Optics and Optical Cavities

Andrew Orr-Ewing (2003)

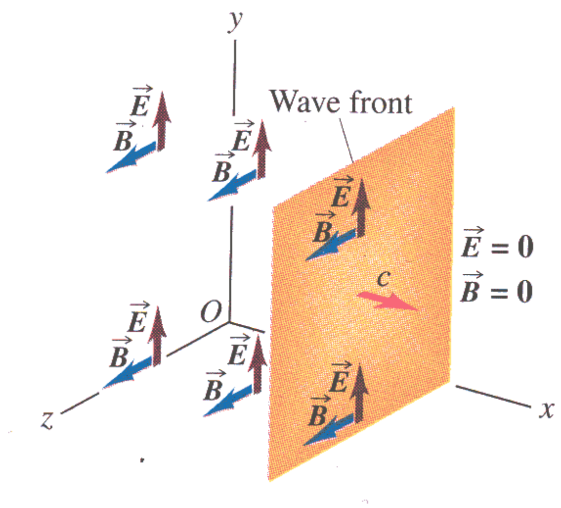
References:
Lasers, A.E. Siegman
Laser Spectroscopy, W. Demtroder
Building Scientific Apparatus, J.H. Moore, C.C. Davis and M.A. Coplan
Quantum Electronics, A. Yariv

1. Electromagnetic (EM) radiation

Light consists of electromagnetic waves that have oscillating electric (**E**) and magnetic (**B**) fields. These waves carry both energy and momentum. The **E** and **B** fields are sinusoidal functions of time and position with a definite frequency and wavelength. Maxwell's equations demonstrate that a time varying magnetic field acts as a source of electric field, and a time-varying electric field acts as a source of magnetic field (e.g., a moving charge generates a **B**-field). Thus, when either an electric or a magnetic field is changing with time, a field of the other kind is induced in adjacent regions of space. These electromagnetic fields propagate in free space at the speed of light.

1.1 Plane electromagnetic waves and the speed of light.

The **E** and **B** fields of the EM radiation are mutually perpendicular and we will take a coordinate system with **E** polarized along the +*y*-direction, **B** along the +*z*-direction and the wave propagating along the +*x*-axis. The wave front is the boundary between the region containing the EM fields and a region of zero field, and moves along +*x* with constant speed *c*. A wave such as this, for which at any instant the fields are uniform over any plane perpendicular to the direction of propagation is called a plane wave. The wave is **transverse** i.e., the **E** and **B** fields are perpendicular to the plane of propagation.



The speed of light is given by

$$c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} = 299792458 \text{ m s}^{-1}$$

where $\varepsilon_0 = 8.85419 \times 10^{-12}$ J^{-1} C^2 m^{-1} is the permittivity of free space and $\mu_0 = 4\pi \times 10^{-7}$ J s^2 C^{-2} m^{-1} is the permeability of free space.

The direction of energy propagation is the **Poynting vector**

$$\underline{S} = \frac{1}{\mu_0} \underline{E} \times \underline{B}$$

and propagation requires no medium. The magnitudes of **E** and **B** are related by E = *c*B and the two fields oscillate in phase. Any wave travelling in the *x*-direction can be represented as a superposition of waves linearly polarized in the *y*- and *z*-directions.

The instantaneous values of the *y*-component of **E** and the *z*-component of **B** at some point along *x* are:

$$E(x, t) = E_0 \sin(\omega t - kx)$$
$$B(x, t) = B_0 \sin(\omega t - kx)$$

where E_0 and B_0 are the amplitudes of the fields, $\omega = 2\pi\nu$ is the angular frequency and $k = 2\pi/\lambda$ is the wave number. The wave is seen to propagate in time because a wave of the form:

$$V = V_0 \cos(kx + \phi)$$

is a standing wave with constant phase (and a wavelength of $\lambda = 2\pi/k$). The waves in the above formulae have a phase that varies with time ($\phi(t) = \omega t$) and thus shifts the wave along the *x*-axis. The phase shifts through a complete cycle of 2π in a time τ such that $\omega\tau = 2\pi$ so $\tau = 1/\nu$. A phase shift of 2π corresponds to advancement of a distance of one wavelength, λ in a time τ. The speed of the wave is thus $v = \lambda/\tau = \nu\lambda$ as required.

The energy density (energy per unit volume) associated with the EM wave in vacuum is

$$u = \varepsilon_0 E^2$$

with equal contributions from the **E** and **B** fields. The energy flow per unit time per unit area is

$$S = c\,\varepsilon_0 E^2 = E\,B/\mu_0$$

with units of W m^{-2}. The Poynting vector describes the magnitude and direction of the energy flow rate. The average value of *S*, which oscillates at the frequency of the EM wave, is the intensity of the radiation (also with units of W m^{-2})

$$I = \langle S \rangle = \tfrac{1}{2} \varepsilon_0 c E_0^2$$

1.2 Electromagnetic waves in matter

We consider EM wave propagation in non-conducting (dielectric) materials. The wave speed is reduced from that in vacuum and is denoted here by *v* rather than *c*. The permittivity and permeability of the medium are given by:

$$\varepsilon = \varepsilon_r \varepsilon_0$$

$$\mu = \mu_r \mu_0$$

where ε_r and μ_r are the dielectric constant and the relative permeability and are dimensionless numbers. The speed of the EM wave is

$$v = \frac{1}{\sqrt{\varepsilon\mu}} = \frac{c}{\sqrt{\varepsilon_r \mu_r}}$$

and because $\mu_r \approx 1$ for most dielectrics, the speed of the EM wave in a dielectric is less than the speed of light. The ratio of the speed in vacuum to the speed in a material is known as the **index of refraction, n**:

$$\frac{c}{v} = n = \sqrt{\varepsilon_r \mu_r} \approx \sqrt{\varepsilon_r}$$

The wavelength in a medium is altered from that in vacuum according to:

$\lambda = \lambda_{vac} / n$

EM waves cannot propagate any appreciable distance in a conductor because the **E** and **B** fields lead to currents that provide a mechanism to dissipate and reflect the energy of the wave. For an ideal conductor, **E** is zero everywhere inside the material and an incident EM wave is totally reflected.

When an EM wave strikes the surface of a conducting reflector, the incident wave induces oscillating currents that give rise to an opposing **E**-field. The net **E**-field is zero everywhere on the surface of the conductor. The currents generate a reflected wave and the superposition of incident and reflected waves gives a standing wave (with fixed nodal planes perpendicular to the *x*-axis) and the **E** and **B** fields oscillate in time 90^o out of phase:

$$E(x,t) = -2E_0 \sin kx \cos \omega t$$
$$B(x,t) = 2B_0 \cos kx \sin \omega t$$

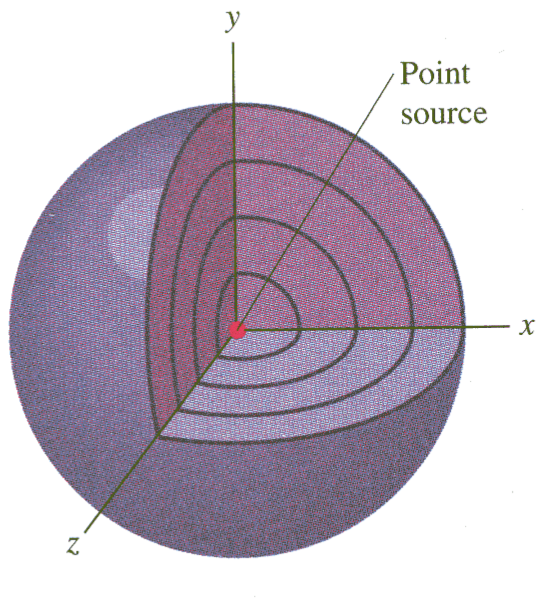
Note that these are standing waves (fixed phase along *x*) but with time-dependent amplitudes given by cos ωt or sin ωt.

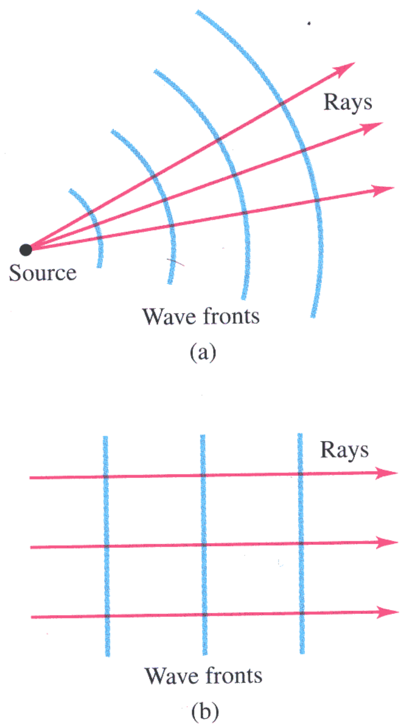
Reflections also occur at an interface between two insulating materials with different dielectric or magnetic properties. The wave is partially transmitted into the second material and partly reflected back into the first.

2. Optics

2.1 Propagation of light

We define the wave front as the *locus of all adjacent points at which the phase of vibration of a physical quantity associated with the wave is the same.* That is, at any instant, all points on a wave front are at the same part of their oscillation (e.g. peaks or troughs). When EM radiation expands out from a point source, any spherical surface concentric with the source is a wave front. Far away from the source, parts of the surface of a sphere look like planes and we can consider plane wave behaviour.





To describe the propagation of light, it is convenient to represent the light wave by rays rather than wave fronts. A ray is an imaginary line along the direction of travel of the wave. In a homogeneous, isotropic material, the rays are always straight lines perpendicular to the wave fronts. At a boundary surface between two materials, the wave speed and the direction of a ray may change. The branch of optics for which the ray description is adequate is termed **Geometric Optics**.

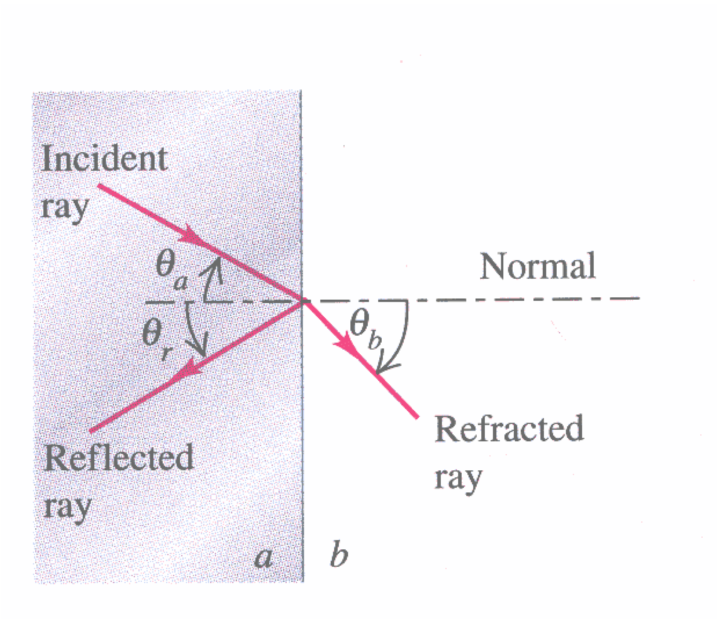
The segments of plane waves can be represented by bundles of rays forming beams of light, and for simplicity we often only draw one ray in a beam.

