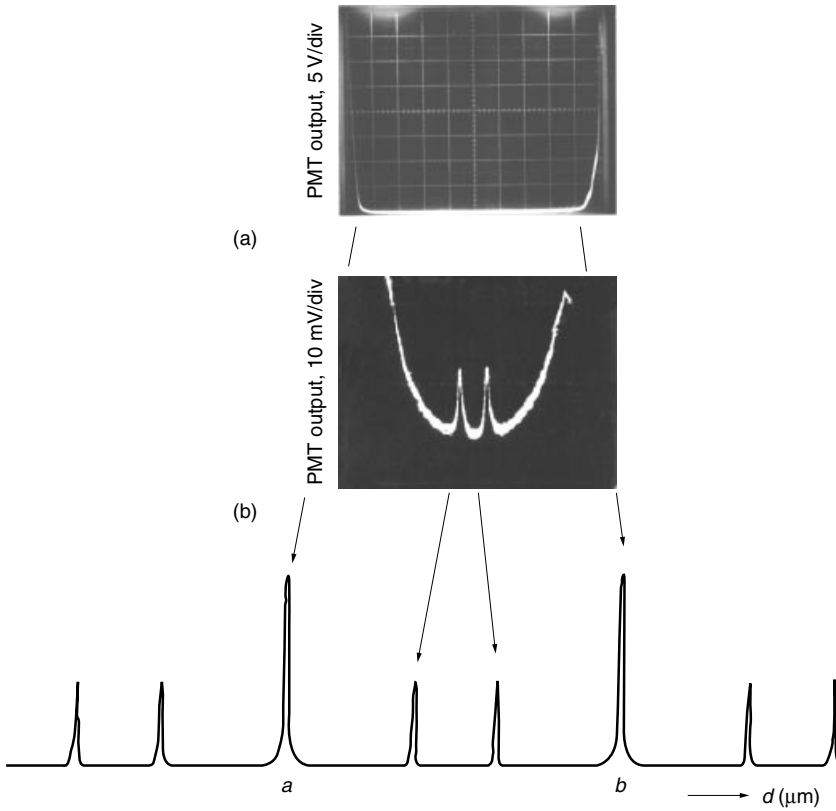


Figure 3.8 Display of a scanning Fabry-Pérot resonator with $d = 600 \mu\text{m}$ and $\lambda = 0.6328 \mu\text{m}$. (After C. Wah, K. Iizuka, and A. P. Freundorfer [15].) (a) Output on the oscilloscope at a vertical scale of 5 V/div; only the carrier is visible. (b) Output on the oscilloscope at a vertical scale of 10 mV/div; the first order sidebands appear.

Solution

(a) From the resonance condition

$$\frac{\lambda_r}{2} m = d_m$$

the scan length between adjacent modes is $\lambda/2$:

$$\frac{\lambda_r}{2} = \frac{0.6328}{2} = 316 \text{ nm}$$

(b) The solution is obtained by referring to Fig. 3.7b and Eq. (3.36). The FSR in wavelength can be converted into that in frequency using Eq. (3.39):

$$\Delta \nu_{\text{FSR}} = \frac{c}{2d}$$

$$\Delta \nu_{\text{FSR}} = \frac{3 \times 10^{14}}{2(600)} = 250 \text{ GHz}$$

From the display, the ratio of the distance to the side lobe to that of the adjacent main peak is measured to be 0.38; hence, the frequency of the phase modulation is $250 \times 0.38 = 95$ GHz. \square

Example 3.3 As shown in Part I of Fig. 3.9, the amplitude of the input spectrum is of a triangular shape with respect to the wavelength λ . This spectrum is observed by a Fabry–Pérot resonator.

- The spacing d between the reflectors is set so that the free spectral range fits the spectral width of the input light. Draw the amplitude of the spectrum when the spacing d is swept by $\lambda/2$.
- Draw the display when the spacing of the reflector is expanded to $2d$, and d is swept by $\lambda/2$.
- Draw the display when the spacing of the reflector is reduced to $d/2$ and d is swept by $\lambda/2$.

Solution The mode lines that are inside the $\Delta\lambda \times (\lambda/2)$ windows are responsible for the output. The answers are shown in Part II of Fig. 3.9.

- When $\Delta\lambda = \Delta\lambda_{\text{FSR}}$, the appropriate output display is obtained.
- When $\Delta\lambda > \Delta\lambda_{\text{FSR}}$, more than one mode line are inside the $\Delta\lambda \times (\lambda/2)$ window. The output displays overlap and the display becomes ambiguous.
- When $\Delta\lambda < \Delta\lambda_{\text{FSR}}$, even though the display is correct, the power of resolution is not the highest.

The lesson to be learned from this example is that when the spectral range of the input light is not known, it is good practice to start the measurement with a small value of d (or a large value of $\Delta\lambda_{\text{FSR}}$) and gradually increase d for improved resolution until overlapping output starts to display. \square

3.2.2 Scanning by the Angle of Incidence

As seen from Eq. (3.14), as θ is increased, ϕ is decreased, and the Fabry–Pérot resonator can be swept by sweeping θ_i . Figure 3.10 shows an arrangement for displaying the resonance rings without any moving mechanism. A diffuser is placed between the source and the Fabry–Pérot resonator. First, the case of monochromatic incident light will be considered. A diffuser creates wavelets of various incident angles to the Fabry–Pérot resonator. Among the incident wavelets, only those whose incident angles match the resonance angles $\theta_m, \theta_{m-1}, \theta_{m-2}, \dots, \theta_{m-i}$ that satisfy Eqs. (3.14) and (3.32) can transmit through the cavity. All other incident angles are reflected back by the Fabry–Pérot cavity toward the source.

The resonance condition is not limited to just the wavelets propagating in the x – z plane; the same condition is satisfied in any plane containing the z axis. The emergent light forms a series of concentric fringe rings on the screen. From the pattern of the fringe rings, the spectrum of the source light is analyzed.

Next, the case of multiple spectra is considered. Light of each wavelength forms its own set of fringe rings, and without the prism, the rings will be superimposed, making it difficult to decode the fringe rings. A dispersive prism, which is placed between the Fabry–Pérot resonator and the output screen in Fig. 3.10, disperses the locations of

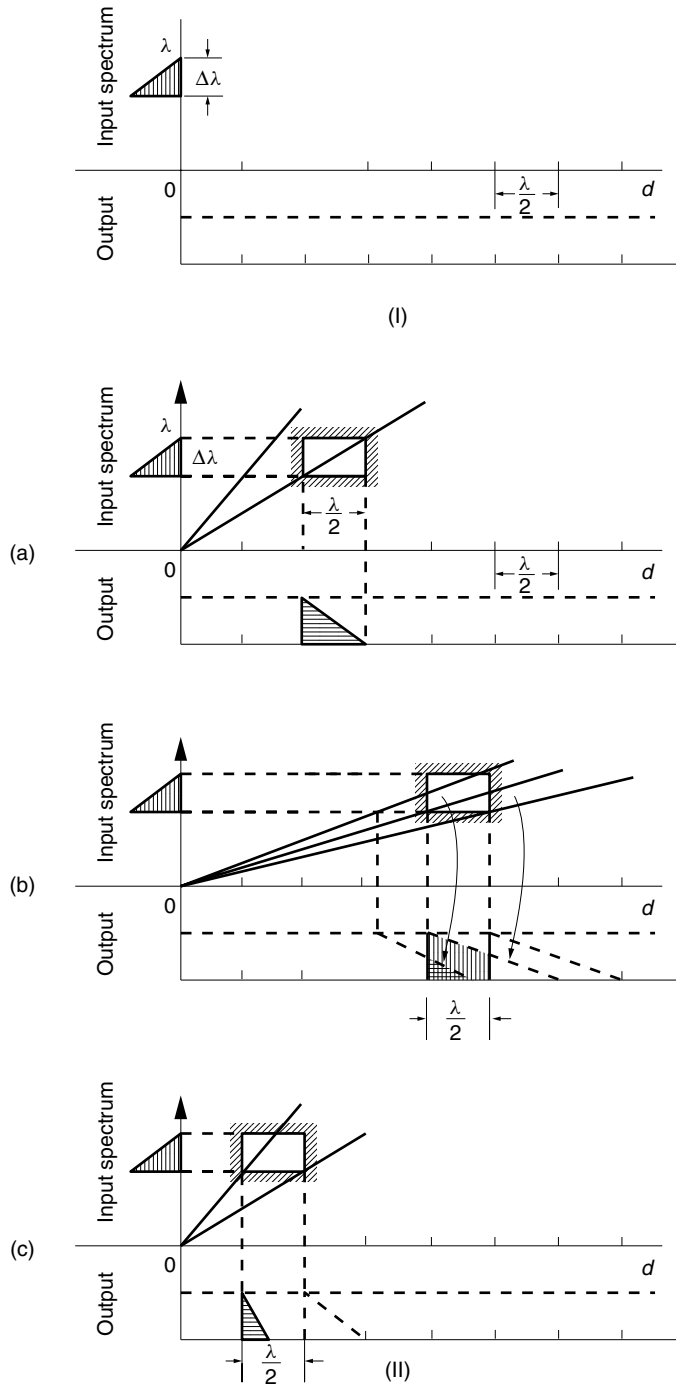


Figure 3.9 Input spectrum (I) and output displays (II) with different values of d . (a) $d = \lambda^2/2\Delta\lambda$. (b) $d = \lambda^2/\Delta\lambda$. (c) $d = \lambda^2/4\Delta\lambda$.

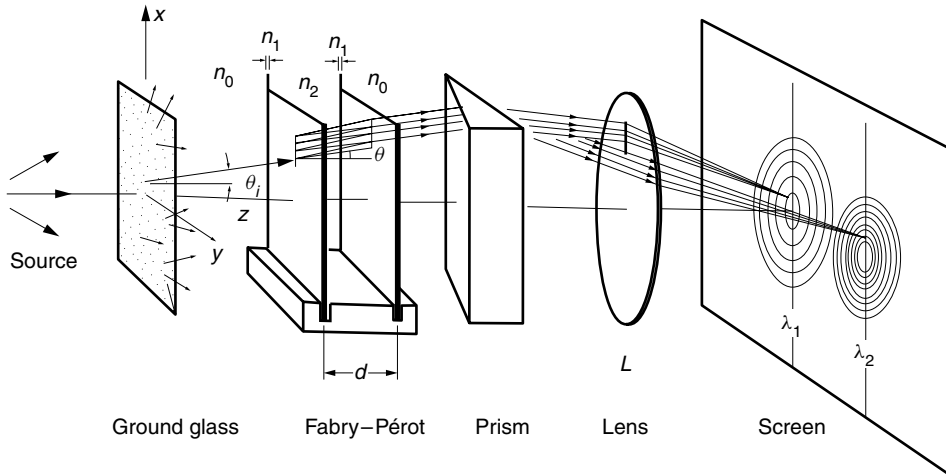


Figure 3.10 Fabry-Pérot spectroscopy using a prism to separate the fringe rings.

the fringe rings belonging to the different wavelengths, so as to avoid the overlapping of the fringe rings.

Example 3.4 Referring to Fig. 3.10, find the radii of the first three maxima of the fringe rings. The wavelength of the diffused source is $0.6328 \mu\text{m}$. The spacing d between the reflectors of the Fabry-Pérot resonator is $30.13 \mu\text{m}$. The refractive index n_2 of the medium in the Fabry-Pérot resonator is 1.05, and the refractive index n_0 of the medium outside the resonator is 1.00. The focal length f of the lens L is 50 mm.

Solution Combining Eqs. (3.14) and (3.32) gives

$$2n_2d \cos \theta = m\lambda$$

With the given parameters, the order of the fringe ring m at $\theta = 0$ is

$$2(1.05)(30.13) = m(0.6328)$$

$$m = 100$$

The angle θ for $m = 99$ is

$$2(1.05)(30.13) \cos \theta = (99)(0.6328)$$

$$\cos \theta = \frac{99}{100}$$

$$\theta = 8.11^\circ$$

In the parallel-plate case, the general Snell's law of

$$n_i \sin \theta_i = \text{constant}$$

holds true. The angle θ_i in air can immediately be related to the angle θ in the center medium without dealing with transmission through the plate with index of refraction n_1 , namely,

$$\begin{aligned} n_2 \sin \theta &= n_0 \sin \theta_i \\ \theta_i &= 8.52^\circ \quad \text{for } m = 99 \end{aligned}$$

Similarly, θ_i for the next two orders is

$$\begin{aligned} \theta_i &= 12.06^\circ \quad \text{for } m = 98 \\ \theta_i &= 14.79^\circ \quad \text{for } m = 97 \end{aligned}$$

The incident wave to the lens in the x direction is

$$E = E_0 e^{jkx \sin \theta_i}$$

The pattern on the back focal plane is the Fourier transform of the input to the lens (see Eq. (1.160)):

$$E_i(x_i) = F\{E\}_{f_x = \frac{x_i}{\lambda f}}$$

The pattern is cylindrically symmetric, and $x_i = r_i$.

$$\begin{aligned} E(r_i, f) &= E_0 \delta \left[\frac{1}{\lambda} \left(\frac{r_i}{f} - \sin \theta_i \right) \right] \\ r_i &= f \sin \theta_i \end{aligned}$$

Thus, the radii for $m = 99, 98, 97$ are

$$\begin{aligned} 7.41 \text{ mm} &\quad \text{for } m = 99 \\ 10.4 \text{ mm} &\quad \text{for } m = 98 \\ 12.8 \text{ mm} &\quad \text{for } m = 97 \end{aligned} \quad \square$$

An angle adjustment scheme using a parallel beam incident to an etalon is shown in the fiber ring laser of Fig. 3.11. The etalon transmits light only at certain wavelengths, which are determined by the angles of incidence specified by Eqs. (3.14) and (3.32). The light of a particular wavelength is fed back to the erbium-doped fiber amplifier (see Chapter 13) and the wavelength of the fiber ring laser can be tuned by rotating the etalon [5].

3.2.3 Scanning by the Index of Refraction

Extreme fine tuning of the Fabry-Pérot resonator without disturbing the reflector alignment can be achieved by controlling the refractive index of the medium inside the Fabry-Pérot resonator. For instance, by increasing the pressure of 100% nitrogen

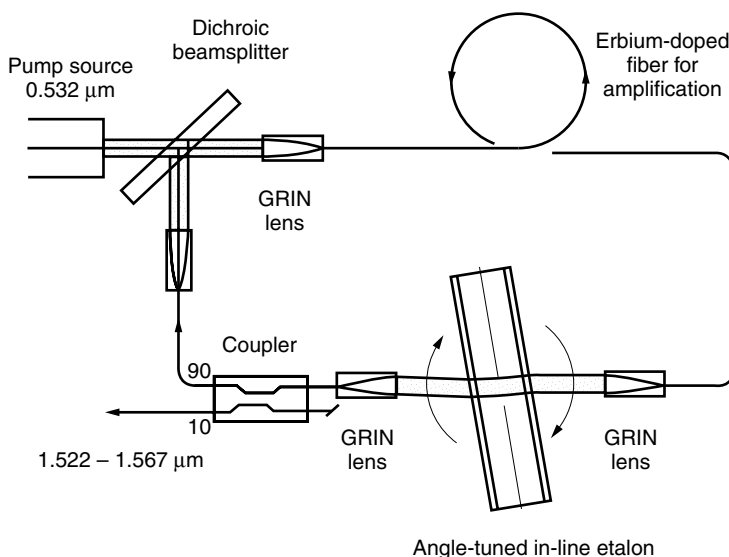


Figure 3.11 A tunable fiber ring laser.

gas by 1 atmosphere, the refractive index changes by 0.0003, and fine control of the refractive index is possible by controlling the pressure of the nitrogen gas.

Another material used is a liquid crystal whose index of refraction can be changed by an external electric field. The electrooptic effect of the nematic-type liquid crystal is large and the control voltage is small, typically 1–5 volts. The refractive index can be controlled by as much as 20% by means of these control voltages.

Figure 3.12 shows a schematic of a liquid crystal fiber filter. A nematic-type liquid crystal is sandwiched between indium tin oxide (ITO) transparent electrodes

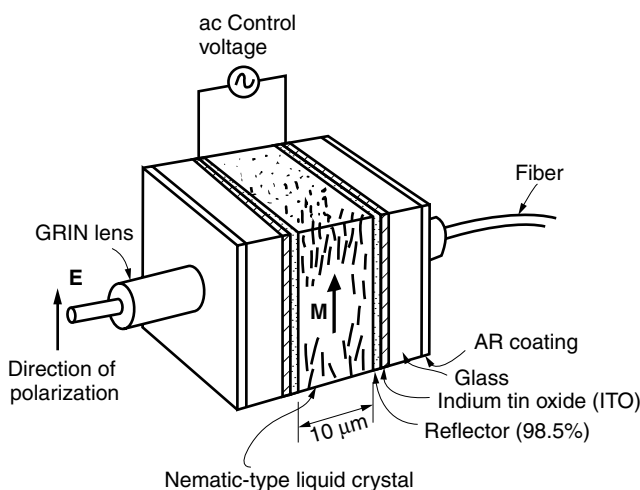


Figure 3.12 Electronically tunable liquid crystal Fabry-Pérot resonator.

deposited on a glass plate. The reflectors of the Fabry-Pérot cavity are deposited over each surface of the ITO electrode. The outer surfaces of the glass plates are antireflection (AR) coated. The whole liquid crystal compartment is connected to GRIN lenses (graded index lenses whose refractive indices are quadratic with the radius and function like thin lenses) [6–11].

Such a tunable filter is useful, for instance, for tuning in to a particular wavelength in a wavelength division multiplexing (WDM) fiber-optic communication system.

Another example of refractive index scanning is found in an electronically tunable semiconductor laser. When the electrons are injected into the guiding layer of a semiconductor laser, its index of refraction is reduced due to the plasma effect (see Section 14.4.3.2) of the electrons. It is decreased as much as 0.35%. A numerical calculation is given in Example 3.5 and the general description can be found in Section 14.10.

Example 3.5 Consider a Fabry-Pérot resonator cavity filled with a semiconductor material with refractive index n , as shown in Fig. 3.13. The spacing between the reflectors is $63.4\ \mu\text{m}$ and the free spectral range around the wavelength $\lambda_m = 1.5517\ \mu\text{m}$ is $\Delta\lambda_{\text{FSR}} = 5.42\ \text{nm}$.

- Find the value n of the refractive index of the semiconductor material.
- Find the value m of the order of the longitudinal mode for $\lambda_m = 1.5517\ \mu\text{m}$.
- With the above arrangement, if an electric current is injected into the semiconductor, the refractive index n of the semiconductor will be decreased due to the plasma effect of the injected current. In this manner, an electronically tunable Fabry-Pérot device can be fabricated. Estimate the refractive index change needed to sweep the wavelength of the resonance peaks by $5.42\ \text{nm}$ around $\lambda = 1.5517\ \mu\text{m}$ by this scheme. Assume the same mode is retained before and after the injection of the current.

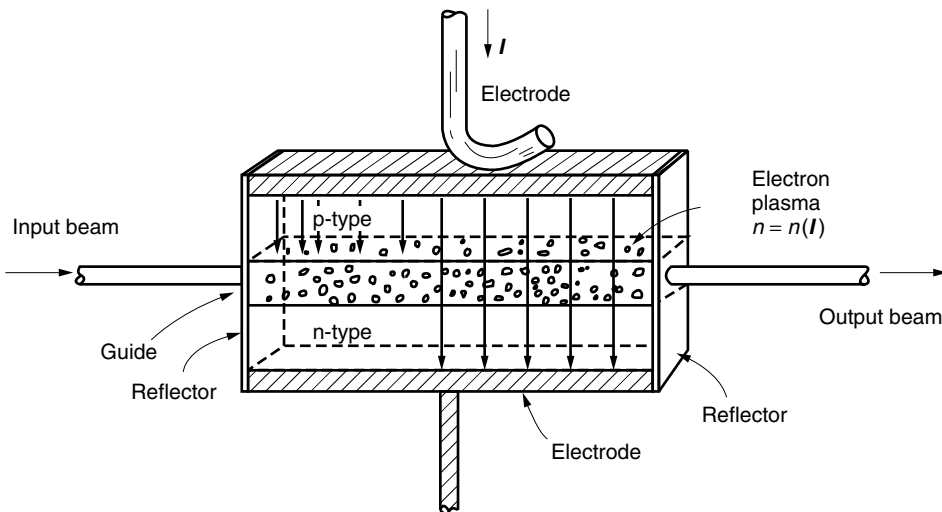


Figure 3.13 Electronically tunable semiconductor Fabry-Pérot cavity.

Solution

(a) From Eq. (3.36) including the index of refraction n , the free spectral range is

$$\begin{aligned}\Delta\lambda_{\text{FSR}} &= \frac{\lambda^2}{2nd} \\ n &= \frac{\lambda^2}{2\Delta\lambda_{\text{FSR}}d} \\ &= \frac{(1.5517)^2}{2(5.42 \times 10^{-3})(63.4)} \\ &= 3.5\end{aligned}$$

(b) To find the value of m , we use

$$\begin{aligned}\frac{\lambda_m}{2}m &= nd \\ m &= \frac{2nd}{\lambda_m} = \frac{2(3.5)(63.4)}{1.5517} = 286\end{aligned}$$

(c) Let λ_m be the wavelength of the resonance before injection of the current and let λ'_m be the wavelength after the injection of current.

$$\begin{aligned}\frac{\lambda_m}{2}m &= nd \\ \frac{\lambda'_m}{2}m &= n'd\end{aligned}$$

Thus,

$$\begin{aligned}n - n' &= \left(\frac{\lambda_m - \lambda'_m}{2d} \right) m \\ &= \frac{(5.42 \times 10^{-3})(286)}{2(63.4)} \\ &= 0.0122\end{aligned}$$

In percentage, the change is

$$\frac{n - n'}{n} = 0.35\%$$

□

3.2.4 Scanning by the Frequency of Incident Light

By scanning the frequency of the incident light, a series of equally spaced resonance peaks are observed in the output of the Fabry-Pérot resonator at frequencies of

$$f_m = m \frac{c}{2nd}$$

This technique has many applications, such as controlling the channel spacing of wavelength division multiplexing (WDM) in a fiber-optic communication system.

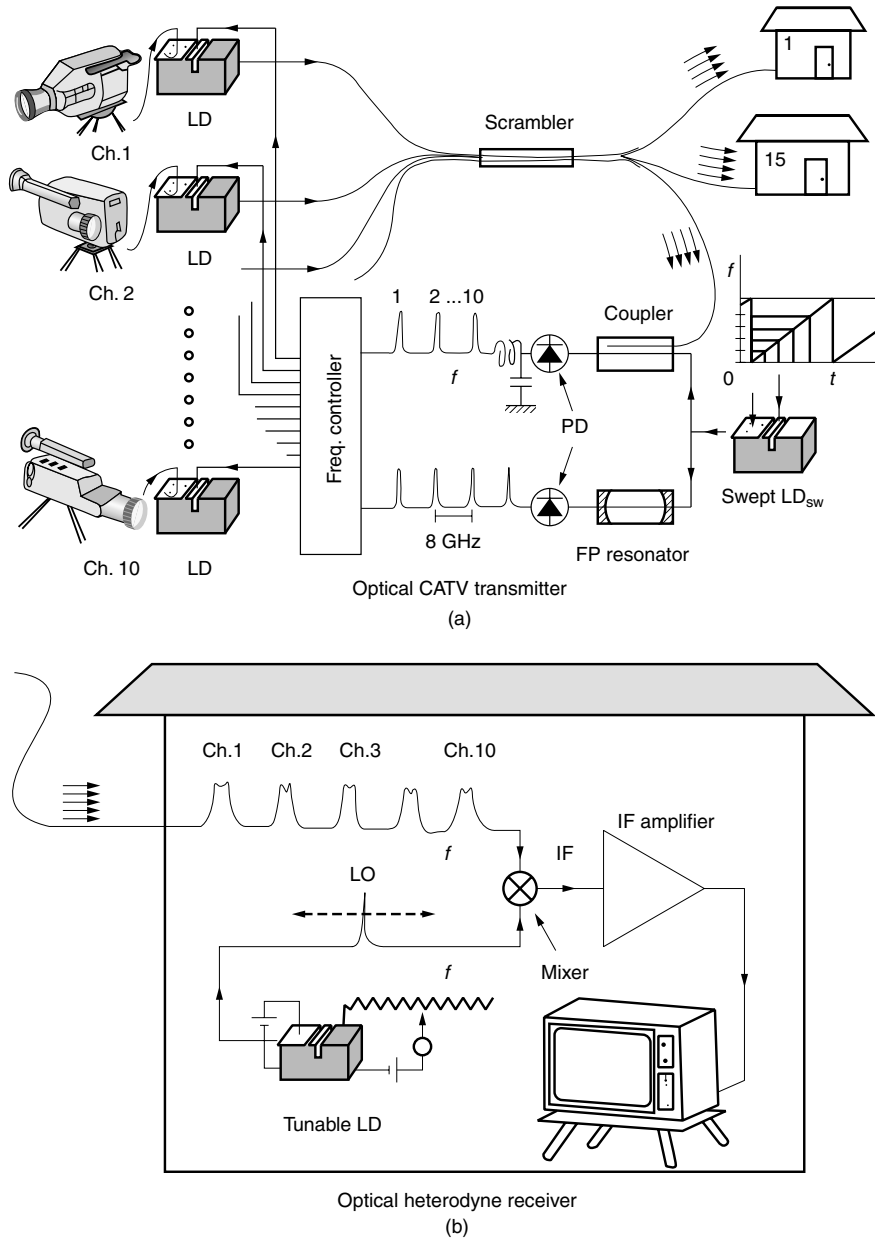


Figure 3.14 Use of a Fabry-Pérot resonator for controlling the channel spacing of the WDM of a fiber-optic communication system. (a) Channel space locking by the reference pulse method. (b) Optical heterodyne receiver at home.

Figure 3.14a shows a block diagram of such a WDM system used in the transmitting station of a fiber-optic cable television system. There are 10 separate tunable laser diodes, which have been tuned to achieve a frequency spacing of 8 GHz between carrier frequencies. Each laser diode is modulated by its own television camera so that 10 TV channels are sent concurrently in one fiber.

The principle of operation is as follows. The outputs from the laser diodes are all fed to the scrambler so that each output fiber from the scrambler transmits the same signal, which contains 10 different TV channels, to each subscriber at home. Subscribers can select any of the 10 channels for viewing at home.

In such a system, it is important to tightly maintain the carrier frequency spacing of 8 GHz so as to avoid crosstalk between the channels. This example demonstrates the reference pulse method [12], which employs a Fabry-Pérot resonator to lock the channel spacing. The reference pulse scheme consists of a swept laser diode LD_{sw} , a coupler, and a Fabry-Pérot (FP) resonator, as shown in the lower right portion of Fig. 3.14a. The carrier frequency of LD_{sw} is linearly swept with respect to time over 80 GHz. One-half of the output from the swept LD_{sw} is fed to the coupler that leads to a photodiode mixer followed by a low-pass filter. The output from one of the scrambler ports is fed to the same coupler. Whenever the frequency of the swept LD_{sw} matches with that of any one of the channels, an electrical pulse appears at the output of the low-pass filter. At each sweep of LD_{sw} , a chain of 10 pulses appears.

On the other hand, the other half of the output from LD_{sw} is fed directly to the Fabry-Pérot resonator. Whenever the swept frequency matches the resonance frequencies of the resonator, the light will reach the photodiode and an electronic pulse appears at the output of the photodiode. The spacing between the resonance frequencies of the Fabry-Pérot resonator is set at 8 GHz. If every pair of pulses from the scrambled signal and from the Fabry-Pérot resonator is synchronized, the frequencies of all channels are properly spaced. The frequency controller compares each pair of pulses. If the pulses of the pair are not synchronized, an error signal is issued to that particular channel to correct the carrier frequency of the laser diode. With this method, the relative frequency fluctuation can be maintained within ± 25 MHz, which is 1.25×10^{-7} of the carrier frequency of the transmitter LD. It may be added that this scheme is much more economical than combining 10 individually frequency-stabilized laser diodes.

Lastly, a word about the scheme of the receiver at the subscriber's home will be added. In Fig. 3.14b, an optical heterodyne receiver (see Chapter 13) is used. The optical heterodyne is essentially a converter from an optical frequency to an intermediate frequency (IF), which is normally in the radiofrequency range. This conversion is achieved by mixing with a local oscillator (LO) light. A particular channel can be selected by tuning the frequency of the tunable local oscillator laser diode.

3.3 RESOLVING POWER OF THE FABRY-PÉROT RESONATOR

This section looks more closely at the problem of resolving two closely spaced spectra. Although the mode lines in Fig. 3.6 are all drawn without spectral width, in reality, the output from the Fabry-Pérot resonator does have finite width. The achievable resolution of the scanning Fabry-Pérot resonator as a spectrum analyzer depends on the lineshape of the transmitted light.

The resolving power R will be calculated. The expression for the transmitted light I_t with respect to the frequency ν is given by Eqs. (3.14) and (3.30) as

$$I_t = I_0 \frac{1}{1 + M \sin^2[(2\pi\nu/c)d]} \quad (3.42)$$

where $A = 0$, $n_2 = 1$, and $\theta = 0$. The lineshape of I_t depends on two factors: one is M , and the other is $2d/c$. The value of M is associated with the reflectance R of the

reflector through Eq. (3.31), and $2d/c$ is exactly the inverse of the FSR in frequency, as given by Eq. (3.39).

Figure 3.15 shows a graph of I_t as a function of ϕ . Let $\phi_{\pm 1/2}$ be the value of ϕ near the m th order mode for which the intensity drops to half its peak value,

$$I_t(\phi_{\pm 1/2}) = \frac{1}{2}I_0 \quad (3.43)$$

Let $\Delta\phi$ be the full width at half maximum (FWHM) in ϕ . From Eq. (3.32), $\phi_{\pm 1/2}$ is expressed as

$$\phi_{\pm 1/2} = 2m\pi \pm \frac{\Delta\phi}{2} \quad (3.44)$$

When the denominator of Eq. (3.30) becomes 2, Eq. (3.43) is satisfied. This condition, in combination with Eq. (3.44), means $\Delta\phi$ must satisfy

$$\begin{aligned} 1 &= M \sin^2 \left(\frac{2m\pi \pm \Delta\phi/2}{2} \right) \\ &= M \left(\sin m\pi \cos \frac{\Delta\phi}{4} \pm \cos m\pi \sin \frac{\Delta\phi}{4} \right)^2 \\ &= M \sin^2 \frac{\Delta\phi}{4} \end{aligned}$$

For $\sin(\Delta\phi/4) \ll 1$, $\Delta\phi$ is approximated as

$$\Delta\phi \div \frac{4}{\sqrt{M}} \quad (3.45)$$

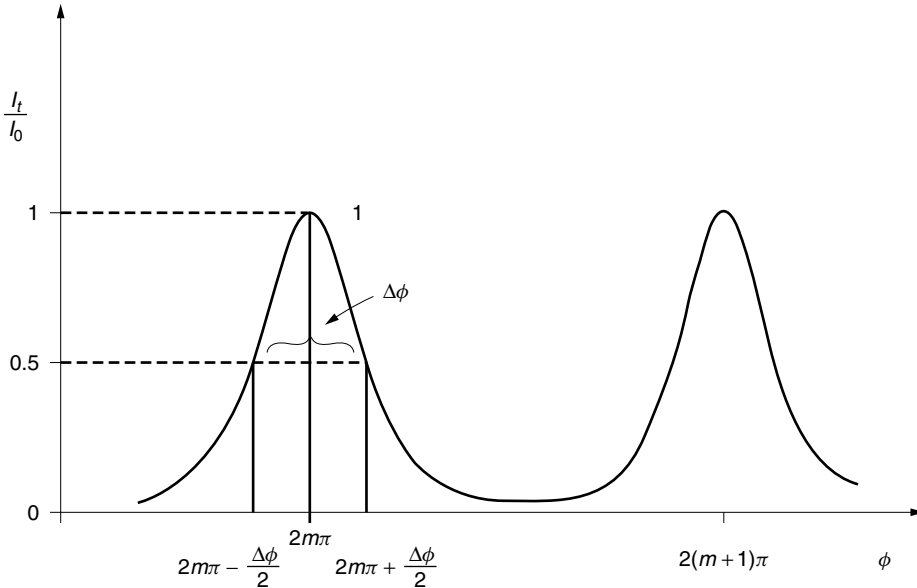


Figure 3.15 Resonance curve of the Fabry-Pérot resonator with ϕ . $\Delta\phi$ is the full width at half maximum (FWHM).

Using Eq. (3.14) with $n_2 = 1$ and $\theta = 0$, ϕ is rewritten in terms of the frequency ν as

$$\phi = \frac{4\pi d}{c} \nu \quad (3.46)$$

Inserting Eq. (3.45) into (3.46), the final result is obtained as

$$\Delta\nu_{1/2} = \frac{\Delta\nu_{\text{FSR}}}{F} \quad (3.47)$$

where

$$F = \frac{\pi}{2} \sqrt{R} \quad (3.48)$$

F is called the *finesse* and is further rewritten with the help of Eq. (3.31) as

$$F = \pi \frac{\sqrt{R}}{1 - R} \quad (3.49)$$

F is sometimes called the reflection finesse F_r to distinguish it from other types of finesse, as mentioned below. The finesse increases with an increase in R . With an increase in F , the FWHM $\Delta\nu_{1/2}$ becomes finer. By rewriting Eq. (3.47) using Eqs. (3.33) and (3.39), the value of the resolving power R_s , which is defined as

$$R_s = \frac{\nu}{\Delta\nu_{1/2}} \quad \text{or} \quad \frac{\lambda}{\Delta\lambda_{1/2}} \quad (3.50)$$

is obtained as

$$R_s = mF \quad (3.51)$$

Thus, the resolving power R_s of the Fabry-Pérot resonator is the product of $m(= 2d/\lambda)$ and F , where m can be thought of as the reflector spacing measured in terms of half-wavelengths, and the finesse F increases with the reflectance R of the reflector. The resolving power R_s is known as the quality value Q of the cavity in the field of electrical engineering. In practice, m is typically of the order of 10^3 , F is of the order of 10^2 , and R_s is of the order of 10^5 . Thus, the scanning Fabry-Pérot resonator has the capability of determining the wavelength of infrared light with a resolution better than 0.01 nm.

The analysis thus far has been idealized in that factors such as flatness, parallelism, or finiteness of the reflectors have been ignored. All of these factors decrease the value of the finesse. Among the neglected factors, the most important factor is the flatness. The effective finesse F_t is approximated as

$$\frac{1}{F_t^2} = \frac{1}{F_r^2} + \frac{1}{F_d^2} \quad (3.52)$$

where F_r is the reflection finesse and F_d is the flatness finesse. These finesses are defined as

$$F_r = \pi \frac{\sqrt{R}}{1 - R} \quad (3.53)$$

$$F_d = \frac{S}{2} \quad (3.54)$$

where the flatness S of the surface of the reflector is defined as the inverse of the depth of deviation d_s in terms of the wavelength λ .

$$S = \frac{1}{d_s/\lambda} \quad (3.55)$$

Example 3.6 The tunable fiber Fabry-Pérot filter [13] shown in Fig. 3.16 is used in an application that requires a FSR of 30 nm in wavelength and a finesse F_t of 100 at a wavelength of $\lambda = 1.55 \mu\text{m}$. Design the spacing between the fiber ends and the coated reflectivity. The flatness finesse is $F_d = 400$.

Solution

$$\Delta\lambda_{\text{FSR}} = \frac{\lambda^2}{2nd}$$

With $n_2 = 1$, the spacing is

$$d = \frac{\lambda^2}{2\Delta\lambda_{\text{FSR}}} = \frac{1.55^2}{2(0.03)} = 40 \mu\text{m}$$

From Eq. (3.52), the effective finesse is

$$\frac{1}{F_t^2} = \frac{1}{F_r^2} + \frac{1}{F_d^2}$$

Solving for F_r gives

$$F_r^2 = \frac{F_d^2 F_t^2}{F_d^2 - F_t^2}$$

With $F_d = 400$ and $F_t = 100$, the value of F_r is

$$F_r = \sqrt{\frac{(400)^2(100)^2}{400^2 - 100^2}} = 103.3$$

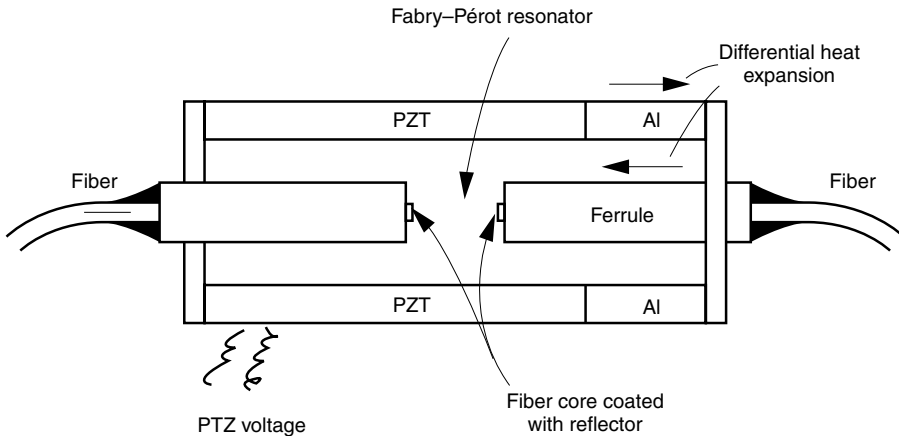


Figure 3.16 Tunable fiber Fabry-Pérot resonator. (After C. M. Miller and F. J. Janniello [13].)

From Eq. (3.53),

$$F_r = \frac{\pi\sqrt{R}}{1-R}$$

and solving for \sqrt{R} gives

$$\begin{aligned}\sqrt{R} &= \frac{-\pi \pm \sqrt{\pi^2 + 4F_r^2}}{2F_r} \\ &= \frac{-\pi \pm \sqrt{\pi^2 + 4(103.3^2)}}{2(103.3)} \\ \sqrt{R} &= 0.985 \\ R &= 0.970\end{aligned}$$

□

Example 3.7 A scanning Fabry-Pérot resonator is used to obtain the output shown in Fig. 3.17, which is the spectrum from an erbium-doped fiber amplifier (EDFA). Determine the Fabry-Pérot spectrometer's value of finesse and the spacing d of the reflectors required to produce such a display. Assume a free spectral range of 110 nm in order to cover the entire tail of the spontaneous emission spectrum.

Solution The spacing d of the reflectors is found from Eq. (3.36):

$$d = \frac{\lambda^2}{2\Delta\lambda_{\text{FSR}}} = \frac{1.55^2}{2(0.11)} = 10.92 \mu\text{m}$$

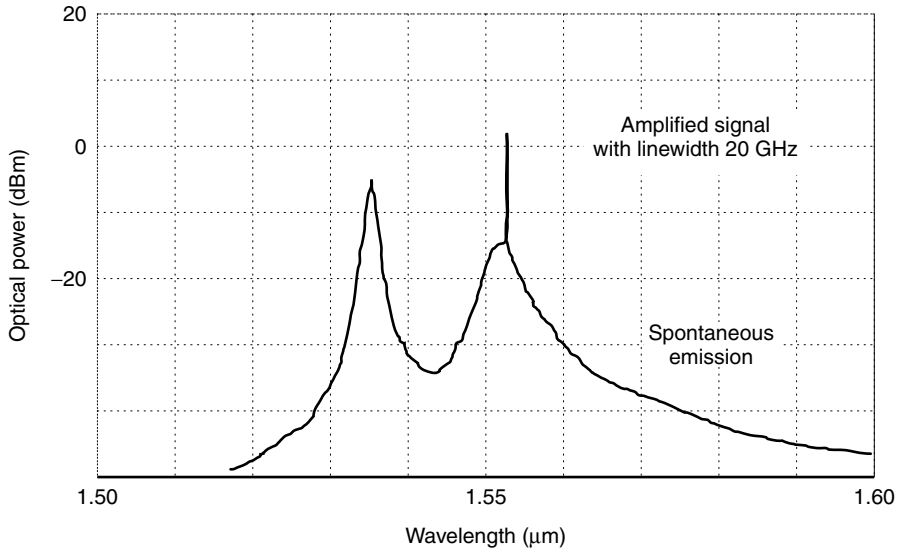


Figure 3.17 Spectrum of the output from an optical EDFA.

The mode number from Eq. (3.33) is

$$m = \frac{2d}{\lambda} = \frac{2(10.92)}{1.55} = 14$$

The linewidth of the amplified signal is 20 GHz, and this determines the required resolving power of the Fabry-Pérot resonator.

The needed resolving power is

$$R_s = \frac{\nu}{\Delta\nu}$$

$$\nu = \frac{c}{\lambda} = \frac{3 \times 10^{14}}{1.55} = 1.94 \times 10^{14} \text{ Hz}$$

$$R_s = \frac{1.94 \times 10^{14}}{20 \times 10^9} = 0.97 \times 10^4$$

From Eq. (3.51), the finesse is

$$F = \frac{R_s}{m} = \frac{0.97 \times 10^4}{14} = 693 \quad \square$$

Example 3.8 Figure 3.18 shows a schematic of an erbium-doped fiber amplifier (EDFA). Both the signal light with wavelength $\lambda_s = 1.554 \pm 0.020 \mu\text{m}$ and the pump light with wavelength $\lambda_p = 1.480 \mu\text{m}$ have to be present in the erbium-doped core of the fiber, and the pump light has to be removed from the amplified light after the amplification. The Fabry-Pérot dichroic filter whose transmission specification is shown in Fig. 3.19 is to be designed. It transmits most of λ_s and is highly reflective at λ_p . In all the calculations, assume $n_2 = 1$, $\theta = 45^\circ$, and $A = 0$.

- Find the reflector spacing of the Fabry-Pérot resonator.
- What is the reflectance R of the reflector if the full width at half maximum (FWHM) of the signal light is 40 nm?

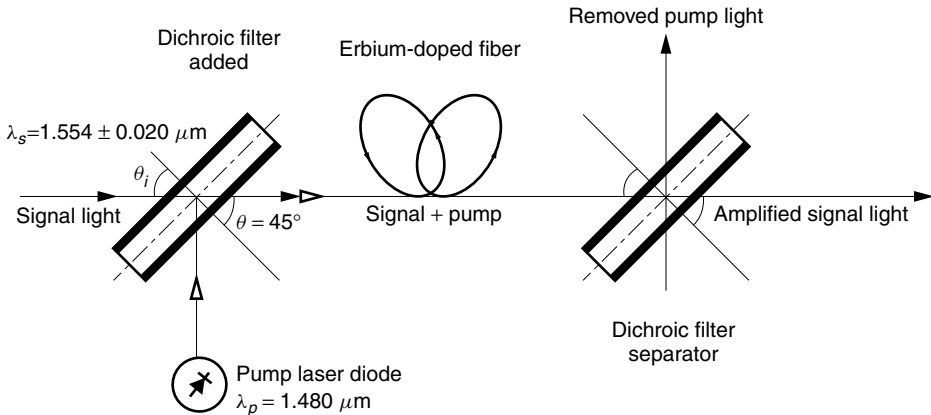


Figure 3.18 Dichroic filter used for an erbium-doped fiber amplifier (EDFA). Passband $\lambda_s = 1.554 \pm 0.020 \mu\text{m}$ and stopband $\lambda_p = 1.480 \mu\text{m}$.

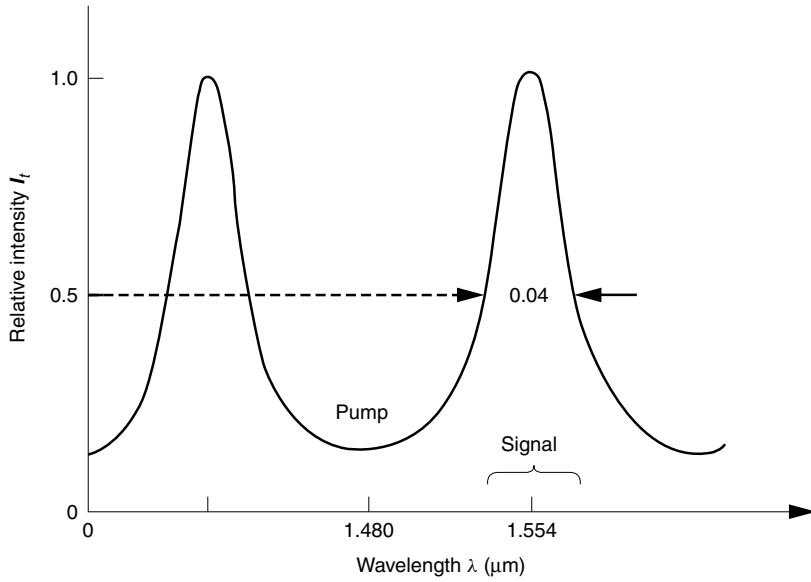


Figure 3.19 Specification of a dichroic filter used for an optical EDFA.

- (c) What is the reflectance I_r/I_0 of the dichroic filter at the wavelength of the pump light?
- (d) What happens if two of the same filters are staggered? Assume that there are no multiple reflections between the filters.

Solution

- (a) From Eqs. (3.14) and (3.32)

$$d \cos \theta = m \frac{\lambda}{2}$$

At the passband, this condition is

$$d \cos \theta = m \frac{\lambda_s}{2}$$

and at the stop band,

$$d \cos \theta = (m + 0.5) \frac{\lambda_p}{2}$$

With $\lambda_s = 1.554 \mu\text{m}$, $\lambda_p = 1.48 \mu\text{m}$, and $\theta = 45^\circ$, the value of m is

$$m = 10.0$$

and the spacing is

$$\begin{aligned} d &= \frac{m \lambda_s}{2 \cos \theta} \\ &= \frac{(10)(1.554)}{2/\sqrt{2}} = 10.918 \mu\text{m} \end{aligned}$$

(b) From Eqs. (3.50) and (3.51), the resolving power is

$$R_s = \frac{\lambda_s}{\Delta\lambda_{1/2}} = mF$$

$$\frac{1.554}{0.04} = 10F$$

$$F = 3.89.$$

From Eq. (3.49),

$$F = \frac{\pi\sqrt{R}}{1-R}$$

$$R = 0.455$$

(c) From Eq. (3.29) with $A = 0$ and $\sin(\phi/2) = 1$, the reflected intensity is

$$\frac{I_r}{I_0} = \frac{4R}{(1-R)^2 + 4R}$$

$$= \frac{4(0.455)}{(1-0.455)^2 + 4(0.455)} = 0.86$$

or

$$\frac{I_t}{I_0} = 0.14$$

(d) The maximum transmittance remains unity. The bandwidth of the passband becomes narrower. The minimum transmittance at the pump light becomes $(0.14)^2 = 0.02$. \square

3.4 PRACTICAL ASPECTS OF OPERATING THE FABRY-PÉROT INTERFEROMETER

This section describes practical techniques for operating the Fabry-Pérot interferometer in the laboratory.

3.4.1 Methods for Parallel Alignment of the Reflectors

The perpendicularity of mirror M_1 with respect to the incident beam can be examined by placing a card with a pinhole as shown in Fig. 3.20. If the laser beam reflected from M_1 goes through the pinhole, M_1 has been adjusted properly.

Mirror M_2 can be adjusted by projecting the transmitted beam onto a distant screen. When the mirror M_2 is aligned, the projected beam converges into a single spot.

Next, Figure 3.21 illustrates a method for fine adjusting a large-diameter reflector. Diffused visible incoherent light is used as the light source. A human eye probes the parallelism of the reflectors. The angles of the emergent light are discrete, as defined by

$$\cos \theta_m = m \frac{\lambda}{2d} \quad (3.56)$$

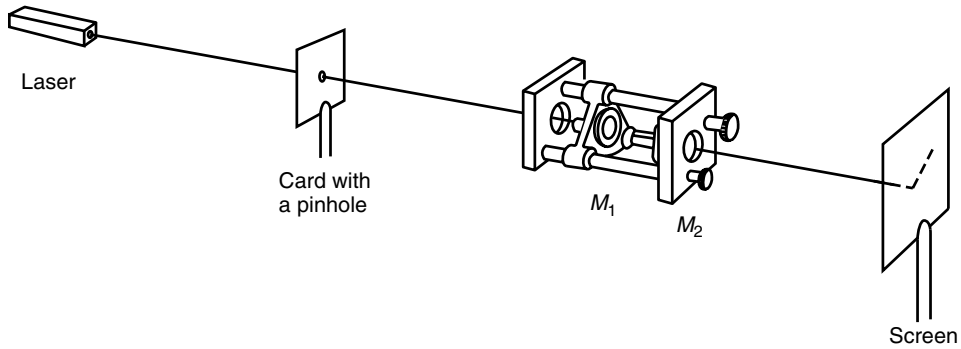


Figure 3.20 A method for aligning the Fabry-Pérot cavity.

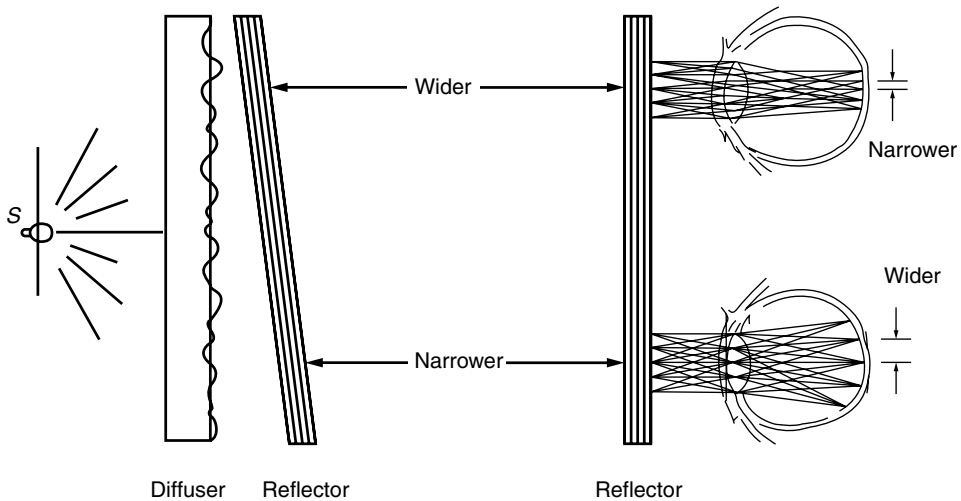


Figure 3.21 A method for the fine adjustment of a pair of large-diameter Fabry-Pérot reflectors.

which is derived from Eqs. (3.14) and (3.32). The discreteness of the angles is also explained in Fig. 3.10 and Example 3.4. The human eye's lens performs the same function as in Fig. 3.10. When the eye is focused on a far away object, an interference fringe is formed on the retina of the eye.

According to Eq. (3.56), regions with a wider reflector spacing project the interference fringe pattern with a narrower period. On the other hand, regions with a narrower reflector spacing, project patterns with a wider period. Parallelism of the reflectors is accomplished by scanning the eye across the reflector and adjusting the tilt of the reflector so that the period of the fringes remains constant.

Next, a method for aligning gas laser mirrors [14] is explained. With gas laser mirrors, not only the mirror parallelism, but also the perpendicularity of the gas tube axis to the mirror surface has to be achieved.

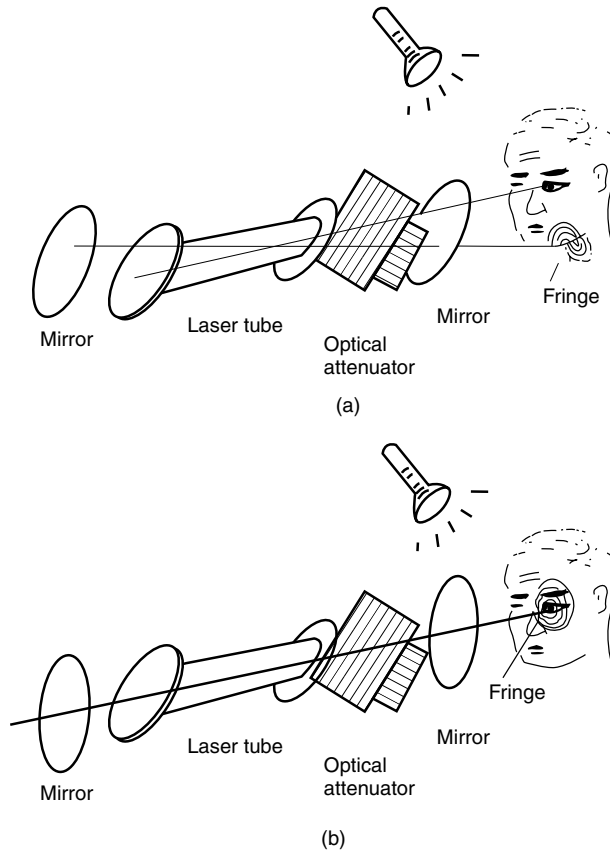


Figure 3.22 A method for aligning laser mirrors. (a) The beam is perpendicular to both mirrors but not collinear to the tube axis. (b) The beam is perpendicular to both mirrors and collinear to the tube axis.

As shown in Fig. 3.22, an optical attenuator is inserted to control the light intensity and protect the eyes from inadvertent laser action. Let us assume that the parallelism of the mirrors has already been established using the previous method, and that interference fringes are already observable in the emergent light.

Figure 3.22a shows a state where the mirrors are parallel but the mirror surface is not perpendicular to the axis of the laser tube. The eye is positioned so that the glow inside the laser tube can be seen cylindrically symmetric. According to Eq. (3.56), the radii of the interference rings are very small for a long laser tube. If the mirror surface is not perpendicular to the laser tube, the interference fringe pattern will be projected onto the observer's face away from the eye. The parallelism of the mirrors is readjusted without moving the position of the eye so that the center of the concentric interference fringe comes to the center of the eyeball. The observer judges whether or not the interference fringe comes to the eyeball by looking at the image of his/her own eye reflected from the end mirror of the laser tube. If the image of the eye is too dark, then the eye is illuminated

by a lamp. Figure 3.22b shows the geometry of the well-adjusted mirrors and laser tube.