2.2 Laws of reflection and refraction:

Reflection at a definite angle from a smooth surface is called **specular reflection**.

1. The incident, reflected and refracted rays and the normal to the surface all lie in the same plane (perpendicular to the boundary surface between two materials) known as the **plane of incidence**.

2. The angle of reflection, θ_r is equal to the angle of incidence θ_a for all wavelengths and for any pair of materials.



3. For monochromatic light and for a given pair of materials *a* and *b* on opposite sides of the interface, **Snell's Law** applies:

$$\frac{\sin\theta_a}{\sin\theta_b} = \frac{n_b}{n_a}$$

Thus, for refraction from vacuum into a medium such as quartz with greater index of refraction, $\theta_{quartz} < \theta_{vac}$ and the ray is bent towards the normal to the surface. In passage from quartz into vacuum, the ray bends away from the surface normal.

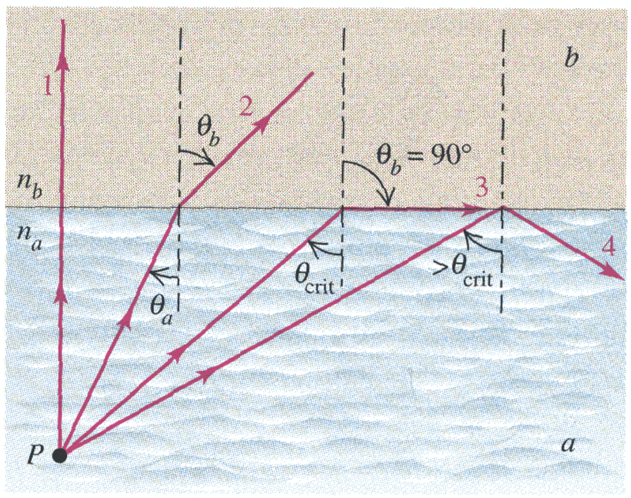
Some indices of refraction of common materials (at 589 nm):

Air:	1.0003
Quartz:	1.544
Diamond:	2.417
Water:	1.333

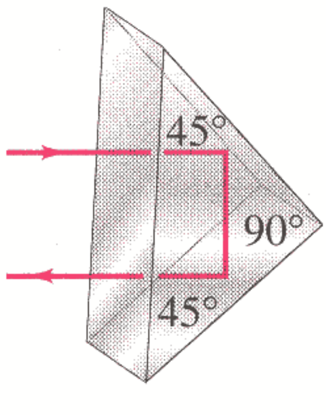
The fraction of intensity reflected or refracted depends upon the polarization of the light, the indices of refraction and the angle of incidence.

When light propagates from a medium *a* of higher refractive index to a material *b* of lower refractive index, the rays bend away from the surface normal. Thus there must be some value of θ_a less than 90^o at which $\theta_b = 90^o$ and the ray emerges into medium *b* parallel to, and grazing the interface. This value of θ_a is called the **critical angle (θ_c)**. If the angle of incidence is greater than the critical angle, the ray cannot pass into material *b* ($\sin\theta_b$ cannot be greater than 1) and thus is trapped in material *a* giving **total internal reflection**. This can only occur for $n_a > n_b$, and is the basis of operation of fibre optics.

$$\sin\theta_c = \frac{n_b}{n_a}$$



(a)

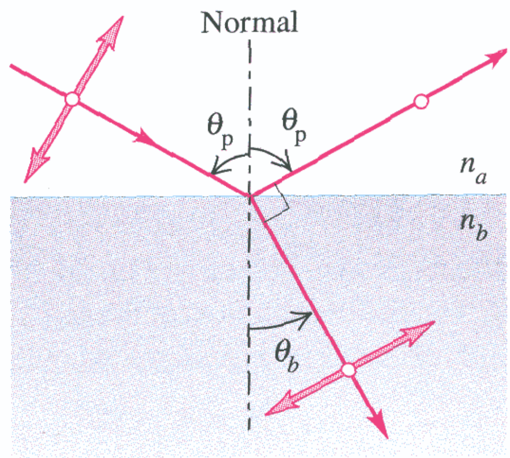


At the glass-air interface, for example, with n = 1.52 for the glass, $\theta_c = 41.1^o$, and a prism with angles of 45^o, 90^o and 45^o can be used as a totally reflecting surface.

The speed of light of different wavelengths in a medium other than vacuum can depend on the wavelength – a property called **dispersion**. The refractive index will also be wavelength dependent. One consequence is separation of different wavelengths of light by a prism because the different indices of refraction result in different deviations of the light from the incident direction as the light enters and exits the prism.

Light reflected from a surface can be polarized. For most angles of incidence, waves for which the electric field vector **E** is perpendicular to the plane of incidence (the plane of the figure on the left) and thus parallel to the reflecting surface, are reflected more strongly than those for which **E** lies in the plane of incidence. At one incidence angle θ_p, called the **polarizing angle**, the light for which E is lies in the plane of incidence is not reflected at all, so the reflected light is completely polarized perpendicular to the plane of incidence. **Brewster's law** gives the angle as

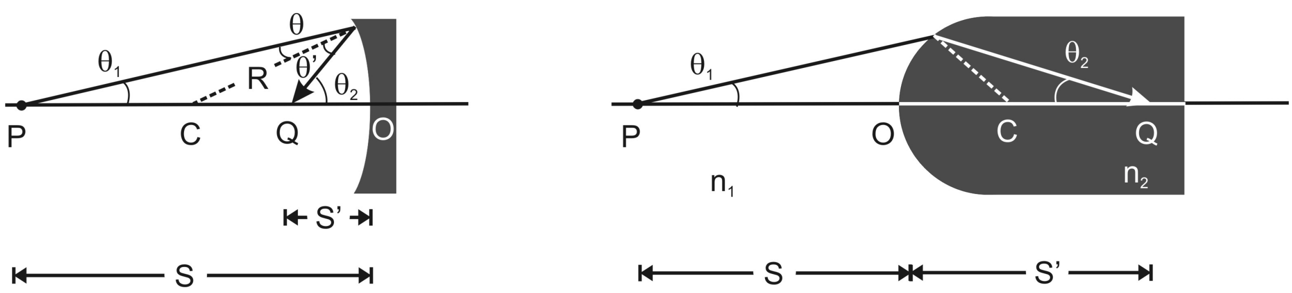
$$\tan\theta_p = \frac{n_b}{n_a}$$



2.3 Paraxial ray analysis

We will denote the propagation direction of a plane wave by the wave vector **k**. In real optical systems, such (infinitely spread) plane waves do not exist because of the finite size of the optical elements. Non-planar optical components cause further deviations of the wave from planarity so the wave acquires a ray direction that varies from point to point on the wave front. In a cylindrically symmetric optical system, **paraxial rays** are those rays whose directions of propagation occur at small enough angles from the cylindrical symmetry axis that $\sin\theta$ (or $\tan\theta$) can be replaced by θ, i.e., we use a small angle approximation.

Virtually all optical instruments in common use contain only two types of optical surface, namely plane and spherical. Ray tracing, using the laws of reflection and refraction, can be used to analyse optical systems composed of such optical elements. The figure below shows the path of a ray originating at a point *P* that lies on the axis of symmetry of a single optical surface – either a spherical mirror or a spherical refracting surface separating two optical media. The ray returns to the axis at a point *Q.* The distance *OP* = *s* is called the **object distance** and the distance *OQ* = *s'* is called the image distance. The radius of curvature of the surface is *OC* = *R*.



For the spherical mirror,

$$R\sin\theta = CP\,\sin\theta_1 = (s - R)\,\sin\theta_1$$
$$R\,\sin\theta' = QC\,\sin\theta_2 = (R - s')\,\sin\theta_2$$

but for reflections, $\theta = \theta'$ and thus

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{(R - s')}{(s - R)}$$

for the relationship between the slope angles of the incident and reflected rays. Similarly for the spherical refracting surface, it can be shown that:

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{(s' - R)}{(s + R)}\frac{n_2}{n_1}$$

for the relationship between the incident and refracted rays.

When a bundle of rays originates from an axial point, the analysis above shows that the image distances are not the same for all rays but rather are functions of the original slope angles θ_1 at the object point. This means that the rays do not come to a single focus – a common phenomenon known as **spherical aberration**. If the angles are small enough for the sines to be replaced by the angles themselves, the important simplification known as the **paraxial approximation** applies.

For the spherical reflector,

$$\frac{1}{s}+\frac{1}{s'}=\frac{2}{R}$$

and for the spherical refracting surface

$$\frac{n_1}{s}+\frac{n_2}{s'}=\frac{n_2-n_1}{R}.$$

If the object distance is infinitely large ($s\to\infty$), so that the incoming rays are parallel, the image distance is $s' = R/2$ which is the **focal length of the mirror**:

$$f = R/2$$

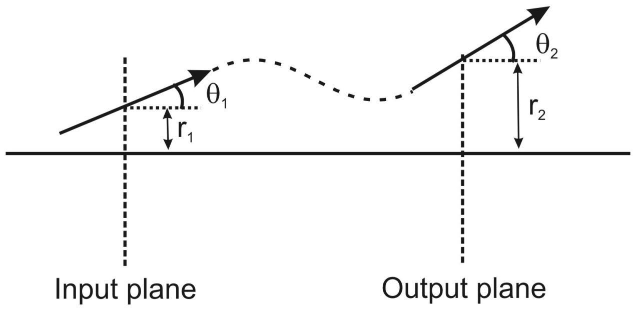
and thus for a spherical mirror of radius of curvature R,

$$\frac{1}{s}+\frac{1}{s'}=\frac{1}{f}$$

2.4 Matrix formulation

In an optical system with a symmetry axis in the z-direction, a paraxial ray at distance z is characterized by its distance **r** from the z-axis and the angle θ it makes with that axis. Suppose the values of these parameters at two planes of the system (an *input* and an *output* plane) are (r_1, θ_1) and (r_2, θ_2). In the paraxial ray approximation there is a linear relation between them of the form:

$$r_2 = Ar_1 + B\theta_1$$
$$\theta_2 = Cr_1 + D\theta_1$$



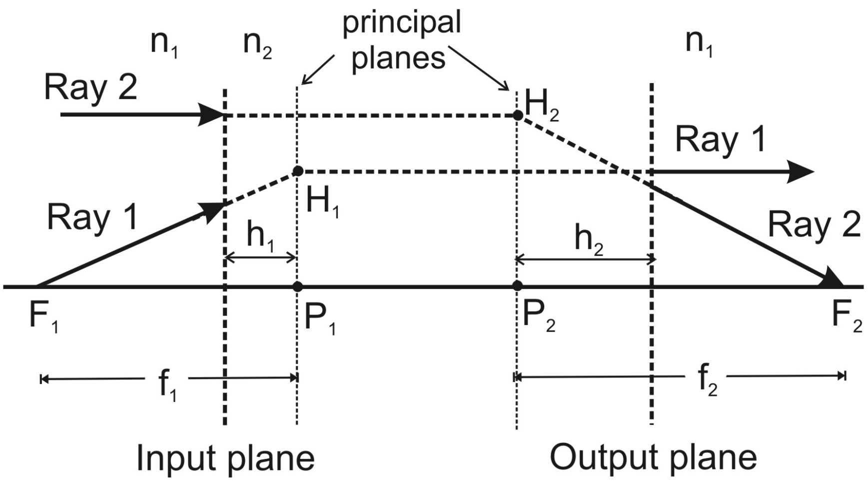
or, in matrix notation,

$$\begin{pmatrix} r_2 \\ \theta_2 \end{pmatrix}=\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} r_1 \\ \theta_1 \end{pmatrix}.$$

The matrix $\mathbf{M}=\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is called the **ray transfer matrix**. Its determinant AD-BC =1. Optical systems made of isotropic materials are generally reversible – a ray that travels from right to left with input parameters (r_2, θ_2) will leave the system with parameters (r_1, θ_1). Thus

$$\begin{pmatrix} r_1 \\ \theta_1 \end{pmatrix}=\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1}\begin{pmatrix} r_2 \\ \theta_2 \end{pmatrix}$$

The figure below illustrates the concepts of input and output planes, focal points and principal planes of an optical system. An input ray that passes through the first focal point F_1 emerges travelling parallel to the axis. The intersection of the extrapolated input and output rays, point H_1 locates the **first principal plane**. Conversely an input ray travelling parallel to the axis will emerge at the output plane and pass through the second focal point F_2. The intersection of the extrapolated rays defines the point H_2 which locates the **second principal plane**. Rays 1 and 2 are called the principal rays. The dashed lines are called virtual ray paths. The axis of the system intersects the principal planes at the principal points P_1 and P_2. The distances from the principal planes to the focal points are denoted as f_1 and f_2, the first and second focal lengths.