3.4.2 Method for Determining the Spacing Between the Reflectors

The true spacing between the mirrors can be difficult to measure if the mirrors are extremely close together or if the mirror surfaces are not readily accessible because of the support structure. When the mirrors have to be set within a few microns of each other, utmost care has to be exercised because the mirrors are easily damaged when they touch each other.

Two methods are suggested here. In one method, the same incident beam is measured with two different angles, and in the other method, the same incident beam is measured with two different mirror spacings avoiding the necessity of the absolute measurement of d . The former is recommended when the mirrors have to be set close to each other and is explained in this section. The latter is explained in the next section. Measurement by the different angles technique has been divided into five steps as follows:

1. Using visible light, the parallelism of the mirrors is first adjusted by the method shown in Fig. 3.21. The visible laser beam and the Fabry–Pérot resonator, whose mirror spacing is to be measured, are placed about 1 meter apart, as shown in Fig. 3.23.
2. The entire body of the Fabry–Pérot resonator is tilted such that the back reflected beam comes about one beam diameter away from the laser aperture. This location of the spot is designated as point P_1 .

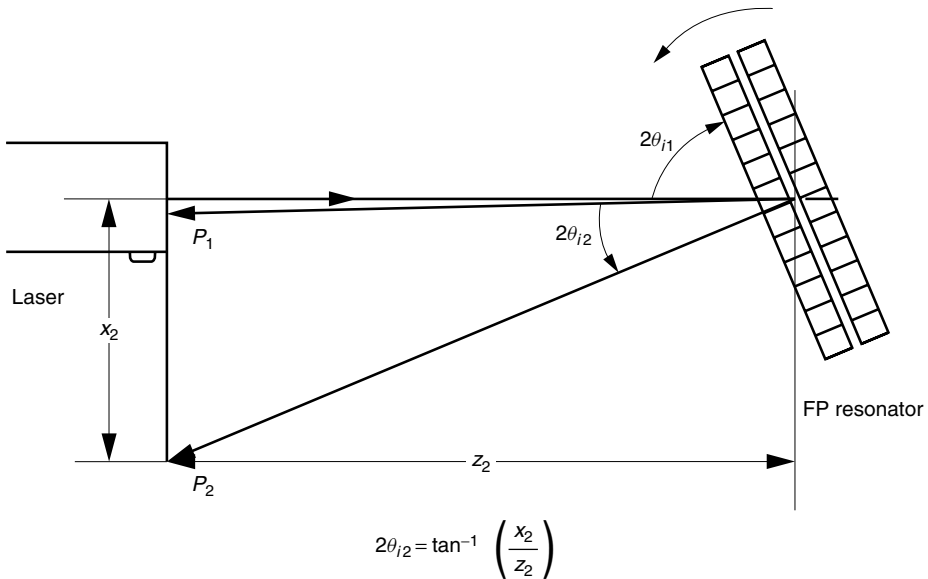


Figure 3.23 A method for determining the reflector spacing by shifting the incident angle.

3. The spacing of the mirrors is adjusted so that the intensity of the spot at P_1 changes to a dark spot. The mirror spacing at the m th resonance condition is, from Eq. (3.14),

$$\frac{4\pi d \cos \theta_1}{\lambda} = 2\pi m \quad (3.57)$$

where θ_1 is the internal angle of the laser beam in the Fabry-Pérot resonator mirror and $n_2 = 1$ is assumed. (θ_1 can be approximated by the external incident angle θ_{i1} .)

4. The entire body of the Fabry-Pérot resonator is then tilted further. Not only the location of the spot shifts, but also the intensity of the spot changes, first increasing to a maximum, then decreasing and becoming dark again. The location of such a dark spot is noted as P_2 , which tells us the internal angle θ_2 of the next higher order resonance mode and

$$\frac{4\pi d \cos \theta_2}{\lambda} = 2\pi(m + 1) \quad (3.58)$$

5. From these two angles, d can be calculated as

$$d = \frac{\lambda}{2(\cos \theta_2 - \cos \theta_1)} \quad (3.59)$$

3.4.3 Spectral Measurements Without Absolute Measurement of d

In this section, another method for measuring the spectrum by a scanning Fabry-Pérot resonator without a knowledge of the absolute value of the spacing of the reflectors will be explained [15].

Measurements are repeated with the same input light but with two different reflector spacings, which are translated by a known amount.

Let us assume that the incident light consists of two spectra, a main peak at λ and an auxiliary peak at $\lambda_a = \lambda + \Delta\lambda$. As explained earlier using Fig. 3.7c, the output display of the scanning Fabry-Pérot resonator would be something like the one shown in the encircled areas in Fig. 3.24. Figure 3.24a corresponds to a reflector spacing of $d = d_1$ while the spacing in Fig. 3.24b is $d = d_2$. These two displays have sufficient information to determine the spectrum, without absolute measurement of the reflector spacing.

The reflector spacing d is electronically scanned using a PZT drive assembly in the vicinity of $d = d_1$ in Fig. 3.24a and $d = d_2$ in Fig. 3.24b. Referring to Fig. 3.24a, as d is scanned, the output peaks appear in the following sequence. The (m_1) th order mode of λ appears first, at $d = d_1$, then that of λ_a appears at $d = d_1 + \Delta d_1$, and finally that of the $(m_1 + 1)$ th order mode of λ appears at $d = d_1 + \lambda/2$. These three peaks are encircled at the bottom of the figure. The expressions describing the locations of these three peaks are, from Eq. (3.33),

$$d_1 = \frac{\lambda}{2} m_1 \quad (3.60)$$

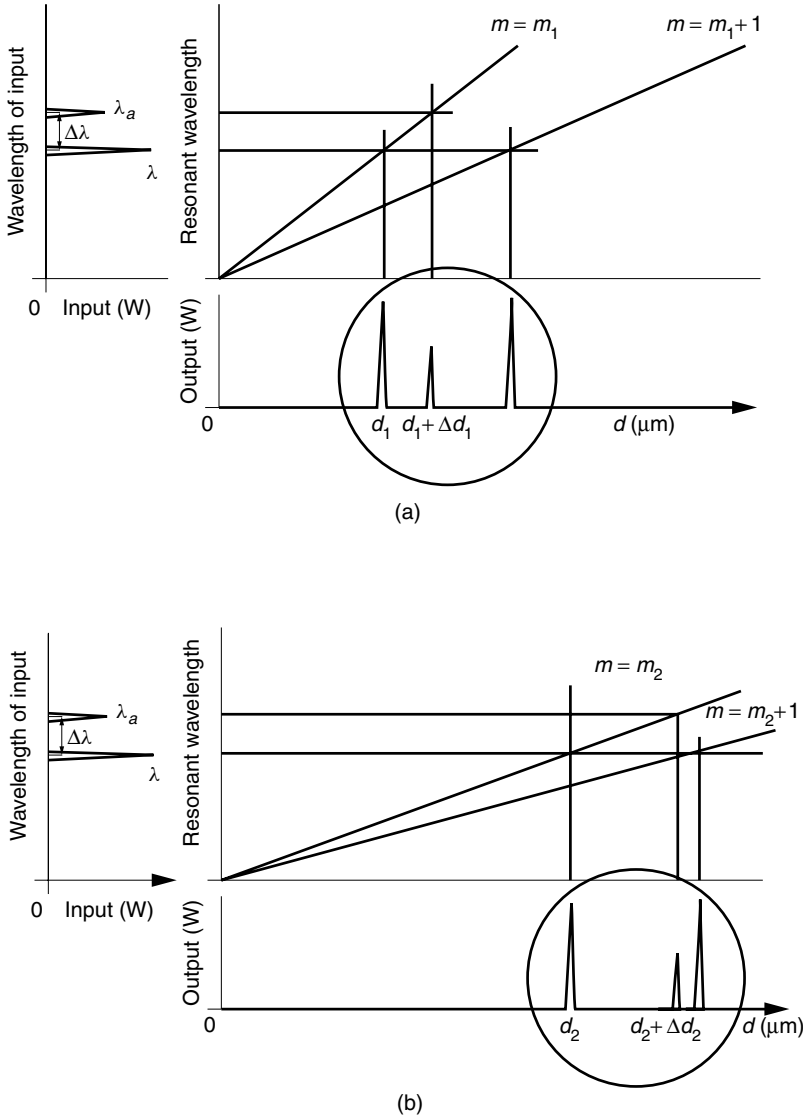


Figure 3.24 Measuring spectrum without absolute measurement of d . Operation of the Fabry-Pérot resonator (a) near $d = d_1$ and (b) near $d = d_2$.

$$d_1 + \Delta d_1 = \frac{(\lambda + \Delta\lambda)}{2} m_1 \quad (3.61)$$

$$d_1 + \frac{\lambda}{2} = \frac{\lambda}{2} (m_1 + 1) \quad (3.62)$$

Having recorded the spectra at $d = d_1$, the reflector is now shifted manually using the adjusting screw on the stationary plate (shown at the left in Fig. 3.5)

to a new location $d = d_2$. The new output display resulting from this operation is illustrated in Fig. 3.24b and the expressions describing the new locations of the three peaks are

$$d_2 = \frac{\lambda}{2}m_2 \quad (3.63)$$

$$d_2 + \Delta d_2 = \frac{(\lambda + \Delta\lambda)}{2}m_2 \quad (3.64)$$

$$d_2 + \frac{\lambda}{2} = \frac{\lambda}{2}(m_2 + 1) \quad (3.65)$$

The combination of the two sets of measurements provides the final result. From Eqs. (3.60) and (3.61), Δd_1 is

$$\Delta d_1 = m_1 \frac{\Delta\lambda}{2} \quad (3.66)$$

and from Eqs. (3.63) and (3.64),

$$\Delta d_2 = m_2 \frac{\Delta\lambda}{2} \quad (3.67)$$

Equations (3.66) and (3.67) give

$$\Delta d_2 - \Delta d_1 = \frac{\Delta\lambda}{2}(m_2 - m_1) \quad (3.68)$$

From Eqs. (3.60) and (3.63), the difference $d_2 - d_1$ is

$$d_2 - d_1 = \frac{\lambda}{2}(m_2 - m_1) \quad (3.69)$$

From Eqs. (3.68) and (3.69), the final result is

$$\frac{\Delta\lambda}{\lambda} = \frac{\Delta d_2 - \Delta d_1}{d_2 - d_1} \quad (3.70)$$

The denominator on the right-hand side of Eq. (3.70) is the manual shift by the adjustment screw on the stationary plate, which can be read from the vernier scale, and the numerator is the amount of electronic sweep by the PZT drive assembly.

The wavelength λ of the main spectrum is obtained from the period $d = \lambda/2$ of repetition of the pattern of the main peaks.

3.5 THE GAUSSIAN BEAM AS A SOLUTION OF THE WAVE EQUATION

In geometrical optics, a straight line represents the propagation direction of the wavefront. If the lightwave is indeed a plane wave, both the \mathbf{E} and \mathbf{H} fields have

to extend to infinity, or if the lightwave is a spherical wave, its wavefront diverges in all directions. These mathematical expressions, although adequate approximations in many situations, are not physically realizable. The Gaussian beam is a more realistic approximation from the viewpoints of both wave and ray optics. The energy of the Gaussian beam is confined within the vicinity of a straight line.

The wavefront of the Gaussian beam is unique in that it behaves like a plane wave in the vicinity of the beam waist, but it gradually converts into a spherical wave as the distance from the waist increases. The beam energy, however, is always confined within a finite divergence angle. The nature of Gaussian beam propagation in free space, as well as inside a cavity, will be explored. The treatment of the propagation in the cavity is especially important because the field inside a Fabry-Pérot resonator with spherically curved reflectors generally has a Gaussian distribution.

3.5.1 Fundamental Mode

Let the expression for a Gaussian beam be [16]

$$u = \psi(x, y, z)e^{jkz} \quad (3.71)$$

The amplitude function $\psi(x, y, z)$ can be a complex function. The Gaussian beam has to satisfy the wave equation:

$$\nabla^2 u + k^2 u = 0 \quad (3.72)$$

By inserting Eq. (3.71) into (3.72), $\psi(x, y, z)$ has to satisfy

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + j2k \frac{\partial \psi}{\partial z} = 0 \quad (3.73)$$

where the variation of the amplitude function ψ with respect to z was assumed slow, and

$$\frac{\partial^2 \psi}{\partial z^2} = 0 \quad (3.74)$$

ψ can be obtained by solving Eq. (3.73). The function

$$\psi = Ae^{j[P(z) + k(x^2 + y^2)/2q(z)]} \quad (3.75)$$

is tried as a solution of Eq. (3.73). The factor A is assumed constant. This solution is called the fundamental mode. The solutions where A is a function of x and y are called higher order modes and are presented in Section 3.7. The quantity $q(z)$ is a complex number referred to as the *q parameter*. The q parameter controls both the phase and the amplitude distribution in the transversal plane, and it plays an important role in describing the type of field. Insertion of Eq. (3.75) into (3.73) gives

$$2k \left(\frac{j}{q} - \frac{\partial P}{\partial z} \right) + \left(\frac{k}{q} \right)^2 \left(\frac{\partial q}{\partial z} - 1 \right) (x^2 + y^2) = 0 \quad (3.76)$$

The condition that Eq. (3.76) has to be satisfied at any point (x, y, z) gives

$$\frac{\partial q}{\partial z} = 1 \quad (3.77)$$

$$\frac{j}{q} = \frac{\partial P}{\partial z} \quad (3.78)$$

From Eq. (3.77), q satisfies

$$q(z) = z + c \quad (3.79)$$

where c is a constant.

Let

$$q = q_0 \quad \text{at } z = 0$$

Equation (3.79) is written as

$$q(z) = z + q_0 \quad (3.80)$$

Before going further, the nature of the q parameter will be examined. The constant q_0 is expressed explicitly in terms of real and imaginary parts by the real numbers s and t as

$$q_0 = s + jt \quad (3.81)$$

$$q(z) = s + z + jt \quad (3.82)$$

which leads to

$$\frac{1}{q(z)} = \frac{1}{R(z)} - j \frac{1}{Q(z)} \quad (3.83)$$

where

$$\frac{1}{R(z)} = \frac{s + z}{(s + z)^2 + t^2} \quad (3.84)$$

$$\frac{1}{Q(z)} = \frac{t}{(s + z)^2 + t^2} \quad (3.85)$$

ψ is obtained by inserting Eq. (3.83) into (3.75), and then u is obtained from Eq. (3.71) as

$$u = A e^{\underbrace{jP(z)}_{\text{Correction factor}} + \underbrace{jkz + jk(x^2 + y^2)/2R(z)}_{\text{Parabolic phase front}} + \underbrace{k(x^2 + y^2)/2Q(z)}_{\text{Amplitude}}} \quad (3.86)$$

3.5.2 Properties of the q Parameter

A single complex number $1/q(z)$ specifies both the phase and amplitude distribution of the Gaussian beam. The real part specifies the phase and the imaginary part specifies the amplitude. It is just like $\gamma = \beta + j\alpha$ in the case of a plane wave (see Example 2.5). Quantities obtainable from $q(z)$ will be summarized.

3.5.2.1 Beam Waist

If the radius of the beam $W(z)$ is defined as the distance $r = \sqrt{x^2 + y^2}$ from the beam axis where the light amplitude u decays to e^{-1} of that on the axis, then from Eq. (3.86)

$$k \frac{W^2(z)}{2Q(z)} = -1 \quad (3.87)$$

and with Eq. (3.85),

$$W^2(z) = -2 \frac{(s+z)^2 + t^2}{kt} \quad (3.88)$$

From Eq. (3.88), the minimum radius exists at the location where $z = -s$. Thus, the waist W_0 is

$$W_0^2 = -2 \frac{t}{k} \quad (3.89)$$

In conclusion, from Eqs. (3.82) and (3.89), the q parameter is related to the waist as

$$\text{Im } q(-s) = t = -\frac{k}{2} W_0^2 \quad (3.90)$$

Thus, the imaginary part of $q(z)$ at $z = -s$ provides the size of the waist.

3.5.2.2 Location of the Waist

As obtained above, the location of the waist is where $z = -s$. This is the location where the real part of $q(z)$ given by Eq. (3.82) is zero; that is, where $q(z)$ is a pure imaginary number.

$$\begin{aligned} \text{Re } q(z) &= 0 \\ q(z) &= jt \end{aligned} \quad (3.91)$$

3.5.2.3 Radius of Curvature of the Wavefront

The radius of curvature $R(z)$ is, from Eq. (3.84),

$$R(z) = (s+z) \left[1 + \left(\frac{t}{s+z} \right)^2 \right] \quad (3.92)$$

First, at the location of the waist where $z = -s$, Eq. (3.92) becomes $R(z) = \infty$. Hence, at the location of the waist, the wavefront is more like that of a plane wave and there is no variation in phase in a plane perpendicular to the direction of propagation. For a large $|s+z|$, $R(z)$ is approximated as

$$R(z) = s+z$$

and the phase factor in Eq. (3.86) represents a parabolic or, approximately speaking, a spherical wavefront with a correction factor $P(z)$. Note that $R(z)$ changes its sign at the waist where $s + z = 0$ and

$$\begin{aligned} R(z) &> 0 \quad \text{for } s + z > 0 \\ R(z) &< 0 \quad \text{for } s + z < 0 \end{aligned} \quad (3.93)$$

In the region of $s + z > 0$, the wavefront is diverging, while in the region of $s + z < 0$, the wavefront is converging. At far distances, $R(z) \rightarrow \infty$ and the wavefront starts to resemble that of a plane wave again.

3.5.3 With the Origin at the Waist

If the origin of the z axis is taken at the location of the waist, the expressions become simpler. With all of the following analyses, the origin will be taken at the waist unless otherwise specified.

3.5.3.1 Focal Parameters

If the origin is shifted to the location of the waist, then the new coordinate z' is

$$z' = z + s$$

The prime in z will be suppressed. The q parameter in the new coordinates is, from Eqs. (3.82) and (3.90),

$$q(z) = z - j\frac{k}{2}W_0^2 \quad (3.94)$$

From Eqs. (3.88) and (3.89), the beam radius or beamwidth at z is

$$W^2(z) = W_0^2 \left[1 + \left(\frac{2z}{kW_0^2} \right)^2 \right] \quad (3.95)$$

From Eqs. (3.90) and (3.92), the radius of curvature at z is

$$R(z) = z \left[1 + \left(\frac{kW_0^2}{2z} \right)^2 \right] \quad (3.96)$$

From Eq. (3.95), the distance $z = z_0$ where the cross-sectional area of the beam becomes twice as much as that of the waist is

$$z_0 = \frac{k}{2}W_0^2 \quad (3.97)$$

Such a distance z_0 is called the *focal beam parameter* or the *Rayleigh range*. It is the distance where the cross-sectional area inflates to twice that of the waist and indicates how fast the beam diverges from the waist. The shorter z_0 is, the faster the divergence is; or in other words, the narrower waist diverges faster.

Inserting Eq. (3.97) into Eqs. (3.95) and (3.96) gives

$$W^2(z) = W_0^2 \left[1 + \left(\frac{z}{z_0} \right)^2 \right] \quad (3.98)$$

$$R(z) = z \left[1 + \left(\frac{z_0}{z} \right)^2 \right] \quad (3.99)$$

From Eqs. (3.98) and (3.99), the relationship between $W^2(z)$ and $R(z)$ will be obtained. Equation (3.98) can be rewritten by multiplying by $(z_0/z)^2/(z_0/z)^2$ as

$$W^2(z) = W_0^2 \left[\frac{1 + (z_0/z)^2}{(z_0/z)^2} \right]$$

Inserting Eqs. (3.97) and (3.99) into this equation gives

$$W^2(z) = \frac{4R(z)}{k^2 W_0^2} z \quad (3.100)$$

Equation (3.100) provides the value of $W^2(z)$ from the given values of $R(z)$ and z , or vice versa.

3.5.3.2 Correction Factor

Next, the value of P will be calculated. Inserting Eq. (3.94) into (3.78) gives

$$\frac{\partial P}{\partial z} = \frac{-1}{(k/2)W_0^2(1 + j2z/kW_0^2)}$$

The solution of this differential equation is

$$P = j \ln \left(1 + j \frac{2z}{kW_0^2} \right)$$

The argument in the logarithm is expressed in polar form as

$$P = j \ln \sqrt{1 + \left(\frac{2z}{kW_0^2} \right)^2} - \phi \quad (3.101)$$

where

$$\phi = \tan^{-1} \frac{2z}{kW_0^2} \quad (3.102)$$

With the help of Eq. (3.95), Eq. (3.101) is rewritten as

$$e^{jP} = \frac{W_0}{W(z)} e^{-j\phi} \quad (3.103)$$

3.5.4 Gaussian Beam Expressions

Various expressions and parameters associated with the Gaussian beam will be summarized to facilitate the use of the formulas.

3.5.4.1 Amplitude Distribution

Using Eqs. (3.87) and (3.103), the expression for the amplitude of the Gaussian beam, Eq. (3.86), is rewritten as

$$u(z) = \underbrace{A}_{\text{Constant}} \underbrace{\frac{W_0}{W(z)}}_{\text{Contraction ratio}} e^{\underbrace{-j\phi + jk(z + r^2/2R(z))}_{\text{Spherical wavefront}} - \underbrace{r^2/W^2(z)}_{\text{Transverse amplitude distribution}}} \quad (3.104)$$

where

$$r^2 = x^2 + y^2$$

$$z_0 = \frac{k}{2} W_0^2 \quad (3.97)$$

$$W^2(z) = W_0^2 \left[1 + \left(\frac{z}{z_0} \right)^2 \right] \quad (3.98)$$

$$R(z) = z \left[1 + \left(\frac{z_0}{z} \right)^2 \right] \quad (3.99)$$

$$\phi(z) = \tan^{-1} \left(\frac{2z}{kW_0^2} \right) \quad (3.102)$$

The field distribution is illustrated in Fig. 3.25a.

The first factor in Eq. (3.104) is the amplitude and the second factor is the change in amplitude due to the change in the radius of the beam. The third factor is the phasefront, which approaches that of a plane wave at the waist as well as in the far field. For intermediate distances, the phasefront approaches a spherical wavefront with radius of curvature $R(z)$, but with a correction factor of ϕ . The last factor represents a bell-shaped transverse field distribution.

From Eq. (3.104), we see that only three parameters are needed to specify the Gaussian beam: the size of the beam waist, the distance from the beam waist, and the wavelength of the light.

3.5.4.2 Intensity Distribution

The intensity distribution $I(r, z)$ is given by uu^* and from Eq. (3.104) it is

$$I(r, z) = I_0 \left(\frac{W_0}{W(z)} \right)^2 e^{-2r^2/W^2(z)} \quad (3.105)$$

As shown in Fig. 3.25b, the transverse distribution at a particular value of z is Gaussian. The longitudinal distribution $I(0, z)$ along the beam axis has its maximum at the

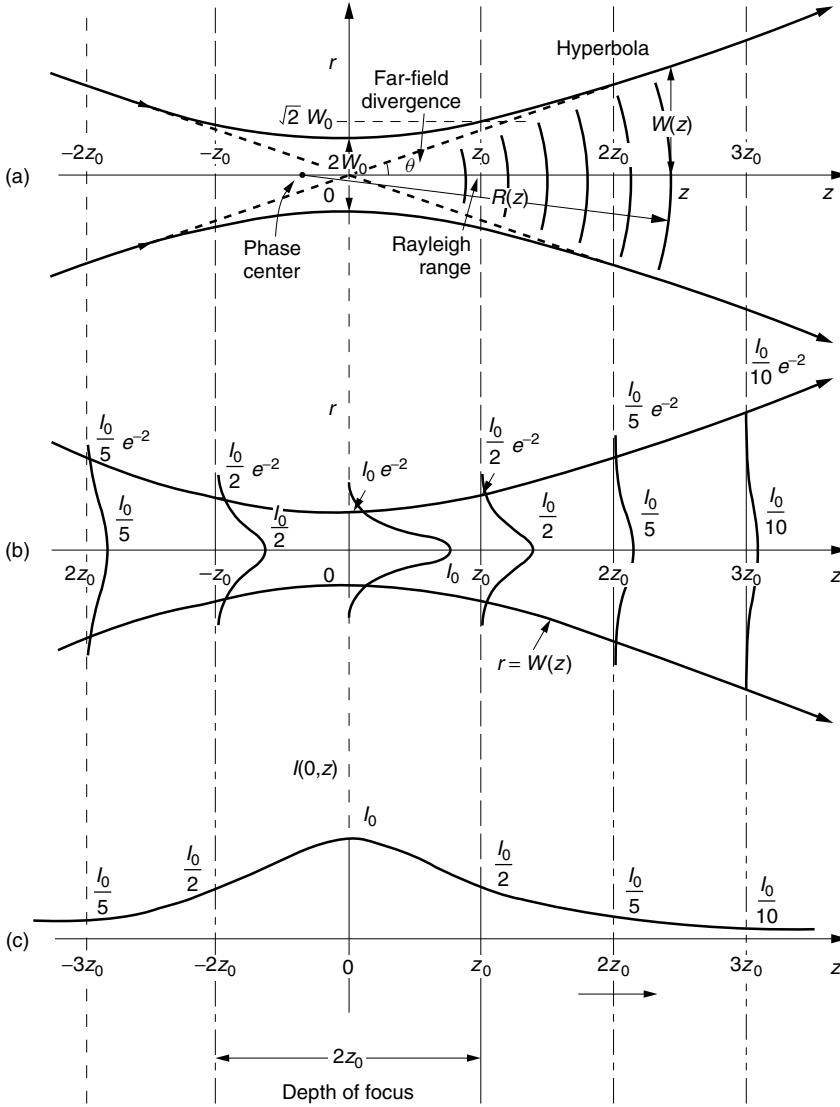


Figure 3.25 Gaussian beam. (a) Amplitude distribution. (b) Cross-sectional intensity distribution and beam radius. (c) Intensity distribution along the beam axis.

waist and decays monotonically as $[W_0/W(z)]^2$, as shown in Fig. 3.25c. The intensity distribution of the Gaussian beam is symmetric with respect to the waist. As soon as the distribution of $I(r, z)$ on one side of the waist is known, that on the other side is obtained by the mirror image.

3.5.4.3 Angle of the Far-Field Divergence

Equation (3.98) can be rewritten as

$$\left(\frac{W(z)}{W_0}\right)^2 - \frac{z^2}{z_0^2} = 1$$

Thus, the beam radius of the Gaussian beam is a hyperbolic function of z . Referring to Fig. 3.25a, the beam expansion at a far distance is

$$\lim_{z \rightarrow \infty} \tan \theta = \lim_{z \rightarrow \infty} \left(\frac{W(z)}{z} \right)$$

and from Eqs. (3.97) and (3.98),

$$\lim_{z \rightarrow \infty} \tan \theta = \frac{2}{kW_0} \quad (3.106)$$

The spread is inversely proportional to the size of the waist W_0 . This is similar to the spread of the far-field diffraction pattern of a plane wave through an aperture with radius W_0 (see Section 1.2).

3.5.4.4 Depth of Focus

The region of the waist bordered by $-z_0 < z < z_0$ is called the depth of focus D_0 and, from Eq. (3.97),

$$D_0 = 2z_0 = kW_0^2 \quad (3.107)$$

A smaller spot size means a shorter depth of focus unless the wavelength of the light is shortened.

Example 3.9 The beamwidth and the radius of curvature of a Gaussian beam with wavelength λ were measured at a certain location z . They were R_1 and W_1 , respectively.

- Find the q parameter.
- Find the beamwidth W_2 and the radius of curvature R_2 at a new location a distance d away from the measured location.

Solution The unknowns in this problem are the distance z from the point of observation to the location of the waist and the size W_0 of the waist. These unknowns will be found first.

- From Eqs. (3.98) and (3.99)

$$z = \frac{R_1}{1 + (z_0/z)^2}$$

$$W_0^2 = \frac{W_1^2}{1 + (z/z_0)^2}$$

Inserting z from Eq. (3.100) and z_0 from Eq. (3.97) into the denominators of these two equations gives

$$z = \frac{R_1}{1 + (2R_1/kW_1^2)^2}$$

$$W_0^2 = \frac{W_1^2}{1 + (kW_1^2/2R_1)^2}$$

The q parameter is obtained by inserting the above two equations into Eq. (3.94) as

$$q(z) = z - j\frac{k}{2}W_0^2$$

(b) From Eq. (3.79), the q parameter at distance d away is obtained just by adding d in the real part, and therefore $q(z + d)$ is

$$q(z + d) = z + d - j\frac{k}{2}W_0^2$$

The desired $R(z + d)$ and $W(z + d)$ will be found from $q(z + d)$:

$$\frac{1}{q(z + d)} = \frac{1}{R(z + d)} - j\frac{1}{Q(z + d)} \quad (3.108a)$$

$$\frac{1}{R(z + d)} = \frac{z + d}{(z + d)^2 + z_0^2} \quad (3.108b)$$

$$\frac{1}{Q(z + d)} = \frac{-z_0}{(z + d)^2 + z_0^2} \quad (3.108c)$$

where Eqs. (3.89) and (3.97) were used.

From Eq. (3.108b), $R(z + d)$ is obtained as

$$R(z + d) = (z + d) \left[1 + \left(\frac{z_0}{z + d} \right)^2 \right]$$

From Eqs. (3.87) and (3.97) and Eq. (3.108c), $W(z + d)$ is obtained as

$$W^2(z + d) = W_0^2 \left[1 + \left(\frac{z + d}{z_0} \right)^2 \right]$$

It should be noted that the solutions for $R(z + d)$ and $W(z + d)$ are obtained by simply replacing z with $z + d$, once the values of z and W_0 are found in part (a). \square

3.6 TRANSFORMATION OF A GAUSSIAN BEAM BY A LENS

A Gaussian beam propagating in free space cannot have more than one waist (just like a human being). However, if a lens is introduced into the beam, then it is possible to form a new beam waist whose location and size are different from the original. Thus, convex and concave lenses and mirrors are valuable tools for transforming Gaussian beams into desired dimensions. The transformation of the Gaussian beam by a lens is the subject of this section.

3.6.1 Transformation of the q Parameter by a Lens

A thin convex lens with focal length f is placed a distance d_0 to the right of the waist of the incident Gaussian beam as shown in Fig. 3.26. The quantities associated with the incident beam will be denoted by the subscript 0, and those of the emergent beam from the lens, by the subscript 1. The radius and phasefront of the emergent light at an arbitrary distance from the lens, as well as the location and size of the new waist, will be calculated.

The lens not only converges the phasefront but also changes the radius of the beam. In other words, the lens changes both $R(z)$ and $W(z)$, and it is best to deal with the basic parameter q . Not only the lens but the propagation itself changes the q parameter. The process of the change in q will be followed. Let the q parameter at the incident waist be

$$q = q_0 \quad (3.109)$$

At the waist, q_0 is a purely imaginary number and its value is $-j(k/2)W_0^2$ from Eqs. (3.90) and (3.91); but at other points, q is a complex number. After propagating a distance d_0 to the front surface of the lens, from Eq. (3.80), q becomes

$$q'_0 = q_0 + d_0 \quad (3.110)$$

Inside the lens, the beam acquires a quadratic phase distribution specified by Eq. (1.139) in Chapter 1, and from Eqs. (1.139) and (3.75) the field immediately after the lens is

$$u = Ae^{j[P(d_0) + k(x^2 + y^2)/2q'_0 - jk(x^2 + y^2)/2f]} \quad (3.111)$$

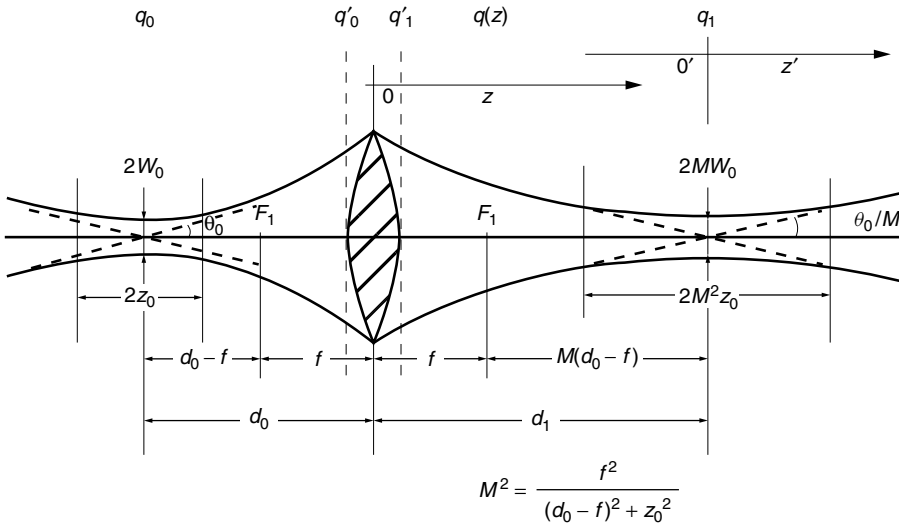


Figure 3.26 Transformation of Gaussian beam by a lens.

Thus, the q parameter just behind the lens is effectively

$$\frac{1}{q'_1} = \frac{1}{q'_0} - \frac{1}{f} \quad (3.112)$$

After further propagation by a distance z , the q parameter is

$$q(z) = q'_1 + z \quad (3.113)$$

where the origin of the z coordinate of the emergent beam is taken at the location of the lens (rather than at the waist).

The rule for calculating the q parameter is that the effect of the propagation is just the addition of the distance, as illustrated in Eqs. (3.110) and (3.113), while the effect of the lens is the addition of the inverse of q and the inverse of the focal length of the lens, as illustrated in Eq. (3.112).

Working backward from Eq. (3.113) to (3.110), the q parameter is expressed in terms of d_0 , z , and f as

$$q(z) = \frac{d_0(f - z) + fz + q_0(f - z)}{(f - d_0) - q_0} \quad (3.114)$$

Insert

$$q_0 = -j\frac{k}{2}W_0^2 = -jz_0$$

into Eq. (3.114) to obtain

$$q(z) = \frac{(z - f)[(d_0 - f)^2 + z_0^2] - f^2(d_0 - f) - jf^2z_0}{(d_0 - f)^2 + z_0^2} \quad (3.115)$$

From the real and imaginary parts of $q(z)$, the size and location of the waist will be found.

3.6.2 Size of the Waist of the Emergent Beam

The size of the waist of the emergent beam will be found. The analysis makes use of the fact that the q parameter becomes a pure imaginary number at the waist. As Eq. (3.89) was obtained, the size of the waist W_1 is found from the imaginary part of Eq. (3.115) at $z = d_1$:

$$W_1^2 = -\frac{2}{k} \text{Im } q(d_1)$$

Inserting Eq. (3.97) into the above equation, the ratio M of the new waist to the old waist is

$$M = \left(\frac{W_1}{W_0} \right) = \frac{f}{\sqrt{(d_0 - f)^2 + (z_0)^2}} \quad (3.116)$$

The curve for M is plotted in Fig. 3.27a with z_0/f as a parameter.

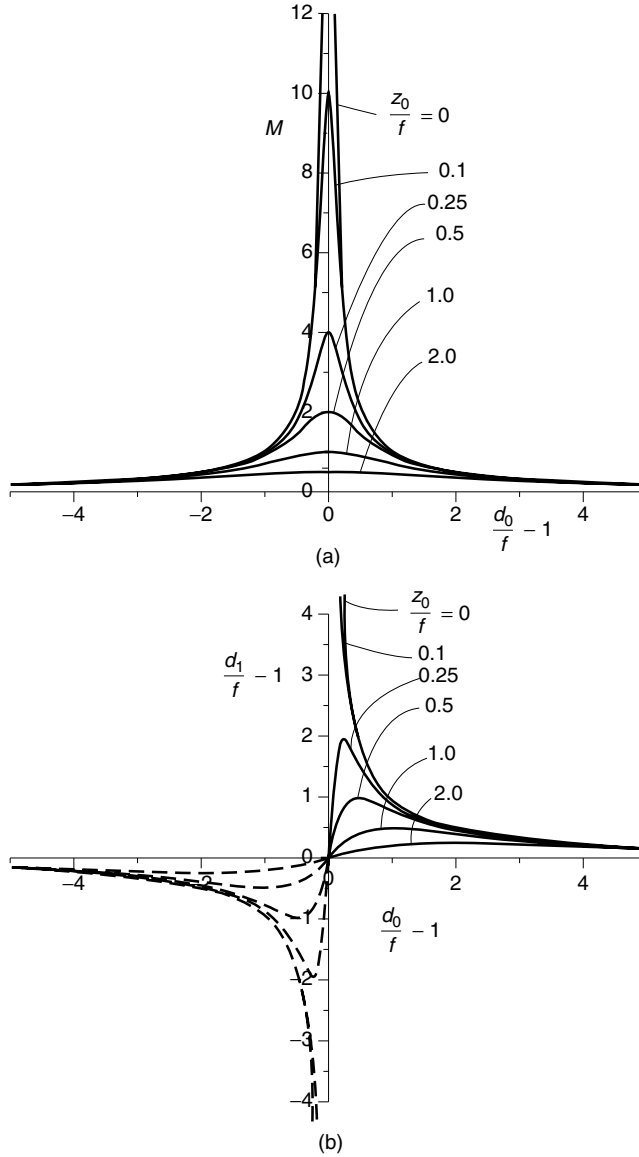


Figure 3.27 Magnification M of the beam waist and the location $(d_1/f - 1)$ of the emergent waist as a function of the location $(d_0/f - 1)$ of the input waist. (a) Magnification of the waist. (b) Location of the emergent waist.

3.6.3 Location of the Waist of the Emergent Beam

From the condition in Eq. (3.91), the location d_1 of the waist is found from the location where the real part vanishes as

$$d_1 = f + \frac{f^2(d_0 - f)}{(d_0 - f)^2 + z_0^2} \quad (3.117)$$

Combining Eq. (3.116) with (3.117) gives

$$d_1 - f = M^2(d_0 - f) \quad (3.118)$$

Graphs of $(d_1/f - 1)$ as a function of $(d_0/f - 1)$ are plotted in Fig. 3.27b, again with z_0/f as a parameter. In the negative $(d_0/f - 1)$ region, as shown by the dotted line, there is a virtual image of the waist.

3.6.4 Rayleigh Range of the Emergent Beam

The Rayleigh range z_1 of the emergent beam is given by the size of the waist W_1 of the emergent beam and

$$z_1 = \frac{k}{2} W_1^2$$

From Eqs. (3.97) and (3.116), the Rayleigh range is

$$z_1 = M^2 z_0 \quad (3.119)$$

The depth of focus D_1 of the emergent beam, from Eqs. (3.107) and (3.119), is therefore

$$D_1 = M^2 D_0 \quad (3.120)$$

Both z_0 and D_0 increase by a factor of M^2 .

3.6.5 Angle of the Far-Field Divergence of the Emergent Beam

The angle of the far-field divergence is obtained from

$$\lim_{z \rightarrow \infty} \tan \theta_1 = \lim_{z \rightarrow \infty} \left(\frac{W_1(z)}{z - d_1} \right)$$

For a small angle θ_1 , the following approximation holds:

$$\theta_1 \doteq \frac{2}{kW_1} \text{ rad}$$

and, hence,

$$\theta_1 = \frac{\theta_0}{M} \text{ rad} \quad (3.121)$$

3.6.6 Comparison with Ray Optics

In regions far away from the focal depth, the image formed by a Gaussian beam becomes closer to that formed by ray optics. As $(d_0 - f)$ becomes much larger than z_0 , the expression for the location of the Gaussian beam waist, Eq. (3.117), approaches

$$f^2 = (d_0 - f)(d_1 - f) \quad (3.122)$$

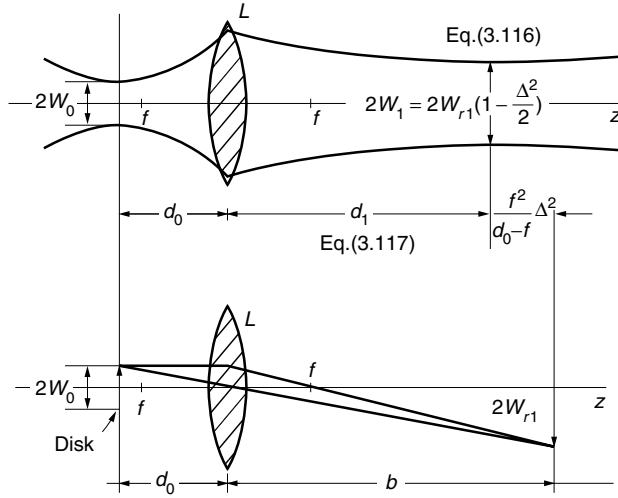


Figure 3.28 Gaussian beam output waist $2W_1$ and ray optics image $2W_{r1}$.

which is identical to Gauss's equation of imaging:

$$\frac{1}{f} = \frac{1}{d_0} + \frac{1}{d_1}.$$

As for the magnification, Gauss's expression for the magnification is

$$\beta = \frac{f}{d_0 - f} \quad (3.123)$$

and Eq. (3.116) also approaches Eq. (3.123) as

$$\Delta = \frac{z_0}{d_0 - f}$$

becomes smaller than unity.

More quantitatively, the case when the input waist is replaced by a circular disk object is calculated using Gauss's ray imaging formula and is compared in Fig. 3.28. The distance d_1 of the output waist of the Gaussian beam is shorter than b of the output disk image by

$$b - d_1 = \frac{f^2}{d_0 - f} \Delta^2$$

The dimension W_1 of the waist given by Eq. (3.116) is smaller than that of the disk approximately by a factor of $(1 - \Delta^2/2)$.

3.6.7 Summary of the Equations of the Transformation by a Lens

Equations describing the transformation are tabulated in Table 3.2. The associated figures are shown in Figs. 3.26 and 3.27.

Table 3.2 Gaussian beam transformation by a lens

Parameter	Input Beam	Output Beam	Equation Number
Beam waist	W_0	$W_{10} = MW_0$	(3.116)
Waist location	$d_0 - f$	$d_1 - f = M^2(d_0 - f)$	(3.118)
Rayleigh range	$z_0 = \frac{k}{2} W_0^2$	$z_1 = M^2 z_0 = \frac{k}{2} W_{10}^2$	(3.119)
Depth of focus	$D_0 = 2z_0$	$D_1 = M^2 D_0$	(3.120)
Angle of far-field divergence	$\theta_0 = \tan^{-1} \left(\frac{2}{kW_0} \right)$	$\theta_1 = \frac{\theta_0}{M}$	(3.121)
Magnification	1	$M^2 = \frac{f^2}{(d_0 - f)^2 + z_0^2}$	(3.116)
		$= \frac{d_1 - f}{d_0 - f}$	(3.118)
		$= \frac{\beta^2}{1 + \Delta^2}$	
		where	
		$\beta = \frac{f}{d_0 - f}$	(3.123)
		$\Delta = \frac{z_0}{d_0 - f}$	

If we shift the origin of the z coordinate from the lens to the emergent beam waist and name the new coordinate z' , Eqs. (3.98) and (3.99) for the beam radius and the radius of curvature for the input beam are converted into those for the emergent beam and expressed as

$$\begin{aligned}
 W_1^2(z') &= W_{10}^2 \left[1 + \left(\frac{z'}{z_1} \right)^2 \right] \\
 R_1(z') &= z' \left[1 + \left(\frac{z_1}{z'} \right)^2 \right]
 \end{aligned}
 \tag{3.124}$$

where W_{10} is the waist of the emergent beam.

3.6.8 Beam Propagation Factor m^2

Focused laser beams with small spot sizes are utilized for such applications as reading and writing a digital video disk (see Section 2.10.7), laser printer heads, and drilling holes in a stainless steel sheet. Figure 3.29 shows a configuration for reducing the waist (spot size) of a laser beam [17]. A large-diameter Gaussian beam waist is incident onto the surface of a lens as shown in Fig. 3.29. The wavefront is parallel to the lens surface and the light intensity along the lens surface decays as the edge of the lens is approached. This bell-shaped distribution contributes to the apodizing effect (see Section 1.4.2) and no harmful side lobes appear in the focused light spot. The location

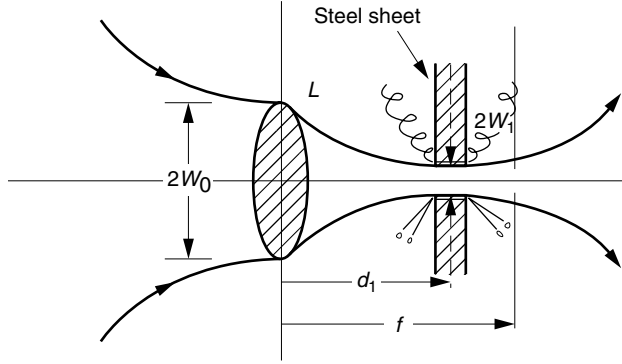


Figure 3.29 Gaussian beam drill spot by a lens. (Input waist is on the surface of the lens).

d_1 of the emergent waist is obtained from Eq. (3.117) with $d_0 = 0$.

$$d_1 = f \frac{1}{1 + (f/z_0)^2} \quad (3.125)$$

The waist appears at a location shorter than the focal length of the lens and the size of the waist is, from Eq. (3.116),

$$\left(\frac{W_1}{W_0}\right)^2 = \frac{1}{1 + (z_0/f)^2}$$

When a short focal length lens with $z_0 \gg f$ is used with Eq. (3.97), the output spot size diameter $2W_1$ approximately becomes

$$2W_1 = \frac{4\lambda f}{\pi D} \quad (3.126)$$

where the input waist $2W_0$ is chosen as wide as the lens diameter D . Thus, a smaller laser beam spot is obtained if a shorter wavelength, or a shorter focal length, or a larger diameter lens is selected.

If a lens with a diameter of $D = 4f/\pi$ is used, a spot size of one wavelength should be obtainable. In practice, however, the input laser beam is not a perfect Gaussian beam. This imperfection has been shown to increase the spot size, and in actual situations, the spot size becomes m times as large as $2W_1$, and the Rayleigh range becomes

$$z'_0 = m^2 z_0 \quad (3.127)$$

The factor m^2 is called the m^2 beam propagation factor and is used for expressing the quality of the incident laser beam. As a matter of fact, when the incident light is a quasi-Gaussian beam, approximate values of $W_1(z')$ and $R_1(z')$ in Eq. (3.124) are calculated by simply replacing z_1 by $m^2 z_1$ [18].

Example 3.10 Write out a sequence of equations (just equations) to transfer the q parameter through a convex lens followed by a concave lens whose focal lengths are f_1 and f_2 and with spacing among the source, lenses, and observation point as shown in Fig. 3.30.

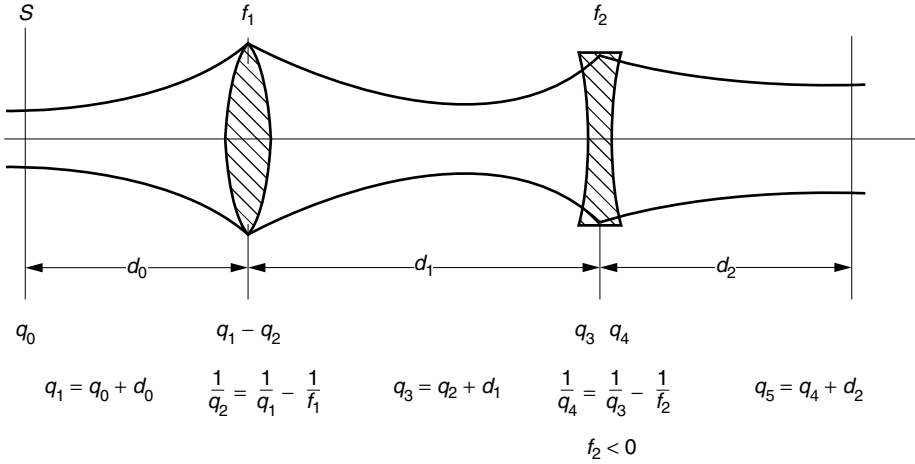


Figure 3.30 Transformation of q parameters; f_1 is a positive quantity, and f_2 is a negative quantity.

Solution Starting from point S , the sequence of equations for the q parameter is

$$\begin{aligned}
 q &= q_0 \\
 q_1 &= q_0 + d_0 \\
 \frac{1}{q_2} &= \frac{1}{q_1} - \frac{1}{f_1} \\
 q_3 &= q_2 + d_1 \\
 \frac{1}{q_4} &= \frac{1}{q_3} - \frac{1}{f_2}, \quad \text{where } f_2 \text{ is a negative quantity} \\
 q_5 &= q_4 + d_2
 \end{aligned}$$

□

Example 3.11 A Gaussian beam with the following parameters is incident onto a convex lens. Find the location and size of the waist of the emergent beam.

$$\begin{aligned}
 f &= 10 \text{ cm} \\
 d_0 &= 20 \text{ cm} \\
 W_0 &= 0.1 \text{ mm} \\
 \lambda &= 0.63 \text{ } \mu\text{m}
 \end{aligned}$$

Solution From Eq. (3.117), the location of the emergent beam waist is calculated as

$$\begin{aligned}
 d_1 - f &= \frac{100^2(200 - 100)}{(200 - 100)^2 + (\pi/0.63 \times 10^{-3})^2(0.1)^4} \\
 &= 80 \text{ mm} \\
 d_1 &= 18 \text{ cm}
 \end{aligned}$$

Note that the contribution of $[(k/2)W_0^2]^2$ is not negligible and, but for this, $d_1 = 20$ cm. From Eqs. (3.116) and (3.118), the emergent beam waist is

$$W_1 = 0.1 \sqrt{\frac{8}{10}} = 0.09 \text{ mm} \quad \square$$

3.7 HERMITE GAUSSIAN BEAM (HIGHER ORDER MODES)

So far, the value of A in the solution

$$\psi = Ae^{j[P+k(x^2+y^2)/2q]} \quad (3.75)$$

was assumed a constant, but this is not the only solution. A more general solution ψ will now be obtained by assuming A is a function of x , y , and z [19]. The derivation of the solution is slightly lengthy, but many interesting tricks are used to find the solution, and the derivation is well worth following. Let

$$\Psi = g\left(\frac{x}{W}\right)h\left(\frac{y}{W}\right)\psi(x, y, z)e^{j\Gamma(z)} \quad (3.128)$$

where $\psi(x, y, z)$ is the solution that has already been obtained with A equal to a constant, given by Eq. (3.75). A was replaced by $g(x/W)h(y/W)e^{j\Gamma(z)}$, where $W(z)$ is presented here without the argument z . The functions $g(x/W)$ and $h(y/W)$ that enable Eq. (3.128) to satisfy

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + j2k \frac{\partial \Psi}{\partial z} = 0$$

will be found.

The partial derivatives of Ψ are

$$\frac{\partial^2 \Psi}{\partial x^2} = \left(\frac{1}{W^2} g'' h \psi + 2j \frac{kx}{Wq} g' h \psi + gh \frac{\partial^2 \psi}{\partial x^2} \right) e^{j\Gamma} \quad (3.129)$$

$$\frac{\partial^2 \Psi}{\partial y^2} = \left(\frac{1}{W^2} gh'' \psi + 2j \frac{ky}{Wq} gh' \psi + gh \frac{\partial^2 \psi}{\partial y^2} \right) e^{j\Gamma} \quad (3.130)$$

$$\begin{aligned} j2k \frac{\partial \Psi}{\partial z} = & \left[j2kx \frac{\partial}{\partial z} \left(\frac{1}{W} \right) g' h \psi + j2ky \frac{\partial}{\partial z} \left(\frac{1}{W} \right) gh' \psi \right. \\ & \left. - 2gh\psi k \frac{\partial \Gamma}{\partial z} + j2kgh \frac{\partial \psi}{\partial z} \right] e^{j\Gamma} \end{aligned} \quad (3.131)$$

where

$$g' = \frac{\partial g}{\partial (x/W)}, \quad h' = \frac{\partial h}{\partial (y/W)}$$

The sum of the last terms in Eqs. (3.129), (3.130), and (3.131) are zero because ψ itself satisfies Eq. (3.73). The total sum divided by Ψ/W^2 gives

$$\frac{g''}{g} + \frac{h''}{h} + j2 \left(x \frac{g'}{g} + y \frac{h'}{h} \right) \left[\frac{kW}{q} + kW^2 \frac{\partial}{\partial z} \left(\frac{1}{W} \right) \right] - 2kW^2 \frac{\partial \Gamma}{\partial z} = 0 \quad (3.132)$$

The factor in the third term can be simplified by noting that

$$\frac{\partial}{\partial z} \left(\frac{1}{W} \right) = -\frac{1}{W^2} \frac{\partial W}{\partial z} \quad (3.133)$$

From Eq. (3.95), $\partial W/\partial z$ is

$$\frac{\partial W}{\partial z} = \frac{1}{W} \frac{4z}{(kW_0)^2} \quad (3.134)$$

With Eqs. (3.100), (3.133), and (3.134), the last term of the last bracket of Eq. (3.132) becomes

$$kW^2 \frac{\partial}{\partial z} \left(\frac{1}{W} \right) = -\frac{kW}{R(z)} \quad (3.135)$$

With Eqs. (3.83), (3.87), and (3.135), the sum inside the last bracket of Eq. (3.132) becomes a single number $j(2/W)$. Thus, Eq. (3.132) finally becomes

$$\underbrace{\frac{g''}{g} - 4 \frac{x}{W} \frac{g'}{g}}_{\text{Function of } x} + \underbrace{\frac{h''}{h} - 4 \frac{y}{W} \frac{h'}{h}}_{\text{Function of } y} - \underbrace{2kW^2 \frac{\partial \Gamma}{\partial z}}_{\text{Function of } z} = 0 \quad (3.136)$$

The differential equation, Eq. (3.136), is solved by the method of separation of variables.