If the refractive indices on the input and output sides are the same, several simplifications arise:

$$f_1 = f_2 = f = -\frac{1}{C}$$

$$h_1 = \frac{D-1}{C}$$

$$h_2 = \frac{A-1}{C}$$

Thus these parameters can easily be determined from knowledge of the transfer matrix.

Sign conventions:

- The ray is assumed to propagate from left to right along the +*z*-axis.
- The distance from the first principal plane to the object is measured as positive from right to left.
- The distance from the second principal plane to the an image is measured positive from left to right.
- The lateral distance of the ray from the axis is positive in the upward direction.
- The acute angle between the system axis direction and the ray is positive for an anti-clockwise motion.
- The radius of curvature of an interface is positive if the interface is convex to the input ray and negative if it is concave to the input ray.

2.4.A Examples of ray transfer matrices for simple optical systems

(i) uniform optical medium:

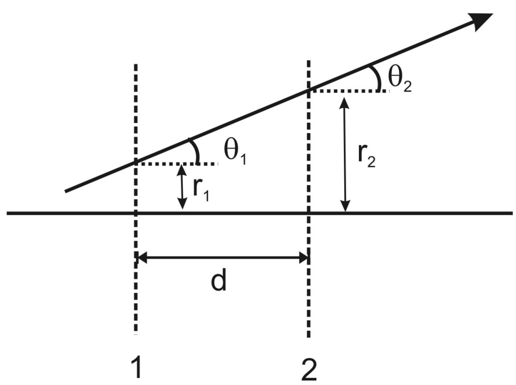
In a medium of length *d*, no change in ray angle occurs so

$\theta_2 = \theta_1$

$r_2 = r_1 + d\theta_1$ (because $d \tan\theta_1 \approx d\,\theta_1$).

so

$$\mathbf{M} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}$$



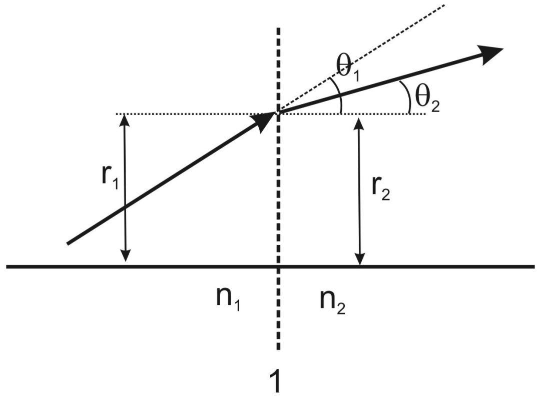
The focal length of the system is infinite and it has no specific principal planes.

(ii) Planar interface between two different media:
At the interface, $r_1 = r_2$ and Snell's law in the paraxial approximation gives

$$\theta_2 = \frac{n_1}{n_2}\theta_1$$

so

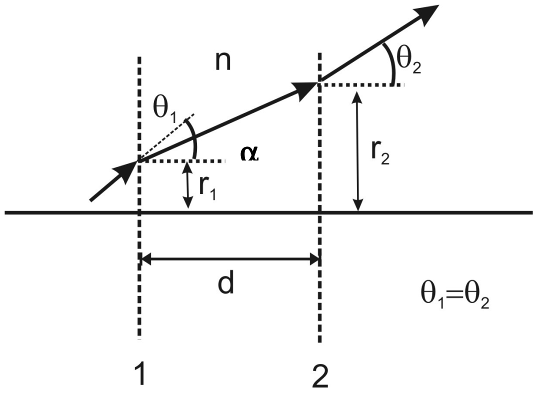
$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ 0 & n_1/n_2 \end{pmatrix}$$



(iii) A parallel-sided slab of refractive index n bounded on both sides with media of index 1:

From Snell's law, $\alpha=\theta_1/n$ so $r_2 = r_1 + d\,\theta_1/n$. Since the refractive indices on both sides of the slab are the same, θ_1 and θ_2 are equal. Hence

$$\mathbf{M} = \begin{pmatrix} 1 & d/n \\ 0 & 1 \end{pmatrix}$$



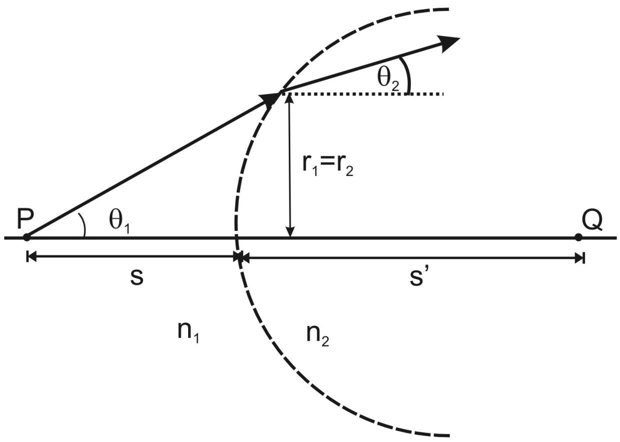
(iv) Curved dielectric interface (with radius of curvature *r*):

At the surface, $r_1 = r_2$ and

$$\theta_2 = \theta_1\frac{n_1}{n_2} - \left(1 - \frac{n_1}{n_2}\right)\frac{r_1}{R}$$

follows from the earlier equation for a spherical refracting surface by putting $\theta_1 = \frac{r_1}{s}$ and $\theta_2 = -\frac{r_1}{s'}$.
Thus

$$\begin{pmatrix} r_2 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{R}\left(1-\frac{n_1}{n_2}\right) & \frac{n_1}{n_2} \end{pmatrix}\begin{pmatrix} r_1 \\ \theta_1 \end{pmatrix}$$



so the ray transfer matrix can be written as:

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ \frac{1}{R}\left(\frac{n_1}{n_2}-1\right) & \frac{n_1}{n_2} \end{pmatrix}$$

with $R>0$ for a convex surface and $R<0$ for a concave surface.

It is convenient to define the **power** of the surface as $D = (n_2 - n_1)/R$ (in units of diopters), so

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ -\frac{D}{n_2} & \frac{n_1}{n_2} \end{pmatrix}$$

(v) A thick lens (consisting of two curved surfaces of radius of curvature R_1 and R_2 separated by a region of material of refractive index n):

The ray transfer matrix is the product of three constituent matrices (note the order):

$$\mathbf{M} = \mathbf{M_3}\,\mathbf{M_2}\,\mathbf{M_1}$$

where

$\mathbf{M_1}$ = matrix for 1st spherical interface

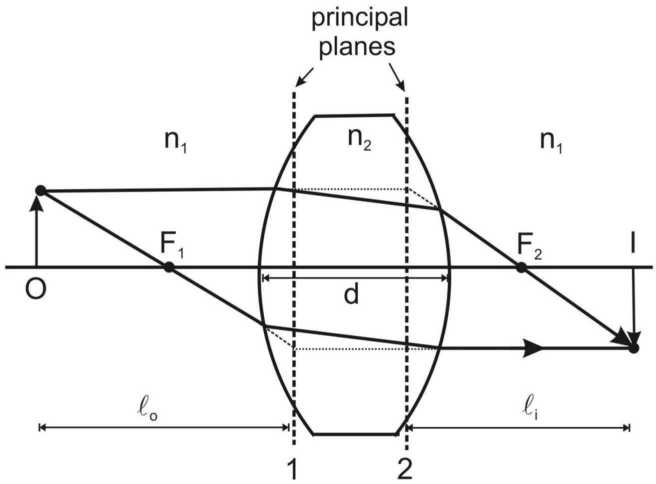
$$\mathbf{M_1} = \begin{pmatrix} 1 & 0 \\ -\frac{D_1}{n_2} & \frac{n_1}{n_2} \end{pmatrix} \quad \text{with } D_1 = (n_2 - n_1)/R_1$$

$\mathbf{M_2}$ = matrix for medium of length d

$$\mathbf{M_2} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}$$

$\mathbf{M_3}$ = matrix for 2nd spherical interface

$$\mathbf{M_3} = \begin{pmatrix} 1 & 0 \\ -\frac{D_2}{n_1} & \frac{n_2}{n_1} \end{pmatrix}$$



$\mathbf{M_1}$ comes on the right because it operates first on the vector describing the input ray. Thus, multiplication of the three matrices gives the ray transfer matrix for the thick lens:

$$\mathbf{M} = \begin{pmatrix} 1 - \frac{dD_1}{n_2} & \frac{dn_1}{n_2} \\ \frac{dD_1D_2}{n_1n_2} - \frac{D_1}{n_1} - \frac{D_2}{n_1} & 1 - \frac{dD_2}{n_2} \end{pmatrix}$$

This can be used to determine the locations of the principal planes, and the focal length is:

$$f = -\left(\frac{dD_1D_2}{n_1n_2} - \frac{D_1}{n_1} - \frac{D_2}{n_1} \right)^{-1}.$$

If ℓ_o is the distance from the object O to the first principal plane and ℓ_i is the distance from the second principal plane to the image, then it can be shown that:

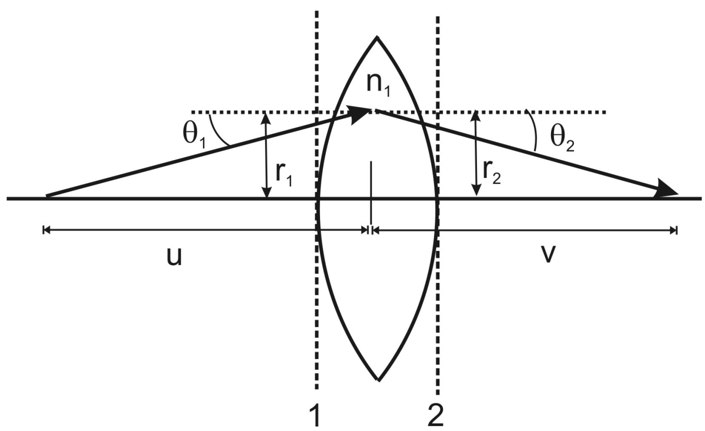
$$\frac{1}{\ell_o} + \frac{1}{\ell_i} = \frac{1}{f}$$

which is the fundamental imaging equation.