The function of x in Eq. (3.136) is

$$\frac{1}{g} \frac{d^2 g(\bar{x})}{d\bar{x}^2} - 4\bar{x} \frac{1}{g} \frac{dg(\bar{x})}{d\bar{x}} \quad (3.137)$$

where

$$\bar{x} = \frac{x}{W}$$

The differential equation of a Hermite polynomial of order n is

$$\frac{\partial^2 H_n}{\partial x^2} - 2x \frac{\partial H_n}{\partial x} + 2nH_n(x) = 0 \quad (3.138)$$

where n is a positive integer.

Table 3.3 The Hermite polynomials of the n th order

n	$H_n(x)$
0	1
1	$2x$
2	$4x^2 - 2$
3	$8x^3 - 12x$

The expressions of $H_n(x)$ are shown in Table 3.3.

We see that $g(\bar{x})$ almost satisfies the differential equation of the Hermite polynomial except that the coefficient is 4 instead of 2 in the first order derivative. Inserting the change of variables

$$s = \sqrt{2}\bar{x}$$

$$\frac{dg}{d\bar{x}} = \sqrt{2} \frac{dg}{ds}$$

$$\frac{d^2g}{d\bar{x}^2} = 2 \frac{d^2g}{ds^2}$$

into Eq. (3.137) gives

$$\frac{2}{g(s)} \left(\frac{d^2g(s)}{ds^2} - 2s \frac{dg(s)}{ds} \right)$$

The coefficient of the first derivative is 2 and fits for the differential equation of the Hermite polynomial. A similar change of variables is made for $\bar{y} = y/W$.

$$\bar{y} = \frac{y}{W}$$

$$t = \sqrt{2}\bar{y}$$

Equation (2.136) now becomes

$$\underbrace{\frac{1}{g} \left(\frac{d^2g}{ds^2} - 2s \frac{dg}{ds} \right)}_{-2n} + \underbrace{\frac{1}{h} \left(\frac{d^2h}{dt^2} - 2t \frac{dh}{dt} \right)}_{-2m} \underbrace{-kW^2 \frac{\partial \Gamma}{\partial z}}_{2(m+n)} = 0 \quad (3.139)$$

The first two terms are functions of s only, the second two terms are functions of t only, and the last term is a function of z only. For Eq. (3.139) to be satisfied everywhere in space, the functions of x , y , and z have to be independently constant: $-2n$, $-2m$, and $2(n+m)$.

$$\frac{d^2g(s)}{ds^2} - 2s \frac{dg(s)}{ds} + 2ng(s) = 0 \quad (3.140a)$$

$$\frac{d^2h(t)}{dt^2} - 2t \frac{dh(t)}{dt} + 2mh(t) = 0 \quad (3.140b)$$

$$kW^2 \frac{\partial \Gamma}{\partial z} = -2(m+n) \quad (3.140c)$$

Hence, from Eq. (3.138),

$$g(s) = H_n \left(\sqrt{2} \frac{x}{W} \right)$$

$$h(t) = H_m \left(\sqrt{2} \frac{y}{W} \right)$$

The value of Γ is found by inserting Eq. (3.95) into Eq. (3.140c):

$$\frac{\partial \Gamma}{\partial z} = - \frac{2(n+m)}{kW_0^2 [1 + (2z/kW_0^2)^2]} \quad (3.141)$$

and

$$\Gamma = -(n+m) \tan^{-1} \left(\frac{2z}{kW_0^2} \right) \quad (3.142)$$

Finally, the expression for a higher order Gaussian beam, which is known as the Hermite Gaussian beam, is

$$\begin{aligned} \Psi = & \frac{W_0}{W} H_m \left(\sqrt{2} \frac{x}{W} \right) H_n \left(\sqrt{2} \frac{y}{W} \right) \\ & \exp \left\{ j \left[\frac{k}{2q} (x^2 + y^2) - (m+n+1) \tan^{-1} \left(\frac{2z}{kW_0^2} \right) \right] \right\} \end{aligned} \quad (3.143)$$

Figure 3.31 shows measured mode patterns inside a cavity with concave mirrors. The TEM_{00} mode in Fig. 3.31 corresponds to the case when $m=n=0$, where $H_0(\sqrt{2}x/W) = H_0(\sqrt{2}y/W) = 1$ and Eq. (3.143) reduces exactly to the fundamental mode given by either Eq. (3.75) or (3.104). The cross-section field distribution has its maximum on the z axis and has a Gaussian distribution, such as shown in Fig. 3.25.

The TEM_{11} mode uses the first order Hermite polynomials in both the x and y directions. Using the values in Table 3.3, the front factor of Eq. (3.143) becomes

$$\left(\frac{W_0}{W} \right) \left(2\sqrt{2} \frac{x}{W} \right) \left(2\sqrt{2} \frac{y}{W} \right) = \frac{8W_0}{W^3} xy$$

Thus, the TEM_{11} mode is zero at the origin.

As a matter of fact, the mode number corresponds to the number of zero fields in the transverse plane. Different modes have different field distributions. The operation of the “light tweezers” mentioned later in this chapter critically depends on the field distribution; hence, the proper mode configuration is essential to the operation of the light tweezers.

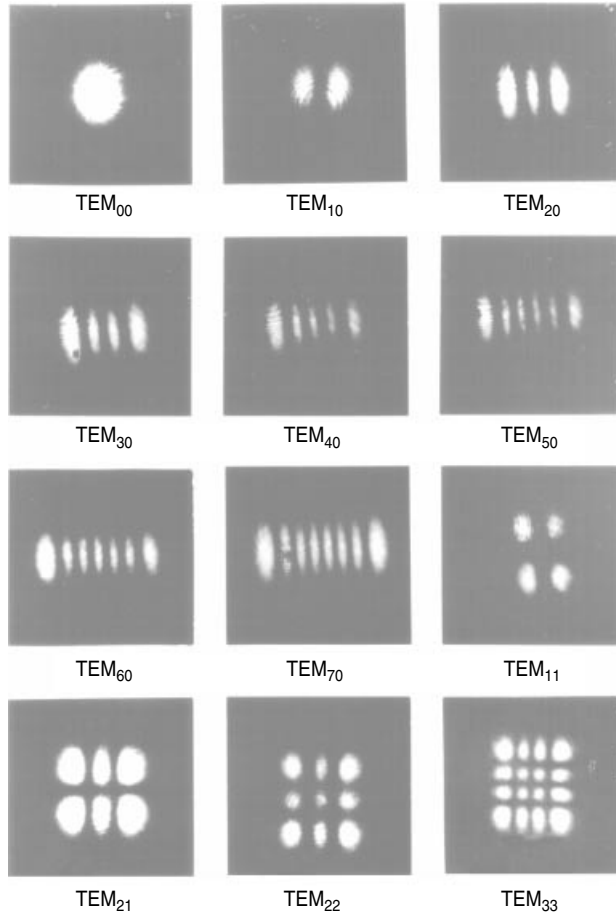


Figure 3.31 Measured mode patterns inside a cavity with concave mirrors. (Courtesy of H. Kogelnik and T. Li [16].)

These different modes have different phase velocities and possess different cavity resonance frequencies. In order to properly interpret the output from a Fabry–Pérot resonator, an understanding of the modes is important.

3.8 THE GAUSSIAN BEAM IN A SPHERICAL MIRROR CAVITY

Since around 1960, spherical mirror cavities have been widely used as Fabry–Pérot interferometers for analyzing the spectrum of light. The sharp frequency dependence of the transmitted light from the cavity is used to find the frequency spectrum of the incident light.

Most of the cavities of gas lasers are made of two concave mirrors facing each other. The actual laser cavities have a gain medium inside the cavity, and the mirrors are not 100% reflective so that some of the laser energy escapes the laser cavity. The characteristics of the laser cavity, however, can be well approximated by those in an

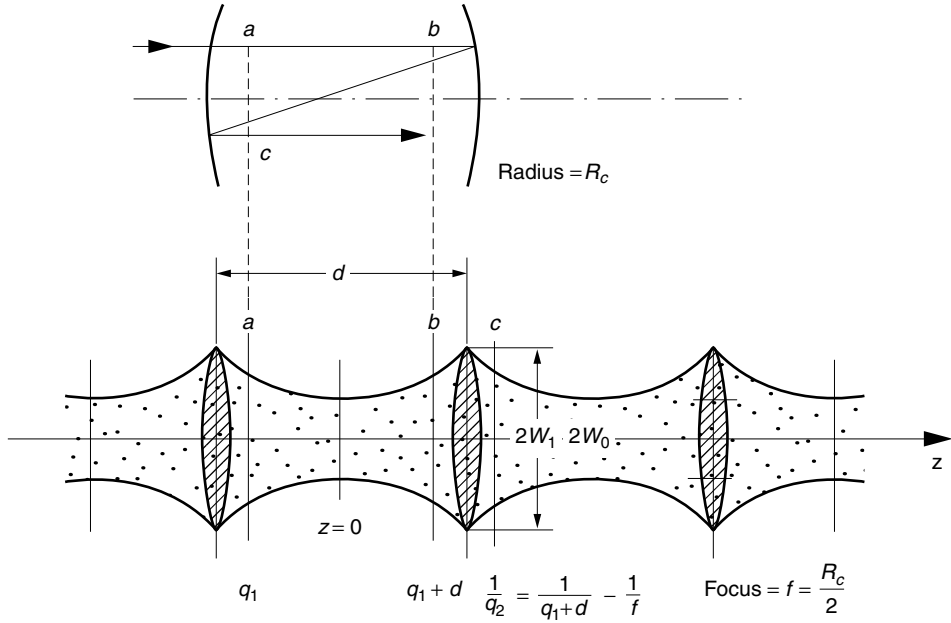


Figure 3.32 Equivalence of the multiple path in a cavity and the path in a lens array with focal length $f = R_c/2$.

ideal air-filled cavity with 100% reflecting, infinitely large concave mirrors. Figure 3.32 shows the structure of a cavity with two identical mirrors. At resonance, the phases of the beam before and after the round trip have to match. This phase match has to take place not only at the center of the beam but also in the entire plane transverse to the laser beam. Such a condition can be realized if the q parameters of the beam before and after the round trip are identical. The resonance condition will be expressed in terms of the radius of curvature and spacing between the mirrors. The repeated reflections by a spheroidal mirror with radius of curvature R_c is equivalent to transmission through an infinite array of lenses with focal length

$$f = \frac{R_c}{2} \quad (3.144)$$

In the present configuration, identical end mirrors are used and the condition after one-half round trip is sufficient.

Referring to Fig. 3.32, let the q parameter immediately after the lens be q_1 and the q parameter immediately after the next lens be q_2 . Then the q parameters q_1 and q_2 satisfy

$$\frac{1}{q_2} = \frac{1}{q_1 + d} - \frac{1}{f} \quad (3.145)$$

The resonance condition is

$$q_1 = q_2 = q \quad (3.146)$$

With Eq. (3.146), Eq. (3.145) becomes

$$\frac{df + q(q + d)}{fq(q + d)} = 0 \quad (3.147)$$

Setting the numerator in Eq. (3.147) equal to zero, and dividing by fdq^2 , the result is

$$\frac{1}{q^2} + \frac{1}{qf} + \frac{1}{fd} = 0 \quad (3.148)$$

Thus, the required q parameter is

$$\frac{1}{q} = \frac{-1}{2f} \pm j\sqrt{\frac{1}{fd} - \frac{1}{4f^2}} \quad (3.149)$$

The $+j$ term in Eq. (3.149) is not physically acceptable, as the field intensity would grow indefinitely as $(x^2 + y^2)$ increases. The real and imaginary parts of $1/q$ represent $1/R$ and $1/Q$ as in Eq. (3.83). At resonance, from the real part of Eq. (3.149), the radius of curvature of the beam just past the lens or just in front of the mirror is $2f$. Using Eq. (3.144),

$$R = -R_c \quad (3.150)$$

The imaginary part of Eq. (3.149) provides the size of the beam W_1 just in front of the mirror and is, from Eq. (3.87),

$$W_1^2 = \frac{2R_c}{k\sqrt{2R_c/d - 1}} \quad (3.151)$$

Next, let us find the beam waist. Equation (3.100) is the relationship developed taking the origin at the waist, as shown in Fig. 3.32. Assuming that the radius of the beam immediately after the lens is the same as that immediately before the lens, W_1 is expressed as

$$W\left(\frac{d}{2}\right) = W_1$$

The radius of curvature immediately before the lens is the same as that immediately after the lens except for the sign reversal and, from Eq. (3.150),

$$R\left(\frac{d}{2}\right) = R_c$$

Thus, Eq. (3.100) with $z = d/2$ becomes

$$W_1^2 = \frac{4R_c}{k^2 W_0^2} \cdot \frac{d}{2} \quad (3.152)$$

Inserting Eq. (3.151) into (3.152) gives

$$W_0^2 = \frac{d}{k} \sqrt{\frac{2R_c}{d} - 1} \quad (3.153)$$

Let us interpret the results obtained. Equation (3.150) states that, at resonance, the radius of curvature of the wave matches the radius of curvature of the end reflector mirror. In other words, the contour of the constant phase matches the surface of the concave mirror. According to Eq. (3.151), for W_1^2 to be real, there is a range restriction on d for a given R_c :

$$0 < d < 2R_c \quad (3.154)$$

Figure 3.33 shows the field inside a resonant cavity as the spacing d is varied for a fixed radius of curvature R_c . As the spacing is increased, the beam radius W on the surface of the end mirrors expands, whereas the size of the waist in the middle shrinks. As soon as d exceeds $2R_c$, the beam is no longer contained inside the cavity. In this case the cavity becomes unstable, and the resonance no longer exists, as shown in Fig. 3.33e.

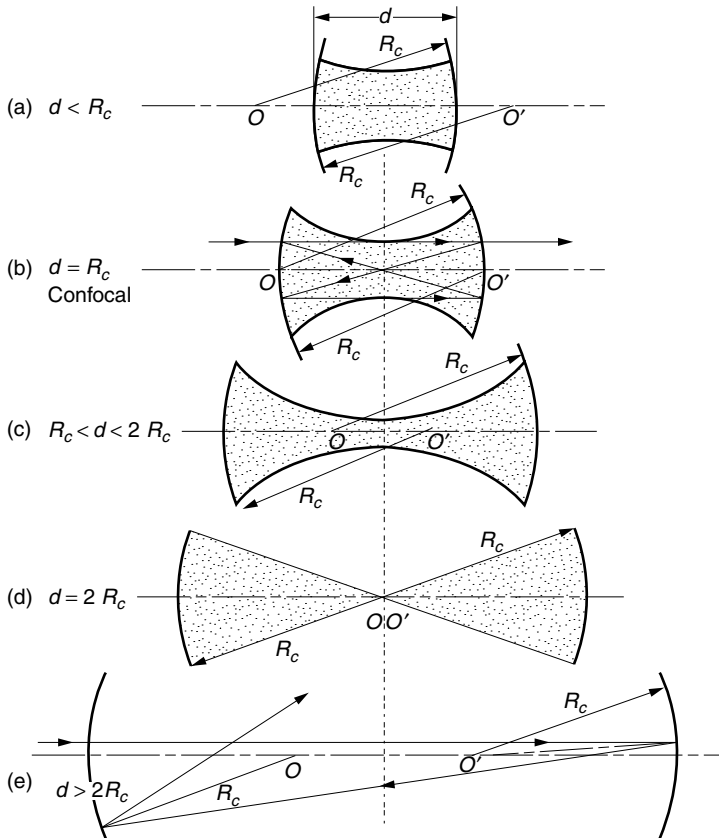


Figure 3.33 Variation of the field distribution inside a cavity with spacing d for a fixed R_c .

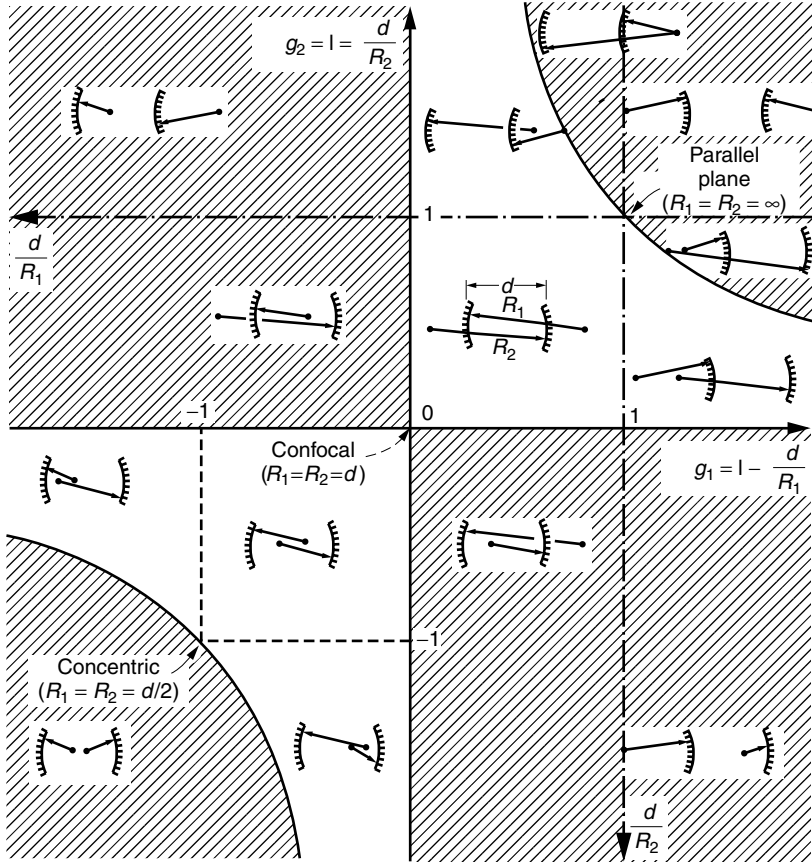


Figure 3.34 Stability diagram. Unstable resonator systems lie in shaded regions. (After H. Kogelnik and T. Li [16].)

The range of existence of the resonance for various combinations of the radii of the end mirrors and the spacing between the mirrors has been studied by Kogelnik and Li [16]. This range is illustrated in Fig. 3.34. The conclusion of the Kogelnik–Li study gives the range of the stable resonance as

$$0 < g_1 g_2 < 1 \quad (3.155)$$

where

$$\begin{aligned} g_1 &= 1 - \frac{d}{R_1} \\ g_2 &= 1 - \frac{d}{R_2} \end{aligned} \quad (3.156)$$

and where R_1 and R_2 are the radii of curvature of the end mirrors. In the present case of

$$R = R_1 = R_2 = R_c$$

the cavity becomes unstable when

$$g_1 g_2 > 1 \quad (3.157)$$

which is equivalent to $d > 2R_c$, as given in Fig. 3.33e.

3.9 RESONANCE FREQUENCIES OF THE CAVITY

Excitation of the cavity by nothing but the fundamental mode or longitudinal mode ($m + n = 0$) can be achieved only when the incident beam is of the fundamental mode and when the mirrors are on the same sphere and perfectly aligned, or the radius of curvature $R_c \rightarrow \infty$ as in the case of planar mirrors. It is easier to achieve this condition when the radius of the beam is narrower. In general, higher order modes ($m + n \neq 0$) are expected to be present.

Let us first find the resonance frequency of the fundamental mode in a cavity with identical end mirrors. Resonance takes place when the phases of the beam before and after a round trip are in phase.

For a cavity with identical end mirrors, the waist is located in the middle of the cavity. The phase distribution in the cavity is readily pictured if the origin of Fig. 3.25a is set in the center of the cavity. The phase of the field at the right end mirror is identical with that in Fig. 3.25a with $z = d/2$, where d is the spacing between the mirrors. The phase correction factor ϕ in Eq. (3.104) or in Eq. (3.143) with $m + n = 0$ can be found from Eq. (3.102) with $z = d/2$.

$$\phi = \tan^{-1} \frac{d}{kW_0^2} \quad (3.158)$$

The value of W_0^2 in Eq. (3.158) can readily be obtained from Eq. (3.153), and ϕ becomes

$$\phi = \tan^{-1} \frac{1}{\sqrt{2R_c/d - 1}} \quad (3.159)$$

For every distance of $d/2$, an additional delay of $-\phi$ exists. The phase delay for one round trip $2d$ has to be an integral multiple of 2π . The round-trip phase delay is

$$2d \frac{2\pi f_p}{v} - 4 \tan^{-1} \frac{1}{\sqrt{2R_c/d - 1}} = 2\pi p \quad (3.160)$$

where v is the velocity of light, p is an integer, and the resonance frequency f_p is called the resonance frequency of the p th longitudinal mode. Solving for f_p , Eq. (3.160) becomes

$$f_p = \frac{v}{2d} \left(p + \frac{2}{\pi} \tan^{-1} \frac{1}{\sqrt{2R_c/d - 1}} \right) \quad (3.161)$$

The frequency spacing Δf between the p th and $(p + 1)$ th longitudinal mode is

$$\Delta f_0 = \frac{v}{2d} \quad (3.162)$$

The resonance frequencies are equally spaced, as shown in Fig. 3.35a, and are the same as those of a planar mirror resonator.

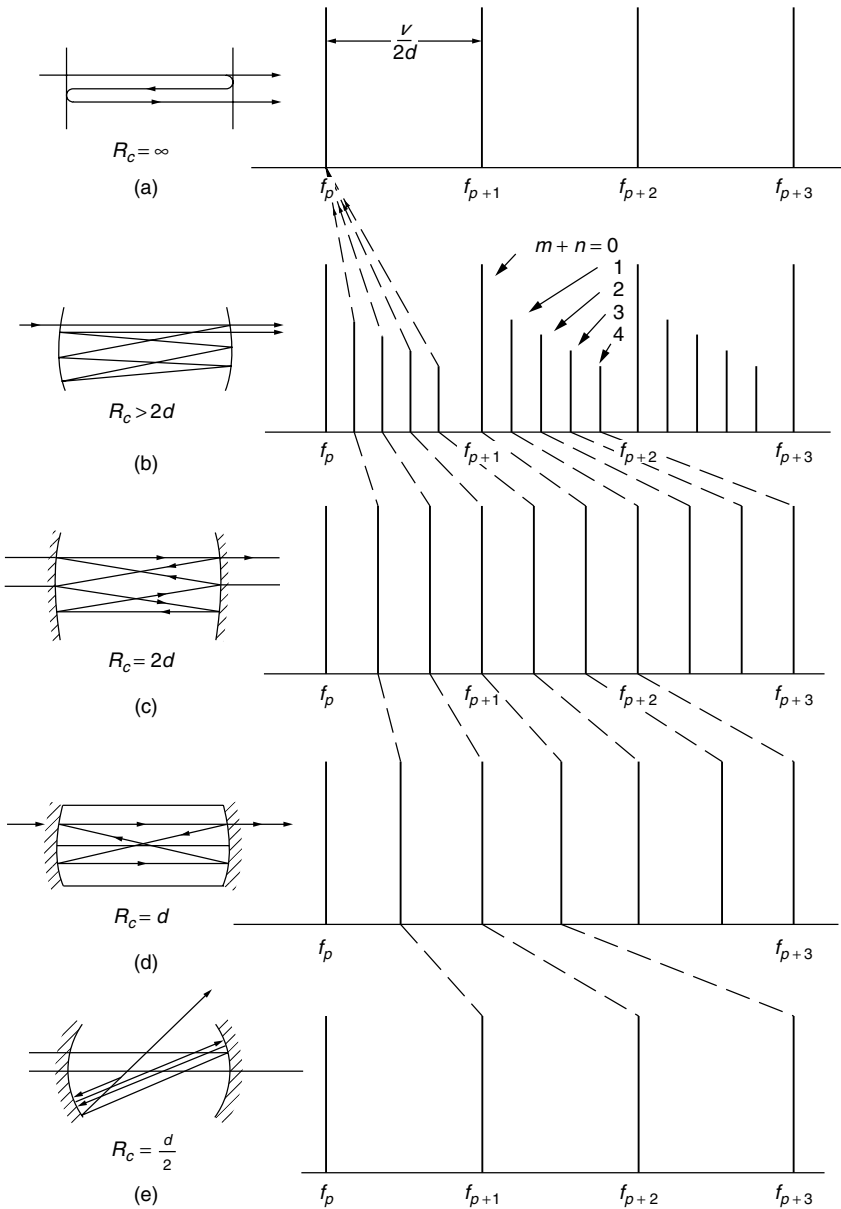


Figure 3.35 Frequency spectrum of a Fabry-perot interferometer. (a) Planar mirror cavity. (b) General spherical mirror cavity. (c) Focal cavity. (Foci are on the confronting mirrors.) (d) Confocal cavity. (Foci of the two mirrors coincide.) (e) Spherical cavity. (Mirrors are on the same sphere.)

Next, the resonance frequencies of the higher order modes ($m + n \neq 0$) or transverse modes will be obtained. The additional phase delay Γ has to be included. From Eqs. (3.142) with $z = d/2$ and (3.153), the phase delay Γ is

$$\Gamma = -(m + n) \tan^{-1} \frac{1}{\sqrt{2R_c/d - 1}} \quad (3.163)$$

which is similar to Eq. (3.159) except for $(m + n)$. The total phase correction factor at every $d/2$ is $\phi + \Gamma$, and the equation of the resonance frequency $f_{p,m,n}$ becomes

$$f_{p,m,n} = \frac{v}{2d} \left(p + \frac{2}{\pi}(m + n + 1) \tan^{-1} \frac{1}{\sqrt{2R_c/d - 1}} \right) \quad (3.164)$$

The spacing between $f_{p,0,0}$ and $f_{p,1,0}$ is

$$\Delta f_1 = \frac{v}{\pi d} \tan^{-1} \frac{1}{\sqrt{2R_c/d - 1}} \quad (3.165)$$

The spacing between $f_{p,1,0}$ and $f_{p,1,1}$ is also the same as Eq. (3.165). As long as the difference $\Delta(m + n)$ is unity, the spacing is given by Eq. (3.165) and if $\Delta(m + n) = 2$ the departure from f_p becomes twice Δf_1 . The spectrum with $f_{p,m,n}$ is shown in Fig. 3.35b. There is a cluster of higher order mode resonances at the higher frequency side of each fundamental mode. These higher order modes are disturbing when the cavity resonator is used as a Fabry-Pérot interferometer. As seen from Eq. (3.165), the spacing Δf_1 can be widened if a smaller value of R_c is chosen for a fixed d , as indicated by the dotted lines in Fig. 3.35. In fact, when $R_c = d$ (confocal cavity), Δf_1 in Eq. (3.165) becomes $v/4d$ and

$$\Delta f_1 = \frac{\Delta f_0}{2} \quad (3.166)$$

which means that the higher order mode shows up midway between adjacent resonant frequencies of the fundamental mode. When $R_c = d/2$, the higher order mode exactly overlaps the fundamental mode; however, as shown in Fig. 3.34, the focal cavity is on the edge of the stability condition and is not recommended. Of the cavities shown in Fig. 3.35, the confocal cavity of Fig. 3.35d is generally the most practical.

3.10 PRACTICAL ASPECTS OF THE FABRY-PÉROT INTERFEROMETER

On the basis of the conclusions reached in the previous section, the practical aspects of the Fabry-Pérot interferometer will be described.

3.10.1 Plane Mirror Cavity

When the mirror is perfectly flat, $R_c = \infty$ and Γ in Eq. (3.163) is zero, which means there is no concern about generating higher order modes. However, when flat mirrors are used for end mirrors, only a parallel beam can be excited in the cavity. In using such an interferometer, it is important to assure the parallelism of the mirrors as well as the perpendicular incidence of the light, otherwise the beam eventually wanders beyond the edges of the mirrors and the sharpness of the resonance (finesse) is poor. For instance, if the reflectivity of the mirror is 99.9%, the beam will bounce back and forth on average 1000 times before leaving the cavity. Unless the mirrors are perfectly parallel, this large number of bounces cannot successfully be completed. If the angle of incidence is other than the normal, the spacing between the mirrors is effectively

changed. This angular dependence is also used to analyze the spectrum but only with poor finesse.

The spacing between the resonant frequencies (free spectral range, FSR) is increased as the spacing between the mirrors is reduced, as indicated by Eq. (3.39). The derivative of Δf_0 with respect to d becomes large for small d , which means the interferometer is more susceptible to mechanical vibrations and temperature fluctuations. The FSR should be selected according to the width of the spectrum of the light under test (see Section 3.2.1.3).

3.10.2 General Spherical Mirror Cavity

Only with a narrow beam and a perfect mirror arrangement can the zeroth order beam be excited, otherwise the higher order modes are excited, as indicated in Fig. 3.35b. If higher order modes are present, the interpretation of the spectrum is more difficult.

One way to interpret the generation of the higher order modes or transverse modes is as follows. A beam of finite size can be thought of as a bundle of narrow beams. The narrow beam at the center represents the longitudinal mode, and it takes a direct path across the cavity. For this direct path, the distance for the resonance condition is one round trip (across the cavity and back). As illustrated in Fig. 3.36, some of the narrow beams take a zigzag path around the cavity such that it takes several traversals of the cavity before the beam actually rejoins itself. Upon rejoining, the resonance condition requires that the rejoined beam be in phase with the original beam. For the zigzag path, the rejoining distance is greater than the round-trip distance of the longitudinal mode. If the zigzag path is equivalent to N round trips of the longitudinal mode, then the spectral separation of the higher order mode is $\Delta f_0/N$ because it is equivalent to a cavity N times as long.

Next, the higher order mode patterns are considered. As a matter of fact, the mode patterns shown earlier in Fig. 3.31 are nothing but standing wave patterns in the transverse plane generated by off-axis propagating wave components. For instance, the rays shown in Fig. 3.36c are represented by two component waves whose propagation vectors are \mathbf{k}_1 and \mathbf{k}_2 .

The transverse component of \mathbf{k}_1 is pointing up while that of \mathbf{k}_2 is pointing down. These two components pointing in opposite directions set up a standing wave pattern in the transverse plane.

3.10.3 Focal Cavity

When the length of the cavity becomes the focal length of the mirror,

$$d = \frac{R_c}{2} \quad (3.167)$$

the cavity is called a focal cavity. For the focal cavity, Eq. (3.165) is $\Delta f_1 = v/6d$ and Δf_1 is one-third of Δf_0 . This can be verified by tracing the beam reflected by concave mirrors of the cavity. As shown in Fig. 3.36c, after three round trips, the beam lines up with the original direction and the separation between the higher order spectra becomes one-third that of the fundamental modes. Every third spectra matches up with the fundamental, and there are two additional spectra between the fundamental modes. The tolerance of the incident angle is not as large as for the confocal cavity, as described next. The focal resonator is also on the borderline of stability.

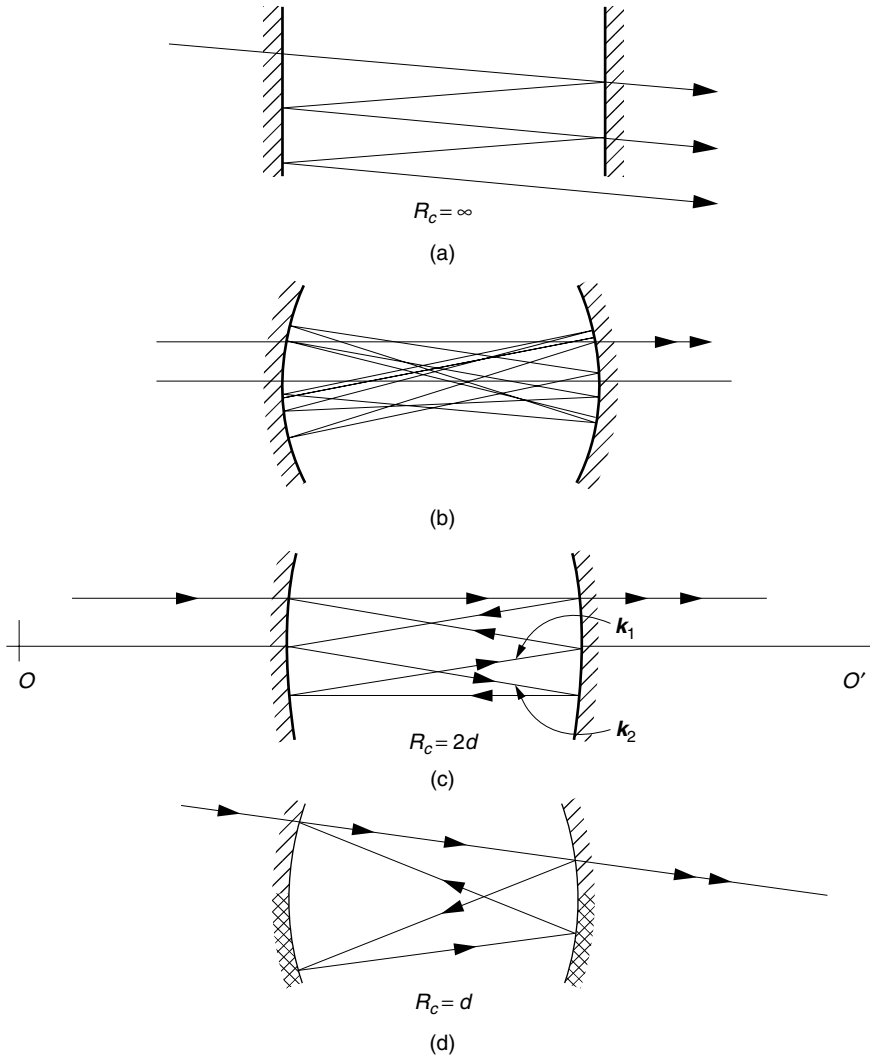


Figure 3.36 Various types of spherical mirrors in a Fabry-perot interferometer. (a) Planar mirror. (b) General. (c) Focal. (d) Confocal.

3.10.4 Confocal Cavity

When the radius of curvature of the spherical mirrors is identical to the spacing between the mirrors,

$$R_c = d \quad (3.168)$$

then such a cavity is called a confocal cavity. For the confocal cavity, Eq. (3.165) is $\Delta f_1 = v/4d$ and Δf_1 is one-half of Δf_0 . The spacing of the resonances associated with the transverse mode becomes exactly one-half of that of the longitudinal modes and the ambiguity due to the transverse mode disappears. This can be explained by tracing the path of light. As shown in Fig. 3.36d, after two round trips, the path of

the emergent beam lines up with the direction of the incident beam. This is as if the resonance took place in a cavity of twice the length, which means half the spectral separation of the fundamental modes. From the figure, the angle of incidence for the confocal cavity is far less critical than for the plane mirror cavity. The large tolerance on the angle of incidence is a definite advantage for the confocal resonator. A variation on this interferometer is to use end mirrors whose reflectivity on the bottom half of the mirror is much larger than the top half of the mirror. This allows easier entrance but higher reflection in the cavity. Such a Fabry–Pérot interferometer has a higher value of finesse.

3.11 BESSEL BEAMS

The Bessel beam propagates in free space with minimum spread in the transverse direction over distances of more than several meters. Because of this unusual property, the beam has been dubbed a diffraction-free beam.

3.11.1 Features of the Bessel Beam

Special features of the Bessel beam are the following:

1. The diameter of the intensity distribution in the transverse plane remains constant over distances of more than several meters.
2. The intensity decays abruptly at a distance $z = z_{\max}$.
3. The propagation constant along the beam is less than that of free space, and the value is adjustable.
4. With an adjustment of the condition of excitation, the n th order beams $J_n(\alpha r)$ can also be excited.

The constant-diameter feature is illustrated in Figs. 3.37–3.39. Figure 3.37 shows the intensity distributions of the Bessel beam in four transverse planes located at $z = 0$, $z = 2$ m, $z = 4$ m, and $z = 5.5$ m, where z is the distance along the beam from the input aperture [20]. The diameter of the main beam remains unchanged at $200\ \mu\text{m}$ over the entire distance of 5.5 m. Figure 3.38 is a photograph of the intensity distribution of a Bessel beam in the plane at $z = 9.6$ m [21] taken with the aid of a projector lens (see Fig. 3.42). The diameter of the main beam is $200\ \mu\text{m}$. In Fig. 3.39, the transverse field distributions of a Gaussian beam (dotted line) and Bessel beam (solid line) are compared in the following planes: (a) $z = 0$, (b) $z = 10$ cm, and (c) $z = 1$ m. The Gaussian beam displays diffraction spread as well as a rapid decrease in intensity. In order for the Gaussian beam intensity to be visible on the graph, the Gaussian curves (b) and (c) have been magnified 30 times and 2000 times, respectively. In contrast, there is no discernable change in the Bessel beam [22].

The abrupt decay of the Bessel beam at $z = z_{\max}$ is shown in Fig. 3.40. The field intensity of the main lobe of the Bessel beam is plotted with respect to the distance z of propagation. Again, for the sake of comparison, the Bessel beam is represented by the solid line and the Gaussian beam, by the dotted line. While the Gaussian beam displays a rapid decay, the Bessel beam maintains its intensity up to $z = z_{\max} = 1$ m, oscillating around a certain value with distance.

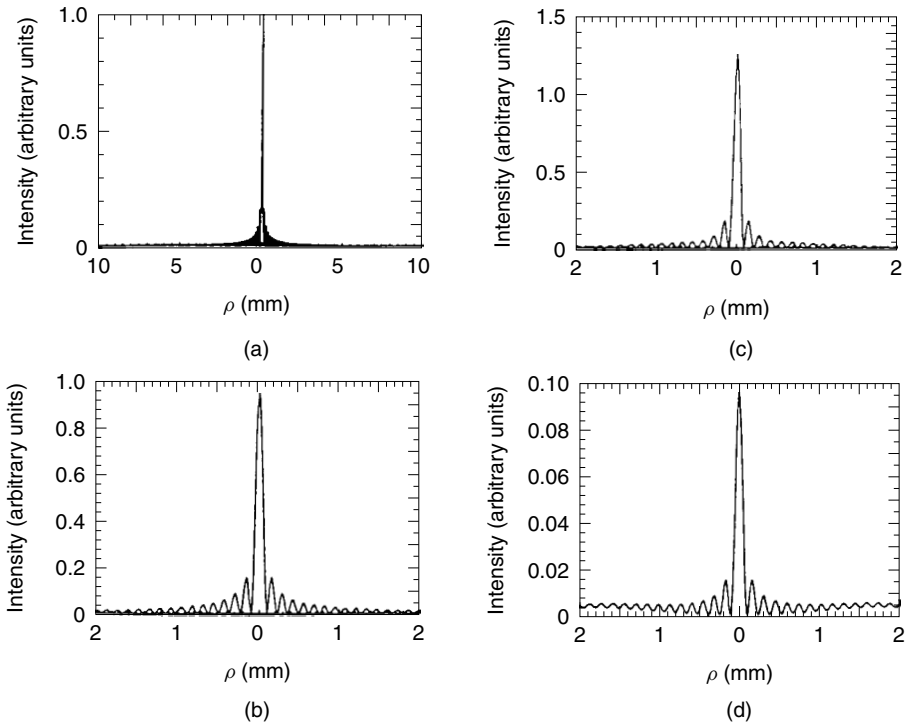


Figure 3.37 Intensity distributions for a Bessel beam: (a) $z = 0$ m, (b) $z = 2$ m, (c) $z = 4$ m, and (d) $z = 5.5$ m. (After J. Durnin [20].)

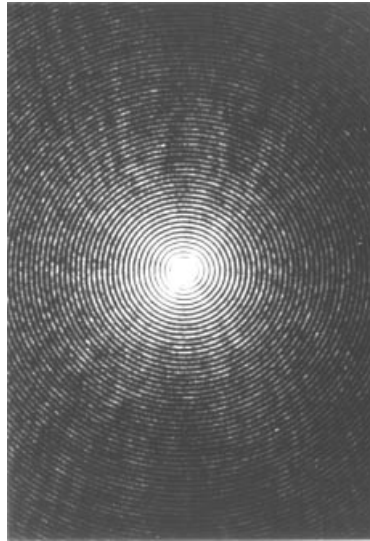


Figure 3.38 Photograph of the Bessel beam at 9.6 m from the aperture using the setup shown in Fig. 3.42c. (Courtesy of R. M. Herman and T. A. Wiggins [21].)

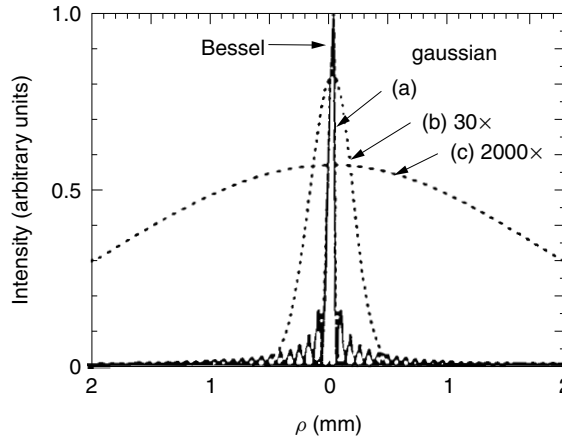


Figure 3.39 The transverse spread of the Bessel beam (solid line) is compared to that of a Gaussian beam (dotted line). The spot sizes of both beams are the same at $z = 0$. The transverse spread of the Gaussian beam is shown at (a) $z = 0$, (b) $z = 10$ cm with $30\times$ magnification in intensity, and (c) $z = 100$ cm with $2000\times$ magnification in intensity. The spot size of the Bessel beam remains the same. (After J. Durnin, J. J. Miceli Jr., and J. H. Eberly [22].)

3.11.2 Practical Applications of the Bessel Beam

A number of applications have been considered for incorporating the special features of the Bessel beam.

3.11.2.1 Precision Optical Measurement

Since the amount of spread during the propagation can be made much smaller than that of the Gaussian beam, the Bessel beam is useful for precision optical measurements [23].

Not only $J_0(\alpha r)$ but also $J_n(\alpha r)$ can be excited. The $J_1(\alpha r)$ beam, which has its null on the beam axis, may be advantageous for a finer definition in the precision alignment.

3.11.2.2 Power Transport

Due to the sharp fall-off of the transmission power at a predetermined distance, the Bessel beam might be useful as a means of power transport.

3.11.2.3 Nonlinear Optics

Bessel beams have been used for nonlinear optics experiments in long liquid cells of carbon disulfide, CS_2 , or acetone, $(\text{CH}_3)_2\text{CO}$ [24–26]. The beam provides not only a long interaction length but also confined high-intensity light, both of which are essential in nonlinear optics experiments.

An additional advantage of the Bessel beam is easy adjustability of the propagation constant, which is often needed for optimizing the nonlinear interaction [26].

3.11.3 One-Dimensional Model

The one-dimensional model in Figs. 3.41a and 3.14b will be used to explain how the Bessel beam is generated. The z axis is taken in the beam direction and the x and

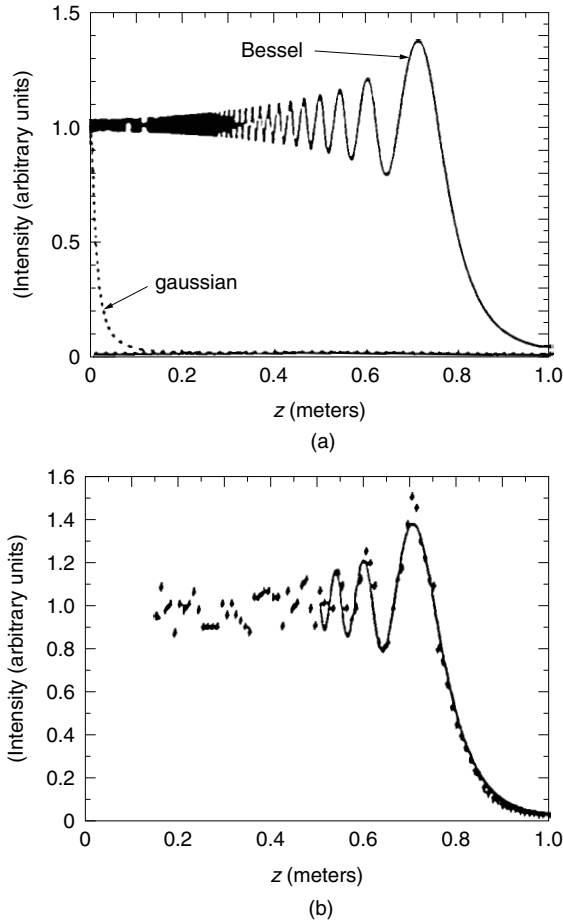


Figure 3.40 (a) Calculated on-axis intensity of the Bessel beam (solid line) and Gaussian beam (dotted line) with respect to the propagation distance z . (b) The corresponding measured results for the Bessel beam. (After J. Durnin, J. J. Miceli, Jr., and J. H. Eberly [22].)

y axes are in the plane transverse to the beam direction. Two delta function sources $\delta(x - a)$ and $\delta(x + a)$ are placed in the front focal plane of a convex lens L , at $x = a$ and $x = -a$. Since the sources are in the front focal plane, the light emerging from the lens is in the form of two parallel beams B_1 and B_2 with angles of inclination $\theta = \pm \tan^{-1}(a/f)$, where f is the focal length of the lens. The resultant field $E(x, z)$ is expressed as

$$\begin{aligned} E(x, z) &= A(e^{j(\beta z + \alpha x)} + e^{j(\beta z - \alpha x)}) \\ &= 2A \cos \alpha x e^{j\beta z} \end{aligned} \quad (3.169)$$

where

$$\beta = k \cos \theta \quad (3.170)$$

$$\alpha = k \sin \theta \quad (3.171)$$

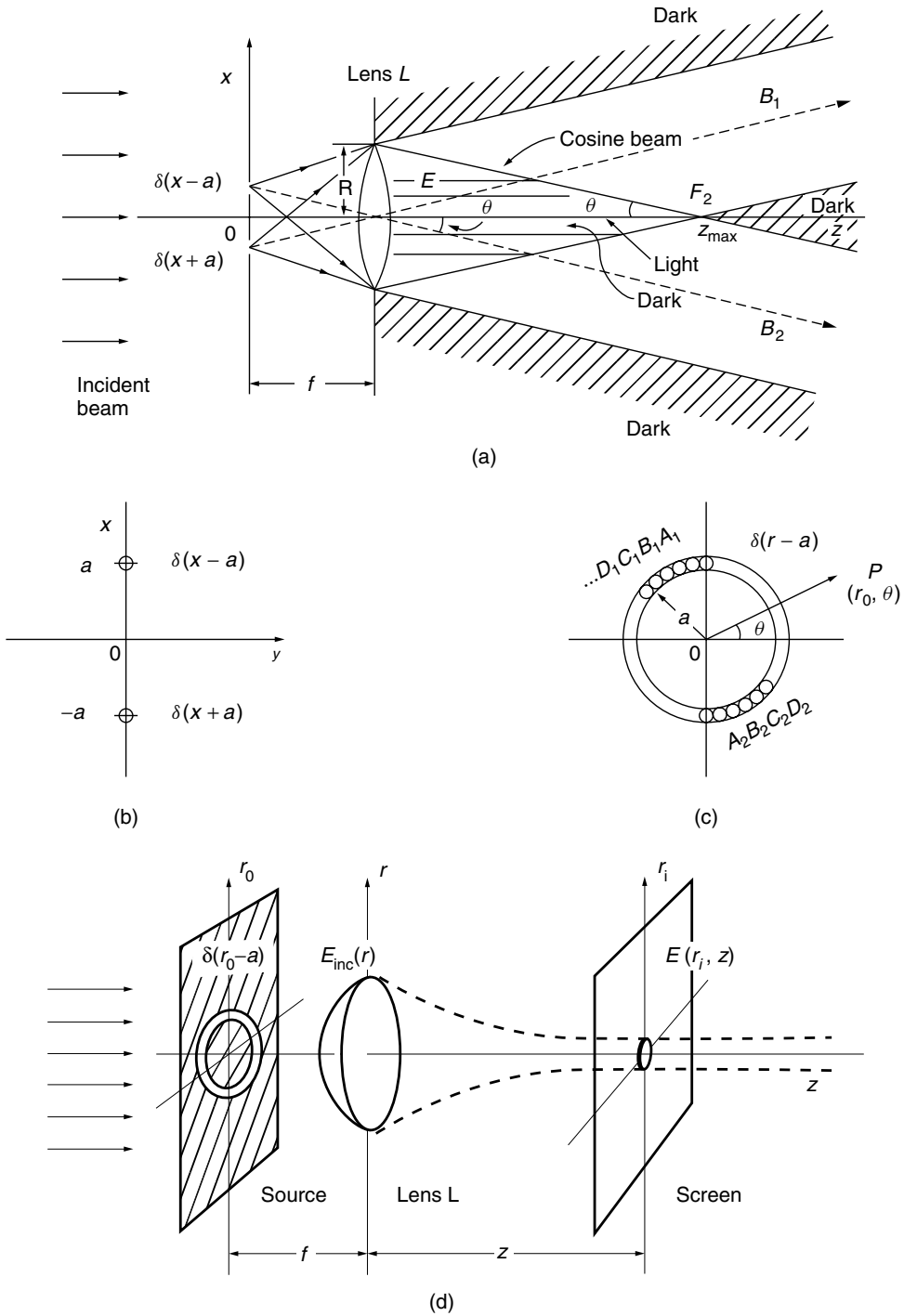


Figure 3.41 Geometries of cosine and Bessel beams. (a) Generation of a cosine beam. (b) Two delta function sources. (c) Bessel beam source $\delta(r-a)$. (d) Generation of the Bessel beam.

As far as the field distribution in the x - z plane is concerned, the resultant field creates a cosinusoidal standing wave pattern in the x direction and a traveling wave in the z direction.

Let this wave be called the cosine beam. The spacing Λ between the standing wave peaks is

$$\Lambda = \frac{\lambda}{2 \sin \theta} \quad (3.172)$$

This one-dimensional model (cosine beam) displays almost all features of the Bessel beam mentioned above.

1. The wave propagates while maintaining the same cosine transverse pattern $2A \cos \alpha x$.
2. The beam abruptly ends at

$$z_{\max} = R \tan^{-1} \theta = R \frac{f}{a} \quad (3.173)$$

where the crossover of the two beams ends. Since the radius R of the lens is finite, z_{\max} is always finite.

3. The propagation constant along the beam (z axis) is less than that of free space and can be adjusted by means of θ .
4. By placing a π -radian phase shifter in front of one of the delta function sources, the cosine beam can be converted into a sine beam.

Next, in order to generate the Bessel beam, the two delta function sources are replaced by a delta function ring source, $\delta(r - a)$. For now, the ring source is divided into paired sections, $A_1, A_2; B_1, B_2; C_1, C_2; D_1, D_2$; and so on, as shown in Fig. 3.41c. The points $A_2, B_2, C_2, D_2, \dots$ are located diagonally opposite to $A_1, B_1, C_1, D_1, \dots$, respectively.

Sections A_1 and A_2 generate exactly the same pattern as the cosine beam mentioned above. Sections B_1 and B_2 do the same but in a plane tilted from the x - z plane. The contributions of all the sections of the ring source are superimposed. Superposition, however, has to be performed taking the phase into consideration.

The distances to all subsections are the same as long as the point of observation is on the z axis, and in this case, the contributions of the subsections are all added in phase. As soon as the point of observation moves away from the z axis, phase mismatch occurs among the contributions of the subsections of the ring source. Thus, the field intensity reaches its maximum on the z axis and decays in a rippling fashion with distance away from the z axis.

Approximately, the field distribution in the transverse plane is a Bessel function and the beam is expressed as

$$E(z, r) = A e^{j(\beta z - \omega t)} J_0(\alpha r) \quad (3.174)$$

where α is given by Eq. (3.171).

3.11.4 Mathematical Expressions for the Bessel Beam

Mathematical expressions for the Bessel beam will be developed using the geometry with the delta function ring source in Fig. 3.41d [27].

First, an expression is sought for the input field $E_{\text{inc}}(r)$ to the convex lens. This expression will be used later for calculating the Bessel beam field distribution. $E_{\text{inc}}(r)$ is calculated using the Fresnel approximation of the Fresnel–Kirchhoff diffraction formula, Eq. (1.38). The Fourier transform Eq. (1.66) in cylindrical coordinates is used.

$$E_{\text{inc}}(r) = \frac{1}{j\lambda f} e^{jk(f+r^2/2f)} 2\pi \int_0^\infty \delta(r_0 - a) e^{jk(r_0^2/2f)} J_0(2\pi\rho r_0) r_0 dr_0 \Big|_{\rho=r/f\lambda} \quad (3.175)$$

where r_0 and r are radial coordinates in the planes of the ring source and the lens L , respectively.

The term $e^{jk(r_0^2/2f)}$ is a part of the point spread function in cylindrical coordinates and corresponds to Eq. (1.40) of rectangular coordinates. The result of the integration from Eq. (1.102) with $dx dy = r dr d\theta$ is

$$E_{\text{inc}}(r) = A e^{jk(r^2/2f)} J_0\left(k\frac{a}{f}r\right) \quad (3.176)$$

where

$$A = \frac{k}{jf} e^{jk(f+a^2/2f)} \quad (3.177)$$

Thus, $E_{\text{inc}}(r)$ on the front surface of the convex lens is the zero order Bessel function of the first kind combined with a quadratic phase factor $e^{jk(r^2/2f)}$. This quadratic phase factor is removed by passing through the convex lens whose transmittance is $e^{-jk(r^2/2f)}$ as given by Eq. (1.139). The field on the output surface from the convex lens is $AJ_0(k(a/f)r)$.

The Fresnel approximation of the field $E(r_i, z)$ emergent from the convex lens is obtained again by the Fresnel–Kirchhoff integral as

$$E(r_i, z) = A' \int_0^R J_0\left(k\frac{a}{f}r\right) e^{jk(r^2/2z)} J_0(2\pi\rho r) r dr \Big|_{\rho=r_i/\lambda z} \quad (3.178)$$

$$A' = A \frac{k}{jz} e^{jk(z+r_i^2/2z)} \quad (3.179)$$

where r_i is the radial coordinate of the point of observation at a distance z away from the lens L .

The expression for the on-axis field distribution ($r_i = \rho = 0$) along the z axis is much simpler than Eq. (3.178).