(vi) Thin lens or mirror of focal length *f*

For a lens that is sufficiently thin that, to a good approximation, d =0, the transfer matrix simplifies to

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ -\frac{D_1 + D_2}{n_1} & 1 \end{pmatrix}$$



The principal planes of such a lens are at the front and rear faces of the lens, and the focal length is given by:

$$\frac{1}{f} = \frac{D_1 + D_2}{n_1} = \left(\frac{n_2}{n_1} - 1\right)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$ **Lens makers' formula**

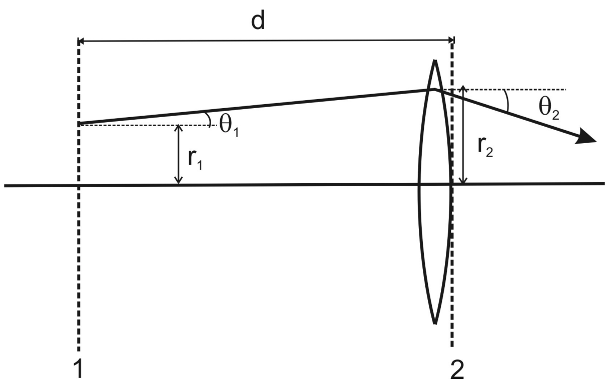
For a quartz or glass lens in air, n_2 =1 (to a good approximation). The matrix for a thin lens or a curved mirror can be written as:

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix}$$ with f=R/2 for a curved mirror.

For a biconvex lens, R_1 is positive and R_2 is negative. For a biconcave lens, R_1 is negative and R_2 is positive.

If u and v are the distances of the image and object from the centre of the lens, then:

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}.$$



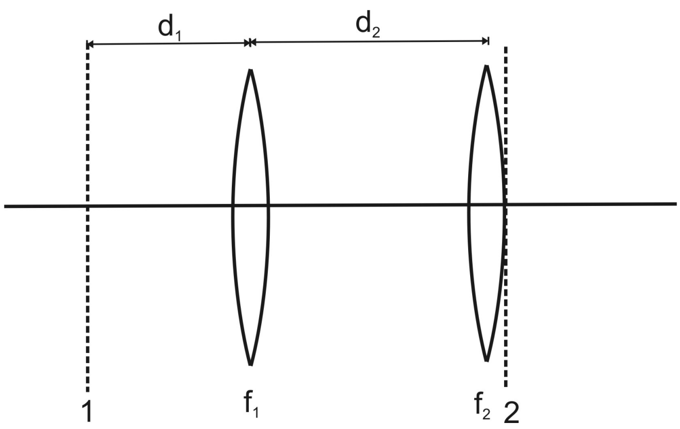
(vii) A length of uniform medium and a thin lens:
The overall transfer matrix is the product of those for the two elements:

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix}\begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & d \\ -1/f & 1 - d/f \end{pmatrix}$$

(viii) Two thin lenses:
The transfer matrix of the combination shown in the figure is:

$$\mathbf{M} = \mathbf{M}_3\,\mathbf{M}_2\,\mathbf{M}_1$$

Where $\mathbf{M}_3$ and $\mathbf{M}_1$ are the matrices for the second and first lenses and $\mathbf{M}_2$ is the matrix for the intervening medium. Thus:



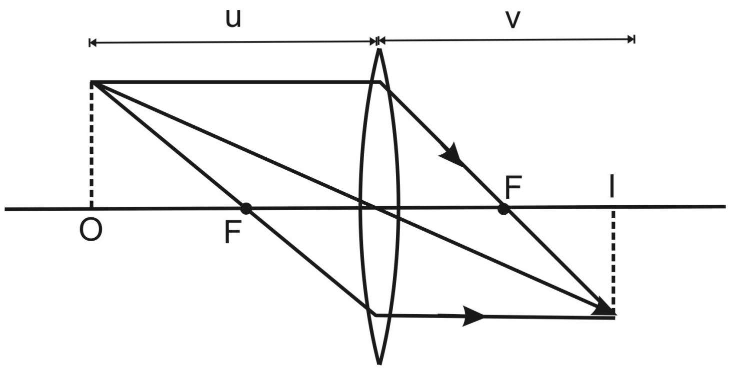
$$\mathbf{M} = \begin{pmatrix} 1 - \frac{d_2}{f_1} & d_1 + d_2 - \frac{d_1 d_2}{f_1} \\ -\frac{1}{f_1} - \frac{1}{f_2} + \frac{d_2}{f_1 f_2} & 1 - \frac{d_1}{f_1} - \frac{d_2}{f_2} + \frac{d_1 d_2}{f_1 f_2} \end{pmatrix}$$

and the focal length of the combination is:

$$f = \frac{f_1 f_2}{f_1 + f_2 - d_2}.$$

2.5 Ray tracing

A few simple rules allow geometrical construction of the principal ray paths from an object point (but neglect any aberrations).



1. The first principal ray from a point on the object to the image is drawn to pass through the first focal point. From the point where this ray intersects the first principal plane, the output ray is drawn parallel to the axis.
2. The second principal ray is directed parallel to the axis. From its intersection with the second principal plane the output ray passes through the second focal point.
3. The intersection of the two principal rays in the image space produces the image point that corresponds to the original point on the object.
4. For a thin lens a third principal ray is useful to locate the image – the ray from a point on the object that passes through the centre of the lens is not deviated by the lens.

The ratio of the height of the image to the height of the object is the **magnification**. For a thin lens m = v/u. For a general system with a ray transfer matrix

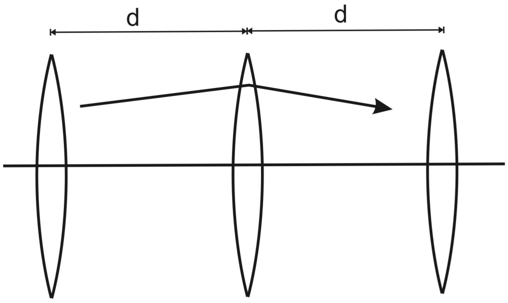
$$\mathbf{M} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

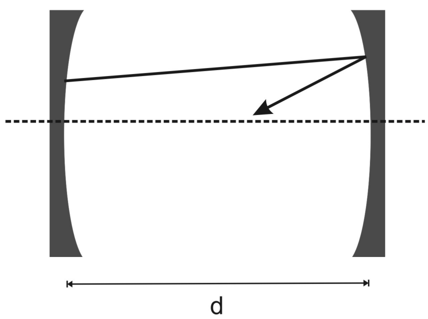
it can be shown that the magnification is m = A. Thus the ray transfer matrix of an imaging system can be written as:

$$\mathbf{M} = \begin{pmatrix} m & 0 \\ -1/f & 1/m \end{pmatrix}$$ (with the value of D chosen so that det(**M**) =1).

2.6 Periodic lens waveguides and optical resonators

If we take a series of lenses each separated by a distance *d* and start a ray at the output of the first lens, the result from section (vii) above can be used to test the stability of the system. An equivalent analysis applies to a ray within an optical resonator constructed from two curved mirrors.





$$\begin{pmatrix} r_2 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} 1 & d \\ -1/f & 1-d/f \end{pmatrix} \begin{pmatrix} r_1 \\ \theta_1 \end{pmatrix}.$$

Is there any initial ray vector such that the output ray vector is equal to the input ray vector multiplied by a constant factor? If there is, then

$$\begin{pmatrix} r_2 \\ \theta_2 \end{pmatrix} = \lambda \begin{pmatrix} r_1 \\ \theta_1 \end{pmatrix}$$

where λ is an eigenvalue of the matrix **M** and the ray vector $\begin{pmatrix} r_1 \\ \theta_1 \end{pmatrix}$ is an eigenvector. Standard matrix algebra requires that, to find eigenvalues,

$$\begin{vmatrix} 1-\lambda & d \\ -\frac{1}{f} & 1-\frac{d}{f}-\lambda \end{vmatrix} = 0$$

which has solutions, written concisely by putting $g = 1 - \frac{d}{2f}$, of:

$$\lambda = g \pm \sqrt{g^2 - 1} \quad \text{for } |g| > 1$$

$$\lambda = g \pm i\sqrt{1 - g^2} \quad \text{for } |g| < 1.$$

If the given ray vector is an eigenvector of the lens/distance combination, and it passes through a number N of such lens/distance combinations, the final output ray is:

$$\begin{pmatrix} r_N \\ \theta_N \end{pmatrix} = \lambda^N \begin{pmatrix} r_1 \\ \theta_1 \end{pmatrix}.$$

For the second eigenvalue solution $\lambda = g \pm i\sqrt{1-g^2}$ with $|g| < 1$, we can use the general expression for a complex number in modulus-argument form:

$$z = a \pm ib = |z|e^{i\phi} \quad \text{with } |z| = \sqrt{a^2 + b^2}$$

to show that $|\lambda| = g^2 + (1 - g^2) = 1$, so

$$\lambda = e^{i\phi} \quad \text{and} \quad \lambda^N = e^{iN\phi} \text{ which has a modulus less than or equal to 1.}$$

The rays will thus remain close to the axis because $r_N \le r_1$. This arrangement is therefore **stable**. The associated eigenvectors will describe the stable paths that the rays can trace through the waveguide or cavity.

For the first eigenvalue solution $\lambda = g \pm \sqrt{g^2 - 1}$ for $|g| > 1$, the modulus of λ is greater than 1 for the positive root, so the ray diverges from the axis as the number of optical elements increase and the arrangement is **unstable**.

We thus concentrate on the case in which the optical arrangement is always stable, i.e., $|g| < 1$. A condition for stability of the chain of lenses (i.e., that the ray does not diverge away from the axis) is therefore:

$|g| < 1$ so with $g = 1 - \frac{d}{2f}$, $\left|1 - \frac{d}{2f}\right| < 1$, which is satisfied for $0 < d < 4f$.

For the optical resonator made of two curved mirrors with $f = R/2$, the condition is therefore:

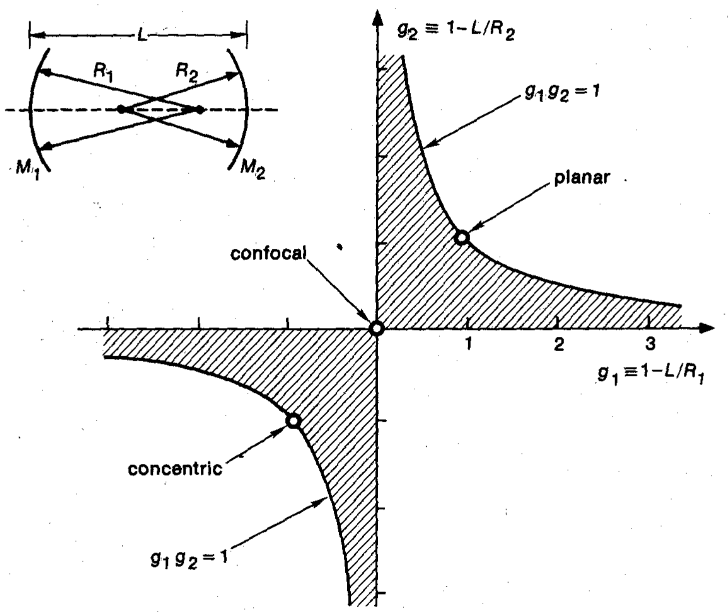
$0 < d < 2r$.