By putting

$$\begin{aligned} r_i &= \rho = 0 \\ q &= \frac{k}{j2z} \\ \alpha &= k\frac{a}{f} \\ x &= \sqrt{q}r \end{aligned} \quad (3.180)$$

Eq. (3.178) is simplified to

$$E(0, z) = \frac{A'}{2q} \int_0^{\sqrt{q}R} 2x J_0\left(\frac{\alpha}{\sqrt{q}}x\right) e^{-x^2} dx \quad (3.181)$$

The Bessel integral formula is available [28] for Eq. (3.181).

$$\int_0^a 2x J_0(\gamma x) e^{-x^2} dx = \begin{cases} e^{-\gamma^2/4} - \sum_{n=0}^{\infty} \left(\frac{-\gamma}{2a}\right)^n J_n(\gamma a) e^{-a^2} & \text{for } \left|\frac{2a}{\gamma}\right| < 1 \\ \sum_{n=0}^{\infty} \left(\frac{2a}{\gamma}\right)^n J_n(\gamma a) e^{-a^2} & \text{for } \left|\frac{2a}{\gamma}\right| > 1 \end{cases} \quad (3.182)$$

By comparing Eq. (3.181) with Eq. (3.182) and by putting

$$\begin{aligned} a &= \sqrt{q}R \\ \gamma &= \frac{\alpha}{\sqrt{q}} \end{aligned} \quad (3.183)$$

the expression for $E(0, z)$ becomes

$$E(0, z) = \frac{A'}{2q} \begin{cases} e^{-\alpha^2/4q} - \sum_{n=0}^{\infty} \left(\frac{-\alpha}{2qR}\right)^n J_n(\alpha R) e^{-qR^2} & \text{for } \left|\frac{\alpha}{2qR}\right| > 1 \\ \sum_{n=0}^{\infty} \left(\frac{2qR}{\alpha}\right)^n J_n(\alpha R) e^{-qR^2} & \text{for } \left|\frac{\alpha}{2qR}\right| < 1 \end{cases} \quad (3.184)$$

Various factors appearing in Eq. (3.184) will be rewritten.

From Eqs. (3.179) and (3.180) and $r_i = 0$, the factor $A'/2q$ becomes

$$\frac{A'}{2q} = A e^{jkz} \quad (3.185)$$

From Eq. (3.180), the factor $\alpha/2qR$ is rewritten as

$$\frac{\alpha}{2qR} = j \frac{za}{Rf}$$

From Eq. (3.173), this factor becomes

$$\frac{\alpha}{2qR} = \frac{jz}{z_{\max}} \quad (3.186)$$

Thus, inserting Eqs. (3.185) and (3.186) into Eq. (3.184) gives

$$E(0, z) = Ae^{jkz} \begin{cases} e^{-j(\alpha^2/2k)z} - \sum_{n=0}^{\infty} \left(\frac{jz}{z_{\max}} \right)^n J_n(\alpha R) e^{jk(R^2/2z)} & \text{for } z > z_{\max} \\ \sum_{n=0}^{\infty} \left(\frac{z_{\max}}{jz} \right)^n J_n(\alpha R) e^{jk(R^2/2z)} & \text{for } z < z_{\max} \end{cases} \quad (3.187)$$

Calculation of Eq. (3.187) produces curves such as the one shown in Fig. 3.40a.

The calculation of the integral of Eq. (3.178) for the off-axis field distribution is more involved and can be found in Ref. 27. Calculated results with various combinations of physical parameters are shown in the same reference.

3.11.5 Methods of Generating Bessel Beams

In order to explain the principle governing the generation of the Bessel beam, a delta function ring source was used in Fig. 3.41 but such a source is inefficient and impractical. Figure 3.42 summarizes more practical methods. The method shown in Fig. 3.42a employs a conical lens [22] whose thickness is linearly reduced as the rim of the lens is approached.

The conical lens generates two parallel beams crossing each other. The pattern shown in Fig. 3.42a corresponds to the field pattern in the x - y plane, but the field patterns in any plane that includes the optical axis are the same as shown in the figure. The conical lens generates the Bessel beam. The region of focus is a line from point F_1 to F_2 on the optical axis. This type of conical lens forms a line image rather than point image and it is sometimes called an axicon, meaning the axis image [29].

Despite the simplicity and high power output, the arrangement with an axicon alone has the disadvantage that the region of the Bessel beam starts immediately behind the axicon. Usually, a certain distance is required from the source to the beginning of the measuring device, as, for example, when the Bessel beam is used for triangulation [23].

The convex projector lens L in Figs. 3.42a relocates the region of the Bessel beam. The region is transformed from the input region F_1 - F_2 to the output region F'_1 - F'_2 . The role of the projector lens is to form the output image F'_n from the input point F_n .

In Fig. 3.42b, positive and negative axicons are combined to transfer the Bessel beam further away from the axicon. The input axicon is a negative axicon and diverges the beam, while the second axicon is a positive axicon and converges the beam. As shown in Fig. 3.42b, there is a hatched dark region in the center of the output axicon so that the two beams meet at a smaller angle and transfer the Bessel beam further away from the output axicon. This combination of a positive with a negative axicon is called a teleaxicon [23].

In Fig. 3.42c, the conical lenses in Fig. 3.42b are replaced by ordinary spherical lenses [21]. The hatched dark region in the center is created by an opaque disk. The resulting beam pattern is quite similar to that from Fig. 3.42b.

In order for the Bessel beam to be created, the region has to be illuminated by nothing but the tilted beams B_1 and B_2 . If the opaque disk is not installed in the center, a third beam propagating parallel to the z axis will be present, in addition to B_1 and B_2 , and the Bessel beam will be disturbed.

A holographic approach to generating a Bessel beam is illustrated in Fig. 3.42d [30]. The Bessel beam is generated by simply illuminating a computer-generated hologram by a parallel laser beam.

Instead of computer-generated holograms, an actual hologram fabricated directly from a Bessel beam [31] can be used. Figure 3.42e shows an arrangement for the fabrication of such a hologram. The laser beam is split into object and reference waves by means of a half-mirror HM. Each wave is spatially filtered and expanded by a combination of two lenses and a pinhole: L_1, P_1, L_2 for the object wave and L_3, P_2, L_4 for the reference wave.

The object wave is converted into a Bessel beam by means of the delta function annular slit placed in the front focal plane of lens L_5 . The photographic plate H is placed in the region F_1-F_2 of the Bessel beam. The off-axis reference wave is added to the photographic plate. After the exposure, the plate is developed. Bleaching after development increases the diffraction efficiency of the hologram.

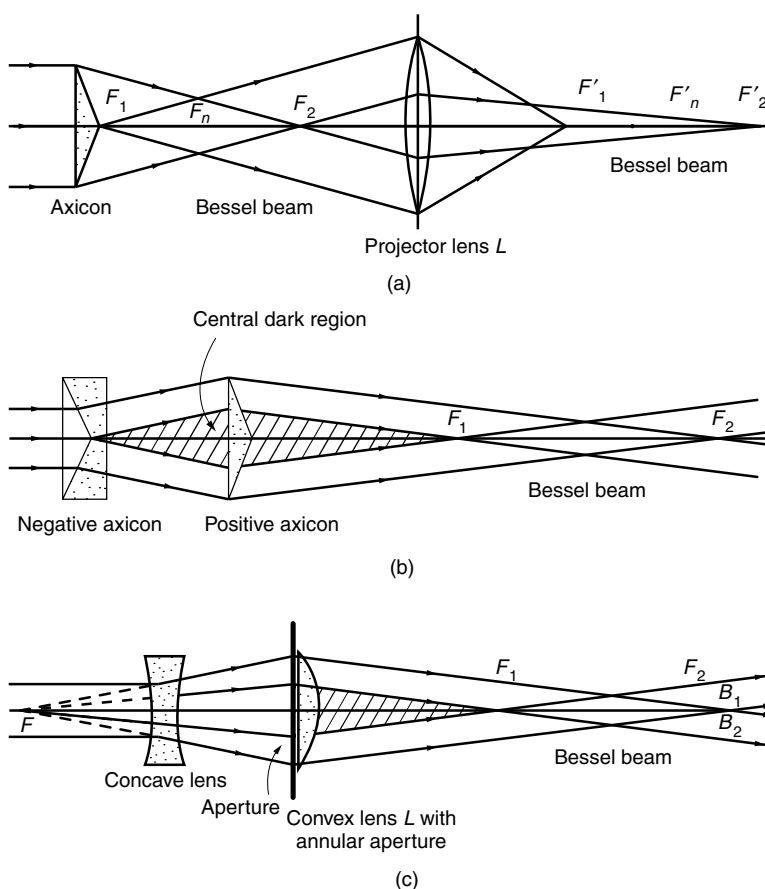


Figure 3.42 Various methods for generating Bessel beams. (a) By means of an axicon and a projector lens. (b) By means of a teleaxicon. (c) By means of a spherical convex lens with an annular iris. (d) By means of a holographic plate. (e) Fabrication of a Bessel beam hologram. (f) By means of β -beam reflector telescope. (After T. Aruga et al. [32].)

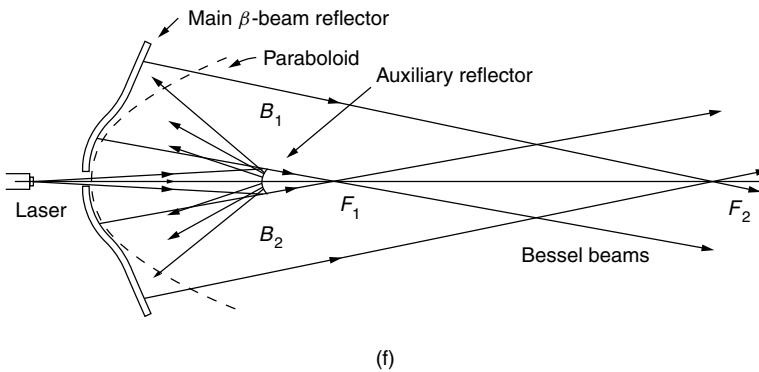
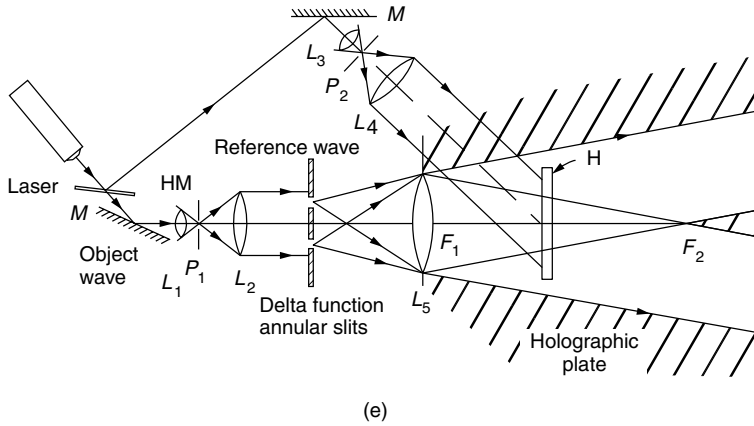
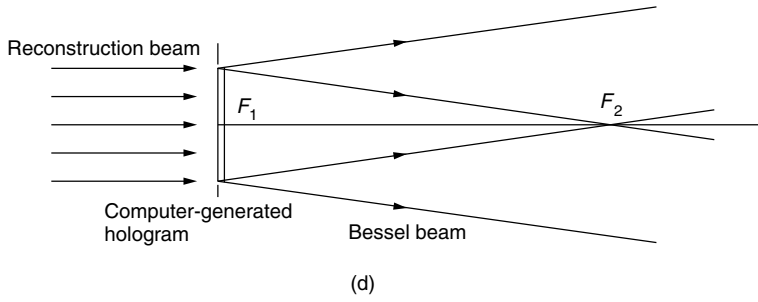


Figure 3.42 (Continued)

The Bessel beam is immediately generated by illuminating this hologram with the reconstruction beam that is the reference beam used for fabrication. The advantage of the holographic Bessel beam is that neither an annular slit nor a convex lens is needed to reconstruct the Bessel beam.

Figure 3.42f shows a reflector-type arrangement [32] for generating a Bessel beam. The surface of the main or auxiliary reflector of the telescope is deformed from the paraboloid so that tilted beams B_1 and B_2 are generated. The reflector type is advantageous over the lens or holographic type when a large-diameter device is needed for higher sensitivity.

Example 3.12 For the Bessel beam geometry shown in Fig. 3.43, derive expressions for (a) the range F_1-F_2 and (b) the diameter of the main lobe using Eq. (3.174).

Solution Figure 3.43 shows a cross section of the x - z plane of the axicon. In order to find the propagation direction θ of beam B_1 , the phase ϕ of a plane wave starting from point S to point P is calculated using wave optics as

$$\phi = knx \tan \Delta + k(z - x \tan \Delta) \quad (3.188)$$

where n is the refractive index of the axicon. The expression for the constant phase ϕ_0 is obtained by setting $\phi = \phi_0$. The vector \mathbf{s} of the propagation constant is the gradient of the constant phase line $\nabla\phi_0$, and

$$\mathbf{s} = \hat{\mathbf{i}}[k(n-1)\tan\Delta] + \hat{\mathbf{k}}k \quad (3.189)$$

where $\hat{\mathbf{i}}$ and $\hat{\mathbf{k}}$ are unit vectors in the x and z directions, respectively. Thus, the direction of propagation θ of beam B_1 is

$$\tan\theta = (n-1)\tan\Delta \quad (3.190)$$

(a) From the geometry, the focal range $\overline{F_1F_2}$ is

$$\overline{F_1F_2} \doteq \frac{R}{(n-1)\tan\Delta} \quad (3.191)$$

where R is the radius of the axicon.

(b) α in Eq. (3.180) is the x component of the propagation constant and is, from Eq. (3.189),

$$\alpha = k(N-1)\tan\Delta \quad (3.192)$$

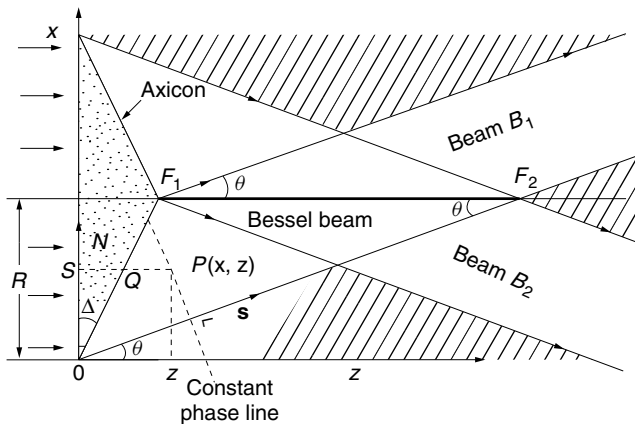


Figure 3.43 Geometry for generating a Bessel beam by means of an axicon lens.

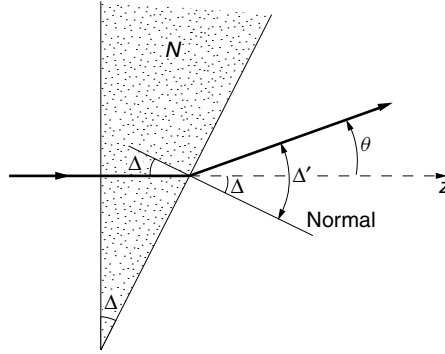


Figure 3.44 Expanded view of the axicon.

$J_0(\alpha r)$ has its first zero at $\alpha r = 2.4$ and the diameter d of the main lobe is

$$d = \frac{2.4\lambda}{\pi(N-1)\tan\Delta} \quad (3.193)$$

It should be pointed out that the same answer is obtained using geometrical optics. An enlarged section is shown in Fig. 3.44. Snell's law gives

$$\begin{aligned} n \sin \Delta &= \sin \Delta' \\ \theta &= \Delta' - \Delta \end{aligned}$$

When $\Delta \ll 1$, θ is approximated as

$$\theta \doteq (n-1)\Delta \quad (3.194)$$

which matches with the approximation of Eq. (3.190) when $\Delta \ll 1$. \square

3.12 MANIPULATION WITH LIGHT BEAMS

In 1619 Johannes Kepler (1571–1630) introduced the concept of radiation pressure to explain why comets' tails are always trailing away from the sun. More recently, the radiation pressure [33,34] of a laser beam has been utilized for trapping micron-sized dielectric spheres. These trapping devices are more commonly known as “*optical tweezers*.” Radiation pressure also has been used for quieting down the random motion of atoms and molecules, a technique known as laser cooling. Even though both optical tweezers and laser cooling use the force generated by the changes in the momentum of the laser beam, the mechanisms of the interaction between the object and the laser beam are different.

3.12.1 Radiation Pressure of Laser Light

Each photon has quantum energy of

$$E = h\nu \quad (3.195)$$

and momentum of

$$p = \frac{h}{\lambda} \quad (3.196)$$

where h is the Planck's constant and is $h = 6.6 \times 10^{-34}$ J·s. Thus, the total number of photons in light with energy W is

$$N = \frac{W}{h\nu} \quad (3.197)$$

The momentum $P_m = Np$ is, from Eqs. (3.196) and (3.197),

$$P_m = \frac{W}{h\nu} \cdot \frac{h}{\lambda} = \frac{n_1 W}{c} \quad (3.198)$$

where n_1 is the index of refraction of the medium.

The force is the time derivative of the momentum, and the force in a medium of index of refraction n_1 is

$$F = n_1 P / c \quad (3.199)$$

where P is the incident light power in watts. The magnitude of the force is not large, but for a particle of a small mass like an atom or molecule, it is significant. The acceleration of a 1- μ m-diameter sphere by the radiation pressure force [33,34] calculated from Eq. (3.199) is over 10^5 times larger than that of gravitation [35].

Next, the force generated by a laser beam transmitted through a sphere will be calculated, referring to Fig. 3.45.

It is assumed that not only the surface reflection is negligible, but also the sphere is perfectly transparent and no heat effect is involved. The refractive index of the dielectric sphere is assumed larger than that of the surrounding medium.

Let vector \mathbf{a} represent the momentum of the incident light and \mathbf{a}' , that of the emergent light. The momentum \mathbf{c}_a transferred to the sphere is obtained from the law of conservation of momentum:

$$\mathbf{a} = \mathbf{a}' + \mathbf{c}_a \quad (3.200)$$

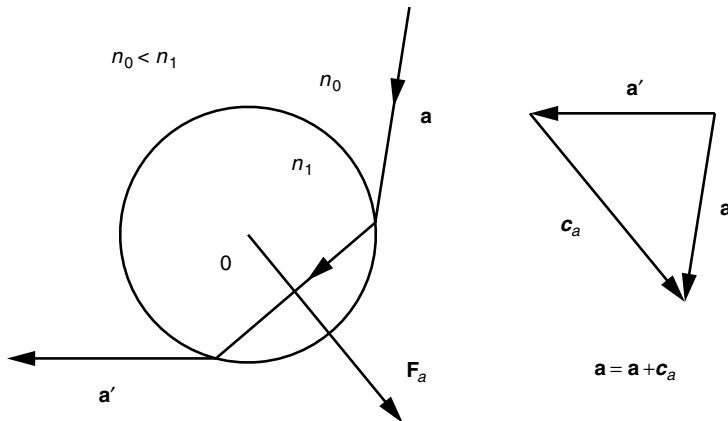


Figure 3.45 The direction of the force on the sphere is known from the conservation of momentum of the laser beam.

The difference c_a between \mathbf{a} and \mathbf{a}' provides the direction of the force \mathbf{F}_a acting on the sphere by the transmitted laser beam. \mathbf{F}_a is parallel to c_a .

3.12.2 Optical Tweezers

A TEM₀₀ mode Gaussian beam is incident onto a large NA value convex lens. The dielectric sphere to be manipulated is placed near the focus f of the lens. Figure 3.46a shows the case when the sphere is displaced above the focus of the lens. (The lens

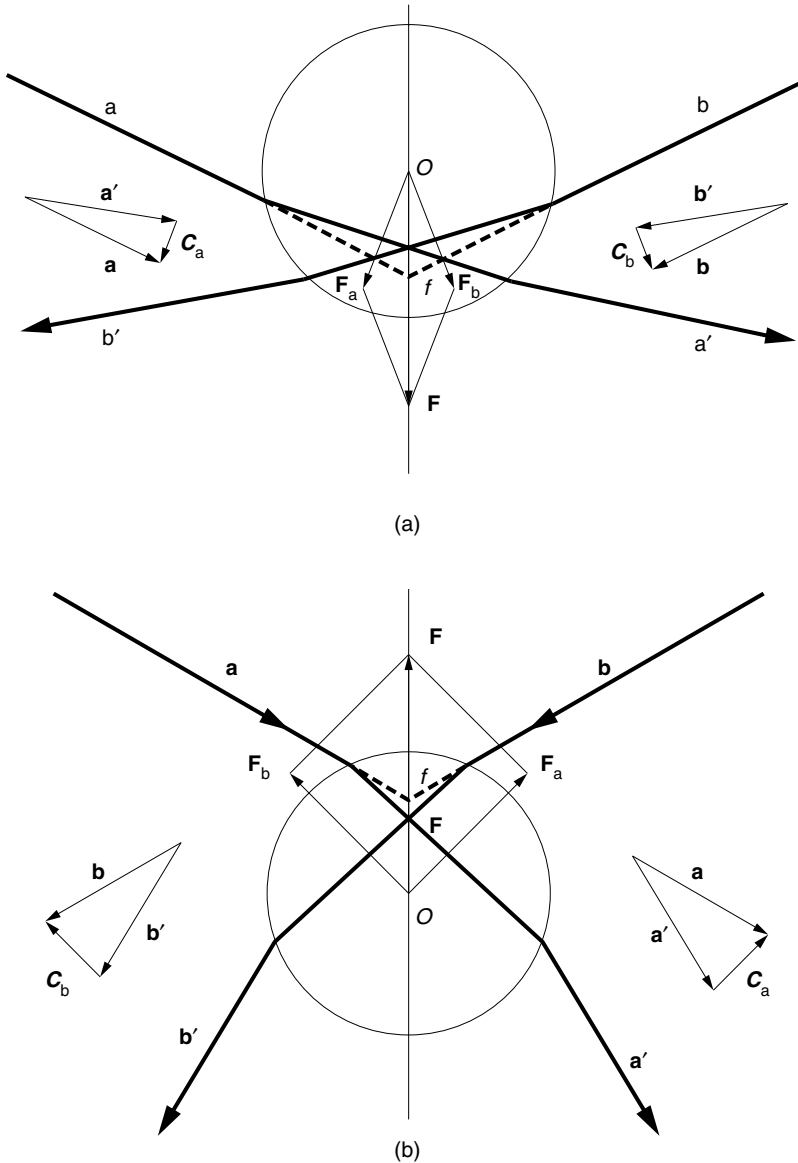


Figure 3.46 Restoring forces from the axial displacements of a dielectric sphere: $n_0 = 1.0$, $n_1 = 1.5$. (a) Displacement above the trap focus f . (b) Displacement below the trap focus f . (After A. Ashkin [35].)

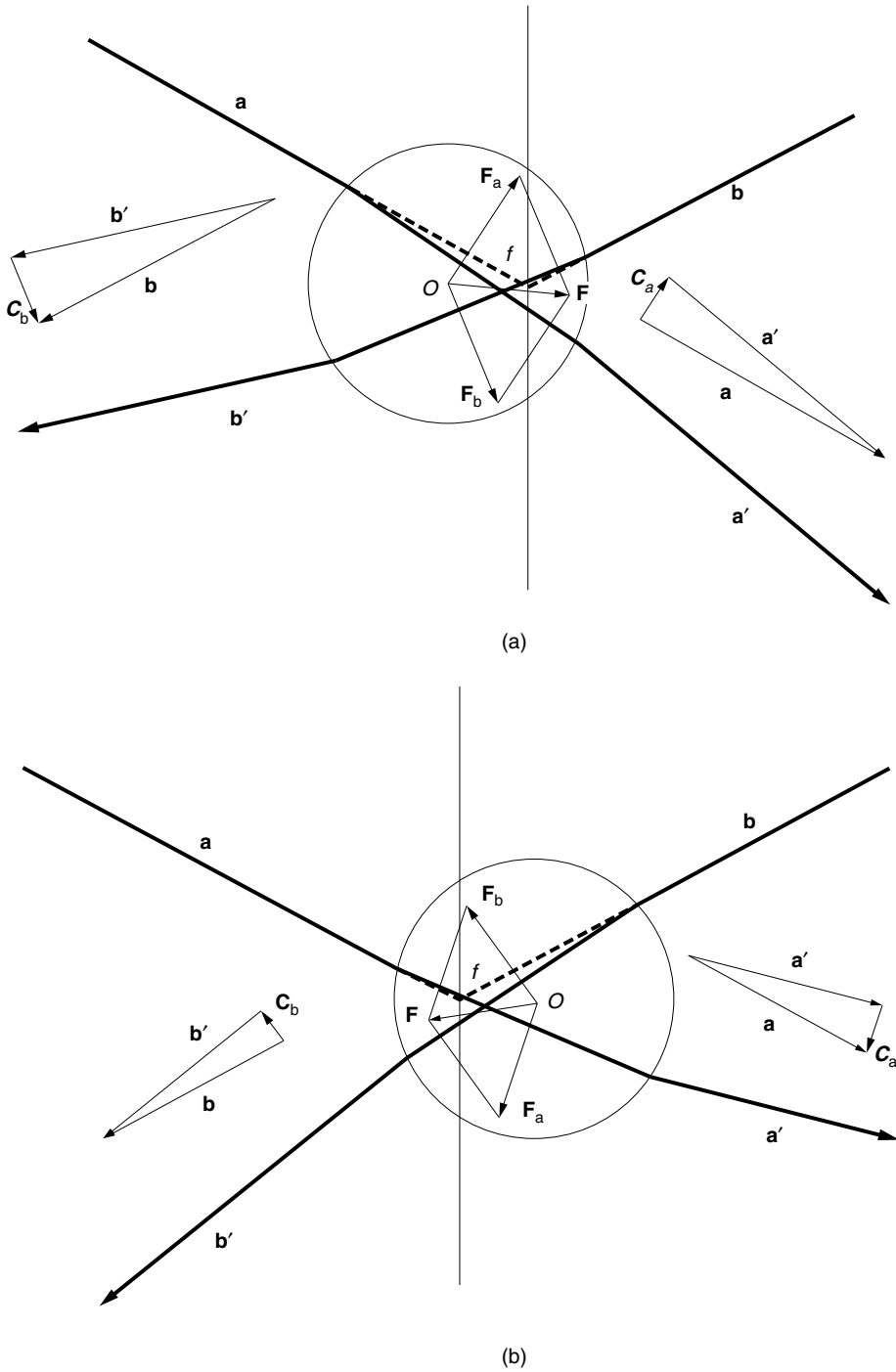


Figure 3.47 Restoring forces from the transverse displacement of a dielectric sphere: (a) Displacement to the left of the trap focus f . (b) Displacement to the right of the trap focus f . (After A. Ashkin [35].)

is above the sphere and is not shown.) Lines a and b are representative component beams from the converging lens. From the difference between the vectors \mathbf{a} and \mathbf{a}' , \mathbf{F}_a of beam a is found, and similarly from vectors \mathbf{b} and \mathbf{b}' , \mathbf{F}_b of beam b is found. The direction of the resultant vector

$$\mathbf{F} = \mathbf{F}_a + \mathbf{F}_b \quad (3.201)$$

is the direction of the force. The direction of \mathbf{F} is downward or, in other words, in the direction of restoring the axial displacement.