An optical cavity will be stable provided the mirrors are separated by a distance less than twice the radius of curvature of the mirrors.

For an unsymmetrical optical resonator with two mirrors of focal length f_1 and f_2 separated by a distance d, the equivalent stability criterion is**:** $\mathbf{0 < g_1 g_2 < 1}$

with $g_1 = 1 - \frac{d}{2f_1} = 1 - \frac{d}{R_1}$ and

$g_2 = 1 - \frac{d}{2f_2} = 1 - \frac{d}{R_2}$.



These conditions can be plotted as a diagram showing regions of stability (see left). There are three important classes of (symmetric) resonator that are identified as key points on the diagram:

Plane parallel:
$R_1 = R_2 = \infty$ so $g_1 = g_2 = 1$

Symmetric confocal:
$R_1 = R_2 = d$ so $g_1 = g_2 = 0$.

Symmetric concentric:
$R_1 = R_2 = ½ d$ so $g_1 = g_2 = -1$.

For symmetric resonators, the formulae describing important parameters such as beam waists in the cavity are summarised in sections 3 and 4 and illustrated for the different cavity types.

3. Gaussian beams

The outputs of lasers are usually better described by Gaussian beams rather than plane waves. Gaussian beams are restricted in their spatial extent in directions perpendicular to the beam propagation direction, even in free space. The field components of a **plane** transverse electromagnetic wave of angular frequency ω propagating in the z-direction are of the form:

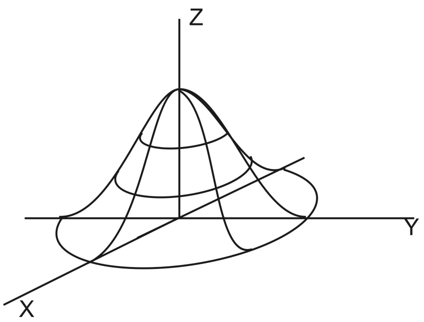
$$V = V_0\, e^{i(\omega t - kz)} = V_0 \{\cos(\omega t - kz) + i \sin(\omega t - kz)\}.$$

V_0 is a constant, independent of x and y for the plane wave.

The amplitude (A) of the field of a (zeroth order) Gaussian beam decays away from the central axis of the beam with a Gaussian like dependence in any direction. Such a beam can be written as:

$$V = A(x,y,z) e^{i(\omega t - kz)} = A(x,y,z)\{cos(\omega t - kz) + i\, sin(\omega t - kz)\}.$$

At any value of z along the beam propagation direction, A^*A gives the relative intensity distribution in the xy plane at that value of z. We will denote the zeroth order intensity distribution in the xy plane as the TEM_{00} mode.



The TEM_{00} mode can be written as:

$$A(x,y,z) = exp\left\{-i\left[P(z) + \frac{k(x^2+y^2)}{2q(z)}\right]\right\}$$

where P(z) is a phase factor, $k = 2\pi/\lambda$, and $q(z)$ is called the **beam parameter** or the **complex radius of curvature**. $q(z)$ is usually written in terms of the **phase front curvature** $R(z)$ of the beam and its **spot size** $w(z)$ as:

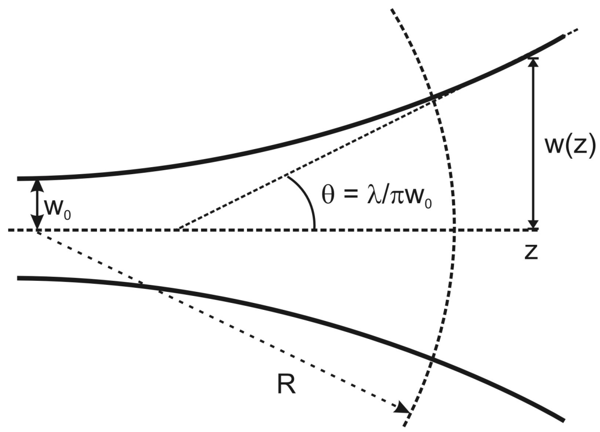
$$\frac{1}{q} = \frac{1}{R} - \frac{i\lambda}{\pi w^2}\ .$$

For a particular value of z the intensity pattern of the TEM_{00} mode is:

$$A^*A = e^{-2(x^2+y^2)/w^2} = e^{-2r^2/w^2}$$

with $r^2 = x^2 + y^2$. Thus $w(z)$ is the distance from the axis of the beam ($x = y = 0$) to a radius at which the intensity has fallen to $1/e^2$ of its axial value, and the fields to 1/e of their axial magnitude.

Every Gaussian beam has a beam waist in the direction of propagation – at the beam waist, the beam radius takes a minimum value w_0. We can arbitrarily chose the plane z=0 to correspond to a beam waist at which the curvature of the wavefront is zero, so $R(z=0)$ is infinite. At a beam waist the curvature switches from negative to positive (or vice versa) and is zero at the waist.



In the vicinity of the beam waist, the Gaussian beam maintains an approximation of a parallel beam with the smallest cross section. Further away it approximates a spherical wave with angular aperture θ. The transition between the two results in the Rayleigh range which is the approximate dividing line between near and far fields. The **Rayleigh range**, z_R is the distance travelled by the beam from z = 0 before the beam radius increases by $\sqrt{2}$ (or the beam area doubles). The value of the Rayleigh range is given by:

$$z_R = \frac{\pi w_0^2}{\lambda}$$

The **confocal parameter** is $b = 2z_R$ for a spot from a focused Gaussian beam.

The beam parameter at the beam waist is then:

$$\frac{1}{q_0} = -\frac{i\lambda}{\pi {w_0}^2} = -\frac{i}{z_R}$$

For any arbitrary value of z, from the fact that the beam parameter $q = q_0 + z$, it can be shown that:

$$w^2(z) = w_0^2\left[1+\left(\frac{\lambda z}{\pi w_0^2}\right)^2\right] = w_0^2\left[1+\left(\frac{z}{z_R}\right)^2\right].$$

The radius of curvature of the phase front at this point is:

$$R(z) = z\left[1+\left(\frac{\pi w_0^2}{\lambda z}\right)^2\right] = z + \frac{z_R^2}{z}.$$

For $z \ll z_R$, $R(z) \sim \infty$; for $z \gg z_R$, $R(z) = z$ (a spherical wave).

The beam waist expands along both the + and $-z$ directions from its beam waist along a hyperbola with asymptotes inclined to the axis at an angle:

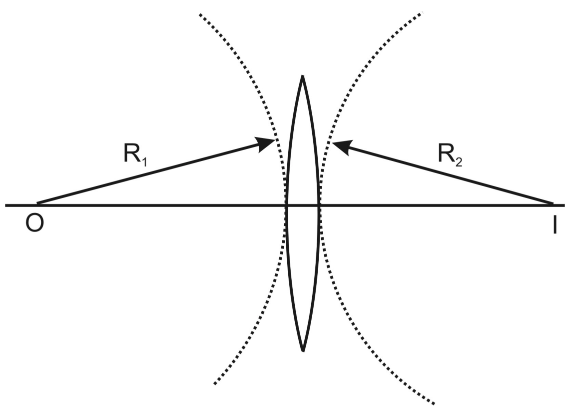
$$\tan\theta = \frac{\lambda}{\pi w_0}.$$

Note that this is generally a very shallow angle – for example for a beam waist of 100 μm and a wavelength of 500 nm, $\theta \sim 0.09^o$.

For $r^2 \ll z^2$ the surfaces of constant phase are spherical with radius of curvature $R(z)$ because for $z \gg \pi w_0^2/\lambda$, $R(z) = z$.

The collimated range of a Gaussian beam = $2z_R = \frac{2\pi w_0^2}{\lambda}$. Thus for w_0=2 mm, the collimated range is ~ 50 m for visible light. In the far field, the beam expands linearly with distance.

A lens or series of lenses can be used to focus a Gaussian beam without changing the transverse intensity pattern – the radius of curvature does, however, change. When a spherical wave of radius of curvature R_1 strikes a thin lens, the object distance (to the point of origin of the wave) is also R_1. The radius of curvature of the output beam after passage through the lens, R_2, must obey:



$$\frac{1}{R_2} = \frac{1}{R_1} - \frac{1}{f}.$$

If the spot size is unchanged at the lens, the beam parameter obeys:

$$\frac{1}{q_2} = \frac{1}{q_1} - \frac{1}{f}$$

from which it can be shown that the minimum spot size of a TEM_{00} Gaussian beam focussed by a lens is:

$$w_f = f\left[\left(\frac{\lambda}{\pi w_1}\right)^2 + \frac{w_1^2}{R_1^2}\right]^{1/2}$$

where w_1 and R_1 are the laser beam spot size and radius of curvature at the input face of the lens. If the lens is far from the beam waist (the object point), this reduces to

$$w_f = \frac{f\lambda}{\pi w_1}.$$

The spot size to which a laser can be focused cannot be reduced without limit just by reducing the focal length because the lens ultimately becomes a sphere and cannot be treated as a thin lens. To focus a laser beam to a small spot, the beam should be expanded and collimated (by a Galilean telescope) before the focusing lens. To prevent diffraction effects, the lens aperture must be larger than the spot size at the lens – a lens diameter D = 2.8w_1 is commonly used. In this case if the lens is placed in a collimated beam:

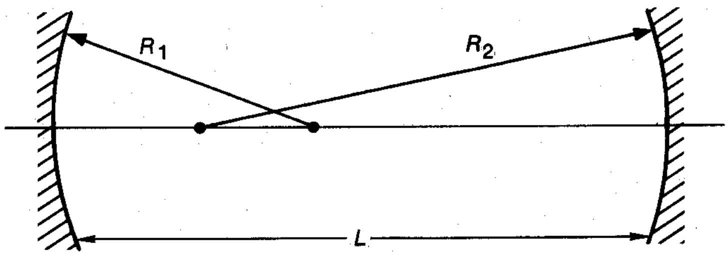
$$2w_f = \frac{5.6f\lambda}{\pi D} = \frac{1.78f\lambda}{D}.$$

In practice it is very difficult to manufacture spherical lenses with very small values of *f/D* (called the ***f* /number**) that achieve the diffraction-limited performance predicted by this equation. Commercial spherical lenses achieve spot diameters $2w_f$ of about 10λ. Smaller spot sizes are possible with **aspheric lenses**. The **depth of focus** is given by $2z_R$.

4. Cavity modes

4.1 Spatial distributions of light in a cavity

The stability diagram (page 14) identified various classes of stable cavity (constructed from two mirrors), and for Gaussian beams propagating in such cavities, various parameters summarising the focusing of the beam within the cavity (for a TEM_{00} mode) can be calculated from straightforward formulae.



For a cavity of length *L* constructed of two concave mirrors of radii of curvature R_1 and R_2 the cavity *g*-parameters are:

$$g_1 = 1 - \frac{L}{R_1} \quad \text{and} \quad g_2 = 1 - \frac{L}{R_2}.$$

The trapped Gaussian beam will have a Rayleigh range

$$z_R^2 = \frac{g_1 g_2 (1 - g_1 g_2)}{(g_1 + g_2 - 2g_1 g_2)^2} L^2$$

and the locations of the two mirrors relative to the Gaussian beam waist will be

$$z_1 = \frac{g_2(1-g_1)}{g_1 + g_2 - 2g_1 g_2} L \qquad\qquad z_2 = \frac{g_1(1-g_2)}{g_1 + g_2 - 2g_1 g_2} L$$

If mirror M_1 is located to the left of the beam waist, z_1 will be negative. The waist spot size is

$$w_0^2 = \frac{L\lambda}{\pi} \sqrt{\frac{g_1 g_2 (1 - g_1 g_2)}{(g_1 + g_2 - 2g_1 g_2)^2}}$$

and the spot sizes on the mirrors at the ends of the resonator are:

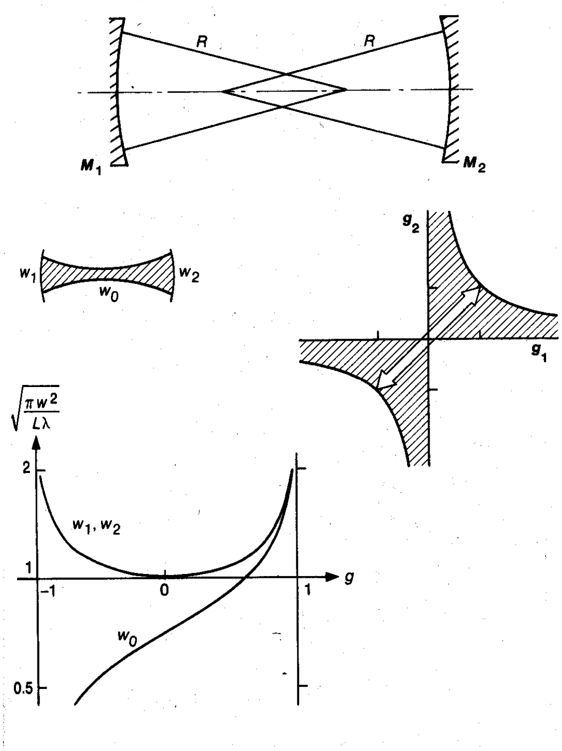
$$w_1^2 = \frac{L\lambda}{\pi} \sqrt{\frac{g_2}{g_1(1 - g_1 g_2)}} \quad \text{and} \quad w_2^2 = \frac{L\lambda}{\pi} \sqrt{\frac{g_1}{g_2(1 - g_1 g_2)}} .$$

These latter two values must be (much) less than the mirror radii to avoid **diffraction losses** from the cavity – i.e., leakage of light around the mirror and diffraction from the mirror aperture. The radii of curvatures of the mirrors should match those of the beam fronts at the mirrors.

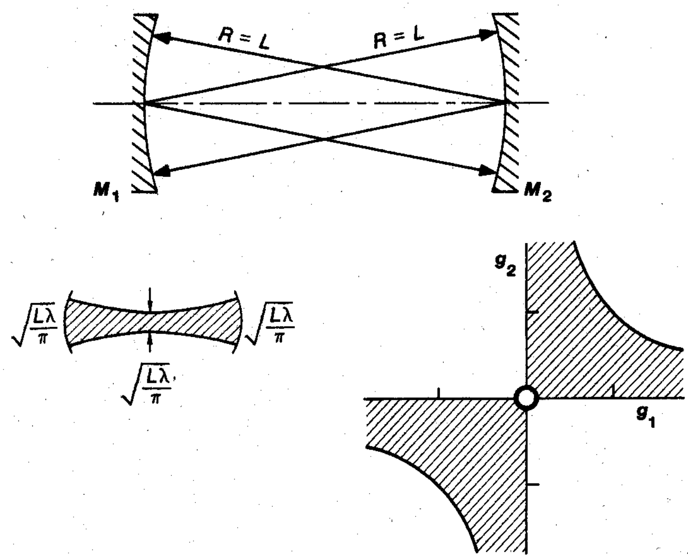
For **symmetric resonators**, $g_1 = g_2$ and the above formulae simplify to (g denotes the single parameter):

$$w_0^2 = \frac{L\lambda}{\pi} \sqrt{\frac{1+g}{4(1-g)}} \qquad \text{and} \qquad w_1^2 = w_2^2 = \frac{L\lambda}{\pi} \sqrt{\frac{1}{1-g^2}} .$$

Symmetric resonators lie along the diagonal in the cavity stability diagram, with allowed ranges from g=1 (planar) through g=0 (confocal) to g=-1 (concentric), as illustrated in the diagrams below.



Symmetric stable resonators lie along the diagonal axis in the g_1,g_2 plane.

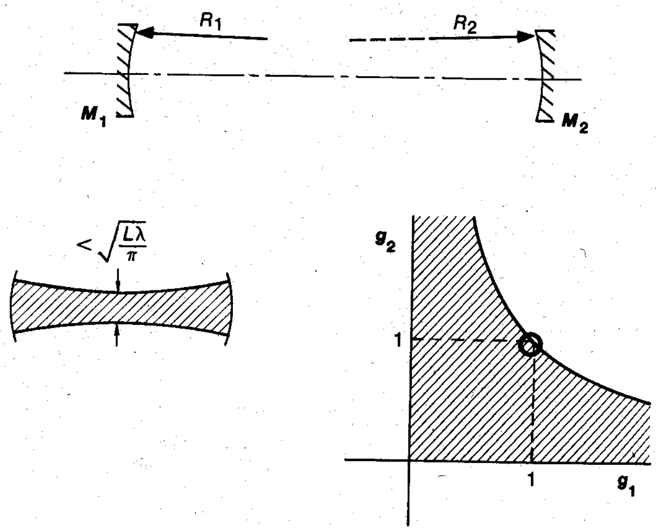


Symmetric confocal resonator. R_1=R_2=L and the focal points of the two end mirrors (f = R/2) coincide at the centre of the resonator. The mirrors are separated by exactly two Rayleigh ranges.

The spot sizes at the centre and end mirrors of a **symmetric confocal resonator** are:

$$w_0^2 = \frac{L\lambda}{2\pi} \qquad w_1^2 = w_2^2 = \frac{L\lambda}{\pi}.$$

The confocal resonator has the overall smallest average spot diameter along its length of any stable resonator, although other resonator designs may have a smaller waist size at one spot. This resonator is highly insensitive to misalignment of either mirror – tilting one mirror leaves the centre of curvature located on the other mirror surface but displaces the optical axis by a small amount.

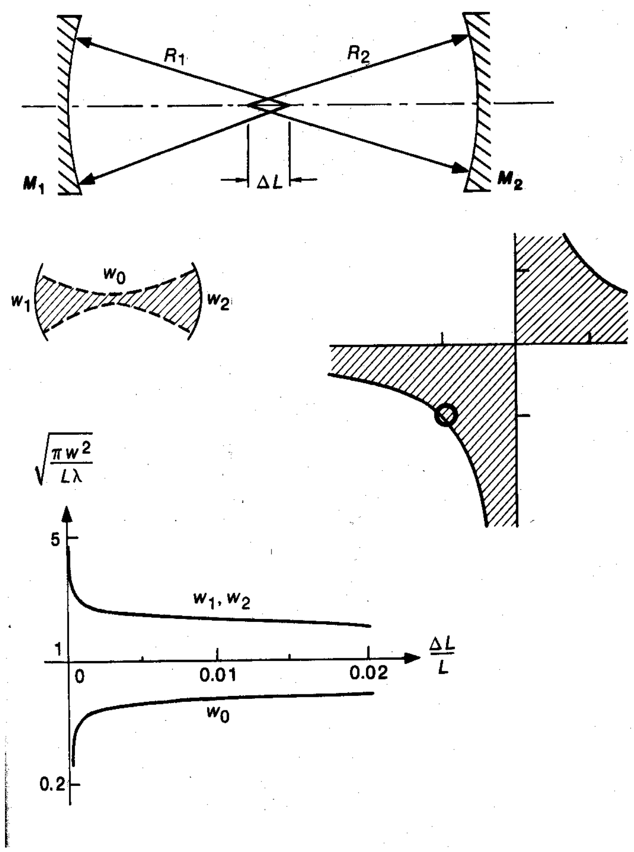


Long-radius (near planar) resonators have $R_1 = R_2 \sim \infty$ and $g_1 = g_2 \sim 1$. Note that we cannot use truly planar mirrors because the beam will walk off them rapidly unless the cavity is absolutely perfectly aligned.

The spot sizes are all large and approximately equal:

$$w_0^2 \approx w_1^2 \approx w_2^2 \approx \frac{L\lambda}{\pi}\sqrt{\frac{R}{2L}} \quad \text{for R >> L.}$$

Very delicate mirror alignment is necessary to ensure cavity stability for this design.



Near concentric resonators give large spot sizes on the mirrors and a very small spot size at the cavity centre. If the cavity length L is less than the sum of the two mirror radii of curvature by a small amount δL then

$$w_0^2 \approx \frac{L\lambda}{\pi}\sqrt{\frac{\delta L}{4L}} \quad \text{for } \delta L << L.$$

The end mirror spot sizes are

$$w_1^2 = w_2^2 \approx \frac{L\lambda}{\pi}\sqrt{\frac{4L}{\delta L}}.$$

This resonator design is very sensitive to mirror misalignments.

4.2 Frequencies of cavity modes

The frequencies of light that can be sustained within a cavity are determined by the condition that the accumulated phase shift for a complete round trip must be some integer multiple of 2π - and thus that the resonator length is equal to an integer or half-integer number of wavelengths so that a stable standing wave pattern is established.

The difference between the phase of a wave at the first and second mirrors of the cavity is

$$\delta\phi = \frac{2\pi L}{\lambda} = \frac{2\pi\nu L}{c} = kL \quad \text{where } k \text{ is the wave number } (k = 2\pi/\lambda).$$

This result can be understood as follows. If the length of the cavity is equal to an integer number (n) of wavelengths plus some fraction a of a wavelength, then $(n+a)\lambda = L$. The phase of the wave after $n + a$ oscillations is $\delta\phi = 2\pi a$ because the n complete oscillations make no change to the initial phase. Thus $\delta\phi = 2\pi(L/\lambda - n) = 2\pi L/\lambda$ because the phase shift of $-2n\pi$ is equivalent to a phase shift of zero.

A round trip of the cavity therefore causes a phase shift in the wave of $2\delta\phi$, but for a standing wave this phase shift must be zero or some integer multiple of 2π (for constructive interference of the circulating wave). Thus for a round trip of the cavity,

$$\frac{4\pi L}{\lambda} = 2\pi q \qquad \text{where } q \text{ is an integer, so}$$

$$\frac{4\pi\nu L}{c} = 2\pi q \qquad \Rightarrow \qquad \nu = \frac{qc}{2L} \quad \text{are the allowed frequencies in the cavity.}$$

The frequency interval between the q^{th} and $(q+1)^{th}$ mode is called the **free spectral range** of the cavity and is

$$\Delta\nu = \frac{c}{2L}.$$

In these formulae, q is generally a very large number.