In the same way, Fig. 3.46b shows the case when the sphere is displaced below the focus f . The resultant force F points up and restores the displacement.

The same is repeated in Figs. 3.47a and 3.47b to show the restoring force for the transversally displaced sphere. Figures 3.46 and 3.47 confirm stable operation as an *optical tweezer*.

Next, the TEM_{00} and TEM_{11} Gaussian modes are compared to see which makes a better optical tweezer. The beam components with a larger apex angle generate a larger trapping force.

The TEM_{11} Gaussian mode beam is more efficient than the TEM_{00} Gaussian mode for trapping the sphere, because as shown in Fig. 3.31 the TEM_{11} mode has a null intensity in the center and the light intensity is more concentrated at the larger apex angles where a larger trapping force is generated.

The optical tweezers provide a high degree of control over the dynamics of small particles and play an important role in physical and biological sciences, especially in manipulating living cells. Figure 3.48 shows the sequence of a DNA molecule being pulled through a polymer solution with the use of an optical tweezer. The DNA molecule is seen to relax along a path defined by its contour [36].

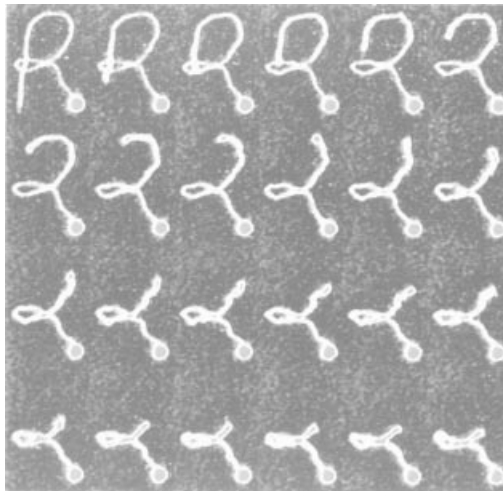


Figure 3.48 DNA molecule being pulled by optical tweezers. The first image in the upper left shows a stained DNA molecule pulled into the shape of the letter R with the optical tweezers. The remaining images (read left to right, top to bottom) show the relaxation of the DNA molecule along its contour. (Courtesy of Steven Chu [36].)

3.13 LASER COOLING OF ATOMS

The random movement of an atom can be slowed down by means of the radiation pressure of a laser beam tuned to an atomic resonance. Such a technique is known as *laser cooling* [37,38]. Because of the minuteness of the target, the laser beam generally passes through the atoms without being disturbed, and no momentum of the laser beam is transferred to the atom. The situation, however, becomes quite different when the frequency of the illuminating laser beam is tuned to the resonance frequency ν_0 of an atomic transition. The transition energy is expressed in terms of the resonance frequency as

$$h\nu_0 = E_1 - E_0 \quad (3.202)$$

where E_0 is the energy level of the ground state of the atom and E_1 is that of the next higher energy level. At the resonant frequency ν_0 , the scattering cross section markedly increases.* The focused laser beam is almost entirely absorbed by the atoms in its path and the laser beam momentum is efficiently transferred to the atoms.

Each time an atom in the ground state absorbs a photon, it makes a transition to the excited state and at the same time it is pushed by the radiation pressure of the photon. The change in the velocity of the motion of the atom due to one push is

$$v_c = \frac{h}{\lambda m} \quad (3.203)$$

where m is the mass of the pushed atom and h/λ is the momentum of the photon.

Once the atom makes a transition to the excited state, the atom no longer interacts with the laser beam until the atom returns to the ground state. The atom returns to the ground state by reradiating a photon with energy $h\nu_0$ by spontaneous emission. The direction of the reradiation is statistically random and the net contribution of the recoils is zero. The atom that returned to the ground state is ready to absorb another photon, and the process continues.

For a typical atom, the spontaneous emission lifetime is 10^{-8} second, which means the repetition of the absorption–emission process takes place 10^8 times a second.

Another important aspect of the mechanism of laser cooling is the differential absorption of the photon energy associated with the Doppler shift. The laser frequency ν_l is set lower than the resonance frequency ν_0 of the atom with an offset frequency v/λ , where the v is the velocity of the atom. When the atom is moving toward the laser source, the atom sees a frequency higher than the laser frequency due to the Doppler shift, and with this offset frequency, the atom sees exactly the resonance frequency. The interaction is high. On the other hand, when the atom is moving away from the laser source, the atom sees a frequency lower than the laser frequency due to the Doppler shift. This atom sees a frequency further away from the resonance frequency, and the interaction is low. The differential absorption generates a net force resisting the movement of the atom.

As the atom is cooled down the velocity reduces and the frequency of the laser source has to be raised accordingly (frequency chirp) in order to maintain optimum interaction

* At the resonance frequency, the scattering cross section reaches a value of approximately λ^2 . Off resonance, the value of the scattering cross section is almost zero.

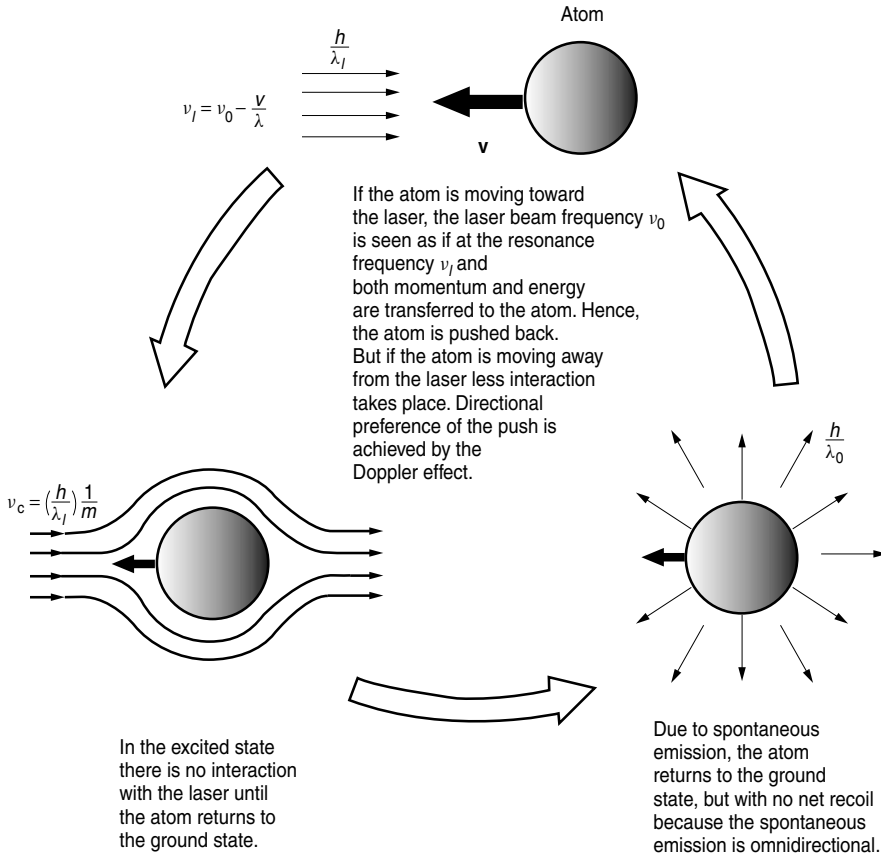


Figure 3.49 Laser cooling cycle.

with the laser beam. This method is called Doppler cooling in order to distinguish it from another method of cooling called polarization gradient cooling [36]. For the atoms whose velocities are reduced by this process, laser cooling quiets down the motion of the atoms to almost complete rest. Figure 3.49 summarizes the laser cooling cycle.

It should be pointed out that the quantum lost from an atom due to spontaneous emission is $h\nu_0$ while that of the laser source is $h\nu_l$. Since ν_0 is larger than ν_l , the lost energy is larger than the supplied energy to the atom. The difference accounts for the cooling of the atom.

Laser cooling has made a significant impact on high resolution spectroscopy and frequency standards like atomic clocks [39].

PROBLEMS

- 3.1** What happens to the pattern on the screen in Fig. 3.10, when the lens L and prism are removed? Assume that the incident light is monochromatic. Draw the patterns in the x - z plane when the screen is very near, moderately far, and very far from the Fabry-Pérot resonator.

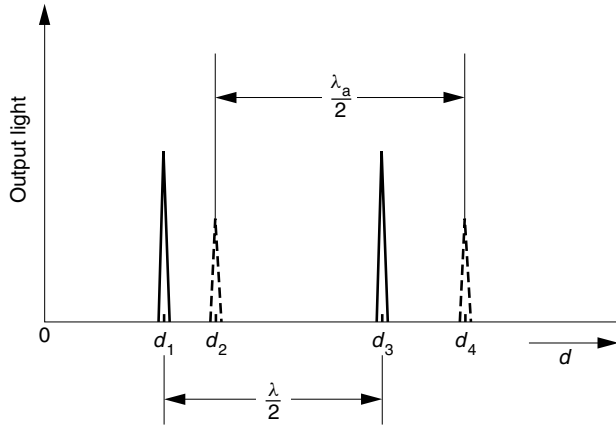


Figure P3.2 Measured peaks using a scanning Fabry-perot resonator.

- 3.2** With the method described in Section 3.4.3, when d is scanned, the same peak pattern repeats at every half-wavelength of the incident light, as shown in Fig. P3.2. Theoretically speaking, λ and λ_a can be determined individually, by measuring

$$\lambda = 2(d_3 - d_1)$$

$$\lambda_a = 2(d_4 - d_2)$$

What is the disadvantage of such a method as far as $\Delta\lambda/\lambda$ is concerned?

- 3.3** The modulation frequency of a He-Ne laser beam was measured by the method explained in Section 3.4.3. The He-Ne wavelength is $\lambda = 0.6328 \mu\text{m}$. Determine the frequency of modulation from the following measured quantities:

$$\Delta d_1 = 0.0829 \mu\text{m}$$

$$\Delta d_2 = 0.1441 \mu\text{m}$$

$$d_1 - d_2 = 300 \mu\text{m}$$

- 3.4** Figure P3.4 shows the display of a scanning Fabry-Pérot resonator when phase modulated light is incident onto the resonator. The mirror spacing is $420 \mu\text{m}$.

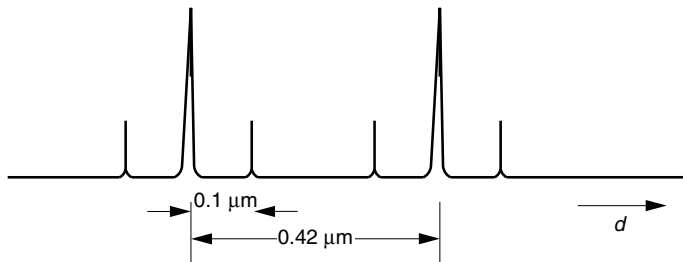


Figure P3.4 The display of the scanning Fabry-perot resonator used to find λ .

- (a) What is the carrier wavelength λ ?
- (b) What is the mode number m of the resonance?
- (c) What is the modulation frequency $\Delta\nu$?
- 3.5** The fringe rings of a Fabry–Pérot spectroscopy are shown in Fig. P3.5. Using the key dimensions in the figure, find the wavelength λ of the incident wave.
- 3.6** It is believed that the gravitational wave causes ground strain when it arrives from extragalactic sources. A 300-m long Fabry–Pérot cavity such as shown in Fig. P3.6 was built [40] to detect the ground strain. What is the resolution

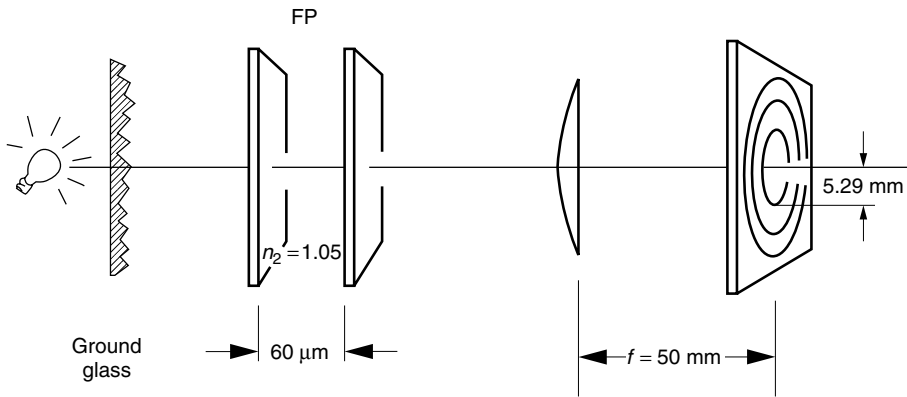


Figure P3.5 Fringe rings of the Fabry–perot spectroscopy used to find λ .



Figure P3.6 The 300-m Fabry–perot cavity of the TAMA gravitational wave detector. (Courtesy of A. Araya et al. [40].)

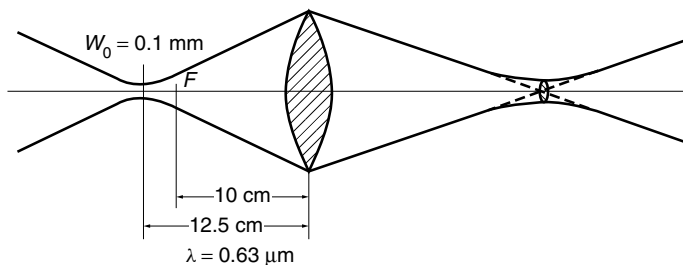


Figure P3.7 Geometry of a Gaussian beam.

$\Delta L/L$ of such a detection system, where L is the length of the cavity and ΔL is the strain for which the output intensity of the cavity drops to one-half of the resonance value. The wavelength of the laser is $\lambda = 1064$ nm. The length of the cavity is 300 m. The finesse of the cavity is $F = 516$.

- 3.7** Using the graphs in Fig. 3.27, obtain M , d_1 , W_1 , z_1 , and θ_1 for the case shown in Fig. P3.7.
- 3.8** A free-space optical communication link is to be established between two satellites. The distance between the satellites is 3.6×10^3 km
- (a) A beam waist of 1 m on the receiver satellite is desired. A 10-m focal length lens is installed on the transmitter satellite, and the light wavelength is $0.63 \mu\text{m}$, as shown in Fig. P3.8a. What are the size and location of the waist of the launching beam?

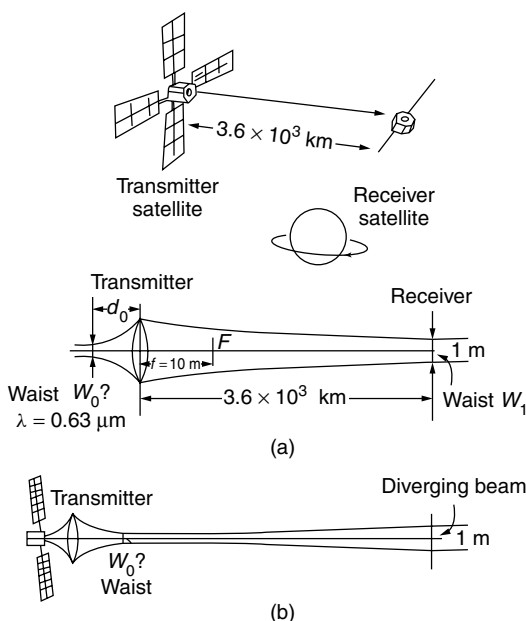


Figure P3.8 Inter satellite free-space optical communication. (a) Waist of the beam is on the receiver side. (b) Waist of the beam is on the transmitter side.

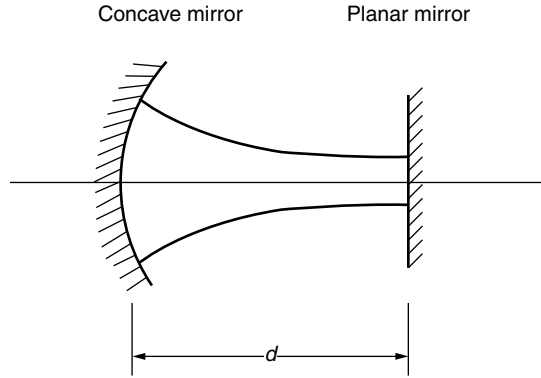


Figure P3.9 Combination of a concave mirror and a planar mirror.

- (b) Assume this time that the beam waist is located at the surface of the transmitting satellite, as shown in Fig. P3.8b. The Gaussian beam expands to a radius of 1 m at the surface of the receiving satellite. Estimate the size of the beam waist launched from the transmitting satellite.
- 3.9** A cavity is constructed with a concave mirror with radius of curvature R_c and a plane mirror spaced by a distance d , as shown in Fig. P3.9. Find the radii of the beam on the end mirrors.
- 3.10** A combination of an annular slit and a convex lens, as shown in Fig. 3.41d, is used to generate a Bessel beam with the following characteristics. The size of the main lobe, which is defined in terms of the dimension ρ from the center to the first zero of the main lobe, is

$$\rho = 60 \text{ } \mu\text{m}$$

The maximum beam distance is

$$z_{\text{max}} = 1 \text{ m}$$

The radius of the annular slit is 2.5 mm. The wavelength is $0.63 \text{ } \mu\text{m}$. Find the focal length and radius of the convex lens.

- 3.11** In 1873 Sir William Crooks (1832–1919) proposed that a radiometer, such as the one shown in Fig. P3.11, could measure radiation pressure. Which way should it rotate if it is indeed measuring the radiation pressure?
- 3.12** An experiment is conducted to demonstrate radiation pressure by showing how much a reflective sphere is pushed by a laser beam (see Fig. P3.12) [35]. The radius of the sphere is equal to the wavelength of the laser beam and its density is 1 g/cm^3 .

An argon ion laser with an output light power of 1 watt, and with $\lambda = 0.5145 \text{ } \mu\text{m}$ is used. Assume that only 7% of the available force is used for pushing the reflective sphere.

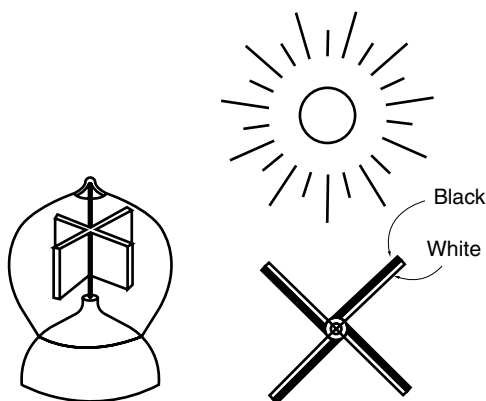


Figure P3.11 Is it a radiometer or a radiation pressure gauge?

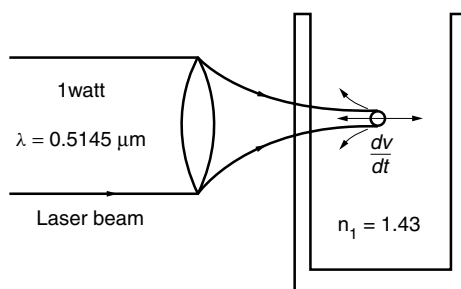


Figure P3.12 Radiation pressure experiment.

- (a) How much is the radiation pressure force acting on the sphere?
 - (b) What is the acceleration that this sphere acquires due to the pushing force of the laser beam?
 - (c) Compare the acceleration by radiation pressure with that by gravity.
- 3.13** A convex lens is used to focus a parallel light beam. What is the direction of the radiation pressure acting on the convex lens?

REFERENCES

1. G. Hernandez, *Fabry-Pérot Interferometers*, Cambridge University Press, Cambridge, 1986.
2. J. M. Vaughan, *The Fabry-Pérot Interferometer, History, Theory, Practice and Applications*, Adam Hilger, Bristol, 1989.
3. A. Yariv, *Optical Electronics*, Saunders, Philadelphia, 1991.
4. M. Mansuripur, "Fabry-Pérot etalons in polarized light," *Opt. and Photonics News* 39–44 (Mar. 1997).
5. C. Y. Chen, M. M. Choy, M. J. Andrejco, M. A. Saifi, and C. Lin, "A widely tunable erbium-doped fiber laser pumped at 532 nm," *IEEE Photonics Technol. Lett.* **2**(1), 18–20 (1990).

6. J. R. Andrews, "Low voltage wavelength tuning of an external cavity diode laser using a nematic liquid crystal-containing birefringent filter," *IEEE Photonics Technol. Lett.* **2**(5), 334–336 (1990).
7. J. S. Patel and Y. Silverberg, "Liquid crystal and grating-based multiple wavelength cross-connect switch," *IEEE Photonics Technol. Lett.* **7**(5), 514–516 (1995).
8. J. S. Patel and M. W. Maeda, "Tunable polarization diversity liquid-crystal wavelength filter," *IEEE Photonics Technol. Lett.* **3**(8), 739–740 (1991).
9. P. J. Collings and J. S. Pantel, *Handbook of Liquid Crystal Research*, Oxford University Press, 1997.
10. J. S. Patel and S. D. Lee, "Electronically tunable and polarization insensitive Fabry–Pérot etalon with a liquid-crystal film," *Appl. Phys. Lett.* **58**(22), 2491–2493 (1991).
11. J. S. Patel, "Polarization insensitive tunable liquid-crystal etalon filter," *Appl. Phys. Lett.* **59**(11), 1314–1316 (1991).
12. S. Yamazaki, M. Shibutani, N. Shimosaka, S. Murata, T. Ono, M. Kitamura, K. Emura, and M. Shikada, "A coherent optical FDM CATV distribution system," *J. Light Technol.* **8**(3), 396–405 (1990).
13. C. M. Miller and F. J. Janniello, "Passively temperature-compensated fibre Fabry–Pérot filter and its application in wavelength division multiple access computer network," *Electron. Lett.* **26**(25), 2122–2123 (1990).
14. K. Matsudaira, *Fundamentals and Experiments on Lasers* (in Japanese), Kyoritsu Publishing, Tokyo, 1973.
15. C. K. L. Wah, K. Iizuka, and A. P. Freundorfer, "94 GHz phase modulation of light by organic MNA (2 methyl 4 nitroaniline) crystal," *Appl. Phys. Lett.* **63**(23), 3110–3112 (1993).
16. H. Kogelnik and T. Li, "Laser beams and resonators," *Appl. Opt.* **5**(10), 1550–1567 (1966).
17. B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, Wiley, New York, 1991.
18. T. F. Johnston, Jr., "Beam propagation (M^2) measurement made as easy as it gets: the four-cuts method," *Appl. Opt.* **37**(21), 4840–4850 (1998).
19. D. Marcuse, *Light Transmission Optics*, 2nd ed., Van Nostrand Reinhold, New York, 1982.
20. J. Durnin, "Exact solutions for nondiffracting beams. I. The scalar theory," *J. Opt. Soc. Am. A* **4**(4), 651–654 (1987).
21. R. M. Herman and T. A. Wiggins, "Production and uses of diffractionless beams," *J. Opt. Soc. Am. A* **8**(6), 932–942 (1991).
22. J. Durnin, J. J. Miceli, Jr., and J. H. Eberly, "Diffraction-free beams," *Phys. Rev. Lett.* **58**(15), 1499–1501 (1987).
23. G. Bickel, G. Häusler, and M. Maul, "Triangulation with expanded range of depth," *Opt. Eng.* **24**(6), 975–977 (1985).
24. S. Sogomonian, S. Klewitz, and S. Herminghaus, "Self-reconstruction of a Bessel beam in a nonlinear medium," *Opt. Commun.* **139**, 313–319 (1997).
25. L. Niggel, T. Lanzl, and M. Maier, "Properties of Bessel beams generated by periodic gratings of circular symmetry," *J. Opt. Soc. Am. A* **14**(1), 27–33 (1997).
26. T. Wulle and S. Herminghaus, "Nonlinear optics of Bessel beams," *Phys. Rev. Lett.* **70**(10), 1401–1404 (1993).
27. P. L. Overfelt and C. S. Kenney, "Comparison of the propagation characteristics of Bessel, Bessel–Gauss and Gaussian beams diffracted by a circular aperture," *J. Opt. Soc. Am. A* **8**(5), 732–745 (1991).
28. A. S. Chai and H. J. Wertz, "The digital computation of the far-field radiation pattern of a truncated Gaussian aperture distribution," *IEEE Trans. Antennas Propag.* **AP13**, 994–995 (1965).

29. J. H. McLeod, "The axicon: a new type of optical element," *J. Opt. Soc. Am.* **44**(8), 592–597 (1954).
30. A. Vasara, J. Turunen, and A. T. Friberg, "Realization of general nondiffracting beams with computer-generated holograms," *J. Opt. Soc. Am. A* **6**(11), 1748–1754 (1989).
31. A. J. Cox and D. C. Dibble, "Holographic reproduction of a diffraction-free beam," *Appl. Opt.* **30**(11), 1330–1332 (1991).
32. T. Aruga, S. W. Li, S. Yoshikado, M. Takabe, and R. Li, "Nondiffracting narrow light beam with small atmospheric turbulence-influenced propagation," *Appl. Opt.* **38**(15), 3152–3156 (1999).
33. A. Ashkin, "Forces of a single-beam gradient laser trap on a dielectric sphere in the ray optics regime," *Biophys. J.* **61**, 569–582 (1992).
34. A. Ashkin, "Optical trapping and manipulation of neutral particles using lasers," *Proc. Natl. Acad. Sci. USA* **94**, 4853–4860 (1997).
35. A. Ashkin, "Acceleration and trapping of particles by radiation pressure," *Phys. Rev. Lett.*, **24**(4), 156–159 (1970).
36. S. Chu, "The manipulation of neutral particles," *Rev. Mod. Phys.* **70**(3), 685–706 (1998).
37. M. Watanabe, "The Nobel Prize in Physics 1997 for development of methods to cool and trap atoms with laser light," *Rev. Laser Eng.* **26**(2), 195–198 (1998).
38. W. D. Phillips, "Laser cooling and trapping of neutral atoms," *Rev. Mod. Phys.* **70**(3), 721–741 (1998).
39. C. N. Cohen-Tannoudji, "Manipulating atoms with photons," *Rev. Mod. Phys.* **70**(3), 707–719 (1998).
40. A. Araya, S. Telada, K. Tochikubo, S. Taniguchi, R. Takahashi, K. Kawabe, D. Tatsumi, T. Yamazaki, S. Kawamura, S. Miyoki, S. Moriwaki, M. Musha, S. Nagano, M. Fujimoto, K. Horikoshi, N. Mio, Y. Naito, A. Takamori, and K. Yamamoto, "Absolute length determination of a long-baseline Fabry-Pérot cavity by means of resonating modulation sidebands," *Appl. Opt.* **38**(13), 2848–2856 (1999).