For a Gaussian mode propagating in the cavity, the treatment of phase shifts is more complicated than the preceding discussion suggests. Any (zeroth order) Gaussian TEM_{00} beam passing through a focus undergoes a phase shift of π in going from the far field, through the focus to the far field limit again. This is known as the Guoy effect. The actual phase at a distance z from the focus (where the phase is chosen to be zero) is given by

$$\psi_{00}(z) = tan^{-1}(z / z_R)$$

where z_R is the usual Rayleigh range. Thus, only for $z >> z_R$ does the phase reach $\pi/2$ of that at the focus. For a large distance in the $-z$ direction, the phase is $-\pi/2$ so overall a phase change of π is accumulated.

Higher order Gaussian beams, denoted by TEM_{nm} (see later) undergo phase shifts different from that of the zeroth order Gaussian beam. The Guoy phase shift is

$$\psi_{nm}(z) = (n + m + 1)\, \psi_{00}(z).$$

The total change in phase between mirrors 1 and 2 in the cavity, a distance L apart, and located at positions z_1 and z_2 is given by the sum of the effect due to the mirror separation (as described earlier) and the Guoy shift:

$$\varphi_{nm}(z_2 - z_1) = \frac{2\pi L}{\lambda} - (n+m+1)[\psi_{00}(z_2) - \psi_{00}(z_1)].$$

As a result of different phase shifts for the various TEM_{nm} modes, these modes resonate at different frequencies and can cause interferences between modes as they exit the cavity.

It can be shown that

$$\psi_{00}(z_2) - \psi_{00}(z_1) = cos^{-1}\left(\pm\sqrt{g_1 g_2}\right)$$

with the plus sign applying for $g_1,g_2 > 0$ and the – sign for $g_1,g_2 < 0$.

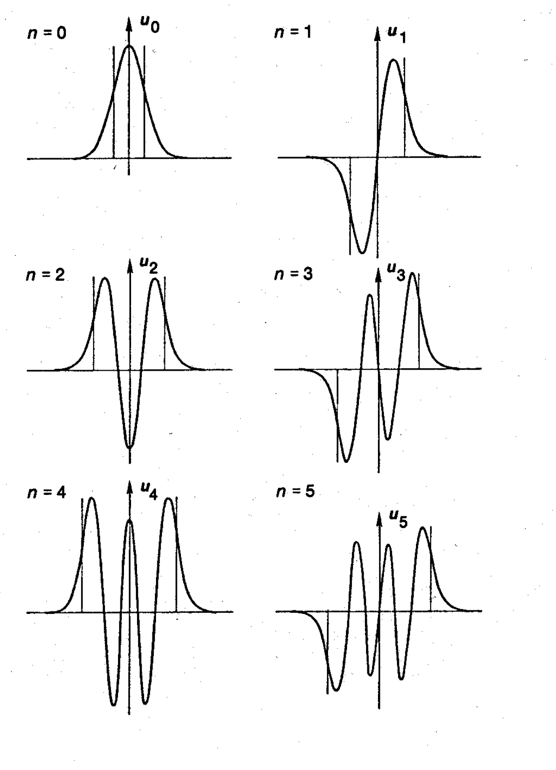
The electric field of a Gaussian wave propagating along the z-axis in a homogeneous medium is:

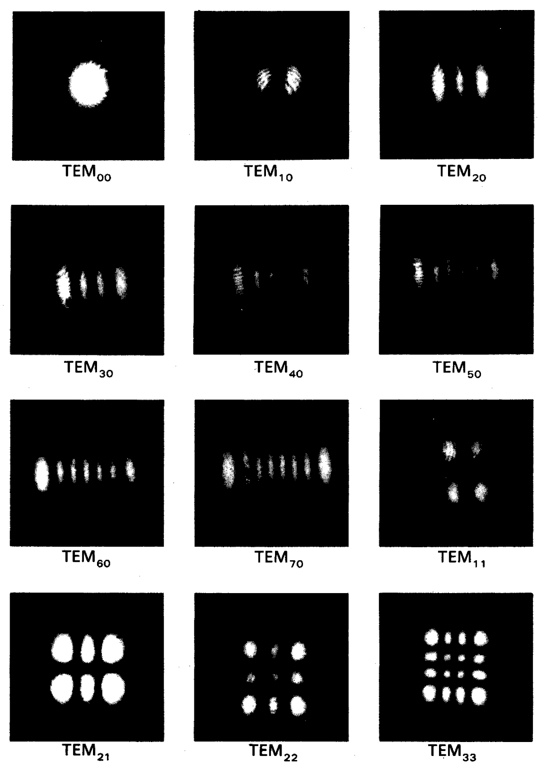
$$E_{nm}(x,y,z) = E_0 \frac{w_0}{w(z)} H_n\left(\sqrt{2}\frac{x}{w(z)}\right) H_m\left(\sqrt{2}\frac{y}{w(z)}\right) exp\left[-\frac{x^2+y^2}{w^2(z)} - \frac{ik\left(x^2+y^2\right)}{2R(z)} - ikz + i(n+m+1)\psi_{00}(z)\right]$$

Here, H_n is a Hermite polynomial of order n. The transverse variation of the electric field along x (or y) is thus of the form:

$$E_n(x) \propto H_n\left(\sqrt{2}\frac{x}{w(z)}\right) exp\left[-\frac{x^2}{w^2(z)}\right]$$

The TEM_{nm} describe the longitudinal and transverse mode frequencies with m and n denoting the number of nodes along x and y. The zeroth order Gaussian modes are described by TEM_{00}, and higher order transverse modes by TEM_{nm} with n, m ≠ 0. Cross sectional views of such modes are shown in the figure.





The resonance condition for a standing wave cavity is that the round trip phase shift must be an integer multiple of 2π, or that the one-pass phase shift must be an integer multiple of π. Thus

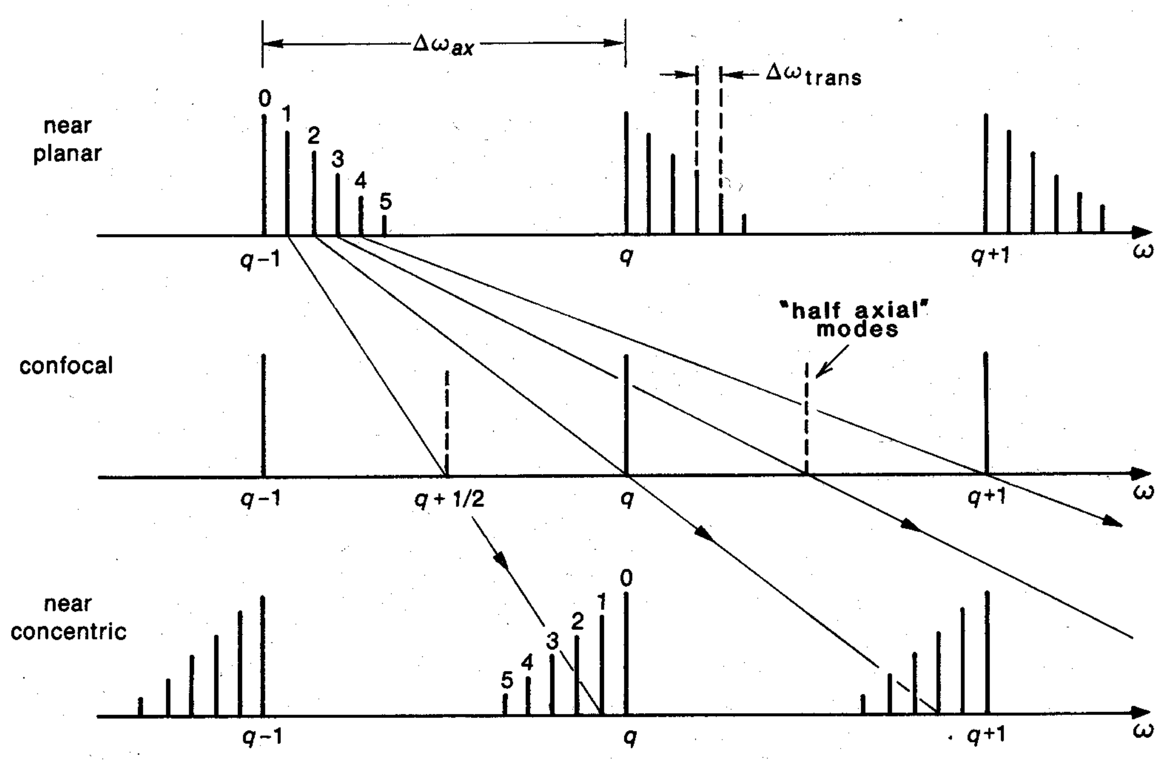
$$\frac{2\pi\nu L}{c} - (n+m+1)\cos^{-1}\left(\pm\sqrt{g_1 g_2}\right) = q\pi .$$

The resonance frequencies of the longitudinal plus transverse modes in the cavity must thus be given by:

$$\nu = \nu_{qnm} = \frac{c}{2L}\left[q + \frac{1}{\pi}(n+m+1)\cos^{-1}\left(\pm\sqrt{g_1 g_2}\right)\right] .$$

The factor $\frac{1}{\pi}\cos^{-1}\left(\pm\sqrt{g_1 g_2}\right)$ arising from the Guoy phase shift takes limiting values of 0 (near planar cavity), ½ (near confocal cavity) and –1 (near concentric cavity).

The diagram below shows the transverse mode frequencies (plotted as angular frequency, ω=2πν) in various stable Gaussian resonators.



In the near planar case, the transverse mode frequencies associated with a single longitudinal (also known as **axial**) mode are all clustered to the high frequency side of the longitudinal mode of frequency ν_{q00}, with equal spacings that are small compared to the longitudinal mode spacing. The mode spot is large and the resonator length short compared to the Rayleigh range, so the transverse modes pick up little additional Guoy phase shift. The longitudinal modes are separated by the usual c/2L factor (denoted as $\Delta\omega_{ax}$ in the figure).

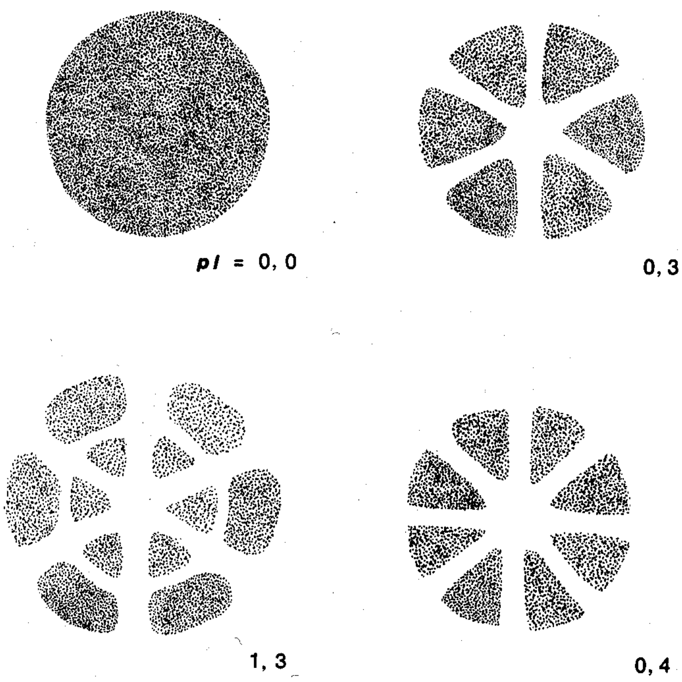
In the confocal resonator, the 01 and 10 transverse modes associated with the q^{th} longitudinal mode fall exactly half way between the *q* and *q* +1 longitudinal modes. The *q*11, *q*02 and *q*20 modes coincide with the q+2,00 mode, etc. All the even symmetry transverse modes are thus degenerate, as are all the odd symmetry modes.

A technical note about mode types

The electric field equation given earlier is obtained by solving the wave equation for the cavity by using an expansion in Hermite-Gaussian polynomial functions (a complete, orthogonal basis set). The 2-D functions are based on Cartesian coordinates (*x*,*y*) and thus most naturally describe fields that vary in rectangular or square symmetry perpendicular to the cavity axis. An alternative but equally valid family of solutions to the wave equation can be written in terms of cylindrical coordinates and are known as the Laguerre-Gaussian solutions. They are expressed in terms of a radial index *p* (the number of nodes along the radial coordinate) and an angular index *m* that describes the number of angular nodes. The modes have cylindrical symmetry, with circles of constant intensity in the radical direction and an $e^{im\theta}$ variation in the azimuthal direction. An arbitrary optical beam can be expanded in either Laguerre-Gaussian or Hermite-Gaussian functions and both methods are equally valid. The former are perhaps more appropriate for problems with cylindrical symmetry (such as a ring-down cavity with spherical mirrors) whereas most laser cavities naturally generate Hermite-Gaussian type modes because elements such as Brewster angle polarizers promote astigmatism between *x* and *y* directions. For completeness, the electric field amplitude expressed in Laguerre polynomials is:

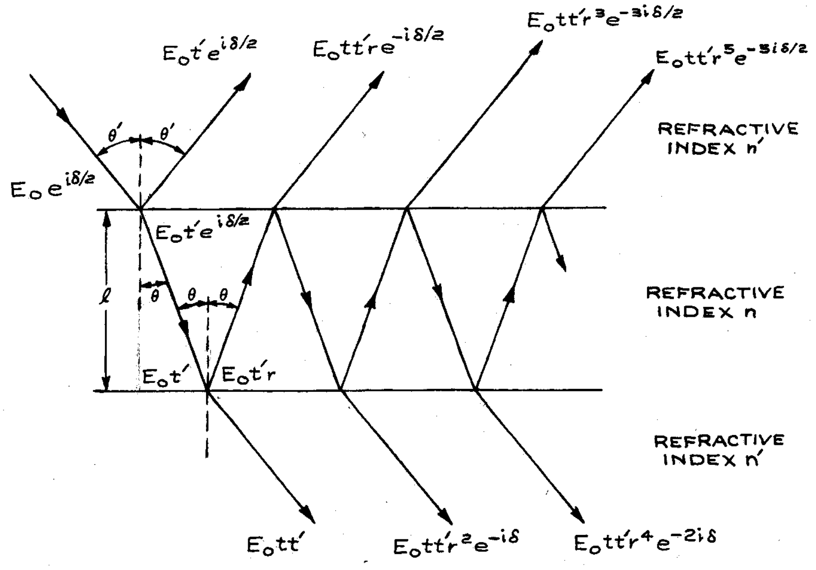
$$E_{nm}(r,\theta,z) = E_0 M_{mp} \frac{w_0}{w(z)} H_n \left(\sqrt{2}\frac{r}{w(z)} \right)^m L_p^m \left(\frac{2r^2}{w^2(z)} \right) exp\left[i(2p+m+1)\psi_{00}(z) \right]$$
$$\times\, exp\left[-ik\frac{r^2}{2R(z)} \right] exp[im\theta]\, exp\left[-\frac{r^2}{w^2(z)} \right]$$

where M_{mp} is a scaling factor. The modes are labelled as TEM_{pm} and examples are shown in the figure.



4.3 Cavity transmission, free spectral range, finesse and mode widths

An optical cavity is a sophisticated version of a Fabry-Perot cavity of the type used in etalons (which provide frequency references in spectroscopy). An etalon consists of a pair of flat, parallel, (partially) reflecting plates. The figure shows the successive reflected and transmitted field amplitudes of a plane EM wave striking an etalon at an angle of incidence θ' (assuming the two plates have equal reflectivities).



The optical path difference between successive transmitted waves is $2n\ell\cos\theta$ where ℓ is the interface spacing and n is the refractive index of the medium between the interfaces. θ and θ' are related by Snell's law. The phase difference between successive transmitted waves is:

$$\delta = 2k\ell\cos\theta + 2\eta$$

where η is the phase change (if any) on reflection. The total resultant transmitted complex amplitude is:

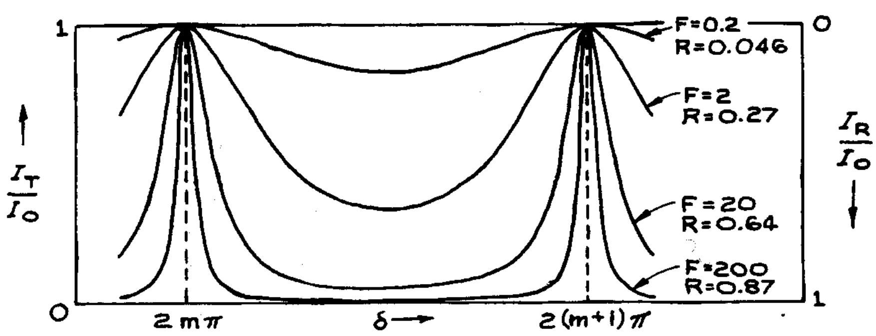
$$E_T = E_0tt' + E_0tt'r^2e^{-i\delta} + E_0tt'r^4e^{-2i\delta} + ... = \frac{E_0tt'}{1-r^2e^{-i\delta}} \ .$$

Here r and t are the reflection and transmission coefficients of the waves passing from the medium of refractive index n to that of refractive index n' at each interface and r' and t' are the corresponding coefficients for passage from n' to n. E_0 is the amplitude of the incident electric field. The total transmitted intensity is:

$$I_T \propto |E_T|^2 = \frac{I_0|tt'|^2}{\left|1-r^2e^{-i\delta}\right|^2} \ .$$

Note that $|tt'| = T$ and $|rr'| = R$ where T and R are the transmittance and reflectance of each interface. If there is no energy lost in the reflection process, T = 1-R, but if there are small absorption losses A then T = 1-R –A. If A = 0 then

$$\frac{I_T}{I_0} = \frac{1}{1+\frac{4R}{(1-R)^2}\sin^2\frac{\delta}{2}} \ .$$



This variation of transmitted intensity with δ is called an **Airy function** and is shown in the figure for different values of R. Note that the transmitted peaks get sharper (as a function of phase) as R increases towards its maximum value of 1. For A ≠0,

$$\frac{I_T}{I_0} = \frac{\left(\frac{T}{1-R}\right)^2}{1+\frac{4R}{(1-R)^2}\sin^2\frac{\delta}{2}} \ .$$

In either case, there is maximum transmitted intensity when

$$\delta = 2m\pi$$

where m is an integer. If the phase change on reflection, η is neglected, this reduces to

$$2\ell \cos\theta = m\lambda.$$

For normal incidence, $2\ell = m\lambda$.

The frequencies of transmission maxima for normal incidence occur at:

$$\nu_0 = \frac{mc}{2\ell}$$

and the frequency between successive transmission maxima is the **free spectral range**:

$$\Delta\nu = \frac{c}{2\ell}.$$

When R is close to 1, all phase angles δ within a transmission maximum differ from the value $2m\pi$ by only a small amount (because the peak is so sharp) so

$$\delta = \frac{4\pi\nu\ell}{c} = \frac{4\pi\nu_0\ell}{c} + \frac{4\pi(\nu - \nu_0)\ell}{c}$$

where ν_0 is the centre frequency of the transmitted peak. This can be written in the form

$$\delta = 2m\pi + \frac{2\pi(\nu - \nu_0)}{\Delta\nu}$$

and thus we obtain:

$$\frac{I_T}{I_0} = \frac{1}{1 + \frac{4\pi^2 R}{(1-R)^2}\frac{(\nu - \nu_0)^2}{\Delta\nu^2}}.$$

Writing the **finesse** of the cavity, F, as:

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

the shape of a narrow transmission maximum can be written as:

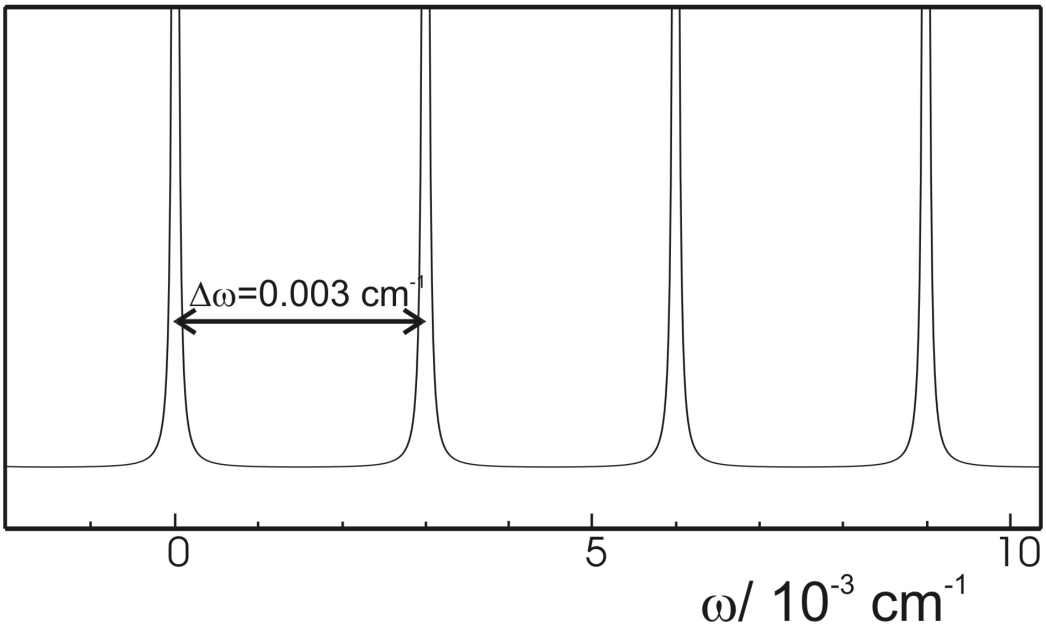
$$\frac{I_T}{I_0} = \frac{1}{1 + \frac{4(\nu - \nu_0)^2}{\Delta\nu_{\frac{1}{2}}^2}}$$

where $\Delta\nu_{\frac{1}{2}}$ is the full width at half maximum (FWHM) of the transmission peak and is given by:

$$\Delta\nu_{\frac{1}{2}} = \frac{\Delta\nu}{F}\,.$$

Thus, the higher the mirror reflectivity, the larger the finesse and the sharper are the transmission peaks.

For a 1-m long cavity with mirrors of reflectivity R = 0.9999, for example, $F \approx$ 31000 and the FWHM of the cavity modes is 4.8 kHz. These modes are separated by a free spectral range of 150 MHz.



Mode structure of a 1.5-m cavity (plotted in wavenumber units).