University of Toronto
ADVANCED PHYSICS LABORATORY

GAUS

Gaussian Beams and Near-field Diffraction

Revisions:

2021 September v1.3:	Robin Marjoribanks
2020 September v1.0:	Robin Marjoribanks
2020 August v0.8:	Robin Marjoribanks
2020 February v0.7:	Robin Marjoribanks (original author)

Please send any corrections, comments, or suggestions to the professor currently supervising this experiment, the author of the most recent revision above, or the Advanced Physics Lab Coordinator.

Overview

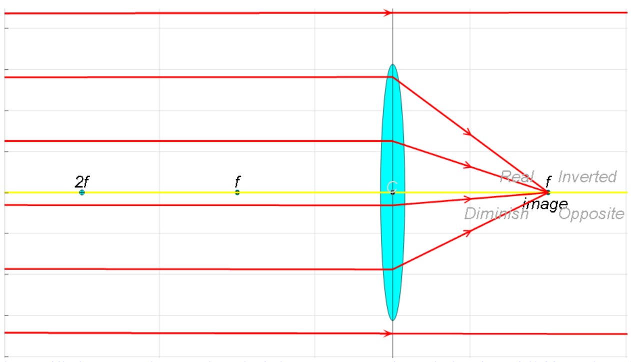
The purpose of this experiment is to extend simplified notions geometric optics (ray-tracing) and of *far-field* (Fraunhofer) diffraction to the *near-field regime* essential for Gaussian beams in lasers, and for Fresnel diffraction which produces the phenomenon of the Spot of Arago (Poisson's spot).

Introduction

Laser beams in science-fiction, used to make an almost-impossible maze that spies must defeat.

It's a common notion that light travels in straight lines, but this is only true in an approximation. In geometrical optics we can do *ray-tracing*, and determine where a bundle of parallel input rays passing through a lens will come to a focus.

The limitation of this approach is obvious — in geometrical optics, the rays all come together at a single point in space, resulting in infinite intensity, which is absurd. More than just 'rays' must be going on, in a more complete viewpoint, and the answer is known to anyone who has passed light through a fine slit: beyond a slit, light spreads out more and more broadly in angle as the slit is narrowed down.

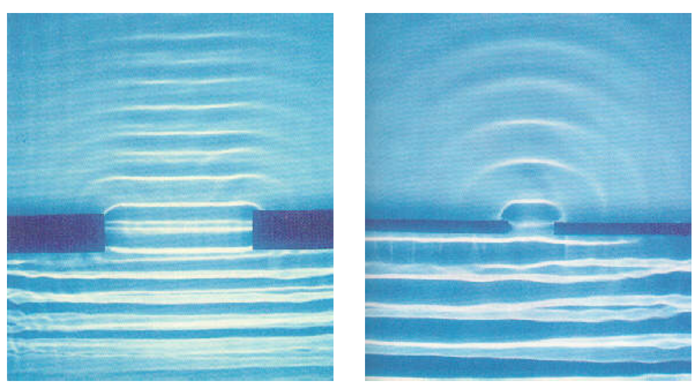


http://labman.phys.utk.edu/phys136core/modules/m9/diffraction.htm

Ray tracing through a lens to a focus. Ray-tracing is quite legitimate for modelling lots of classical optical systems such as DSLR camera lenses.

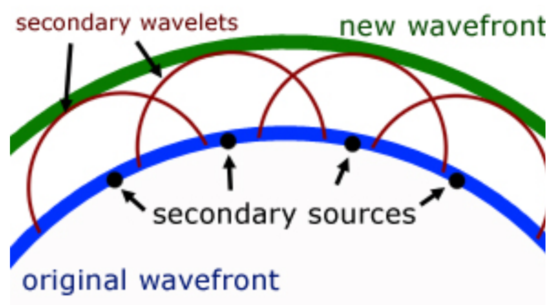
This is true for any wave, and ties both to the time-bandwidth product theorem for a brief musical note or other temporal waveform, and to Heisenberg's Uncertainty Principle in quantum mechanics — once de Broglie had posited that particles have a wave nature, all of the ramifications of wave nature followed immediately, including the H.U.P. for position and momentum.

Waves for the win — It was Christian Huygens (1629-1695) who noticed that if one can draw the wavefront of any wave, one can deduce much about the future of the wavefront. In a small advancement of the wave, the new wavefront can be found by constructing a little spherical wavelet at every point along the wavefront and letting it expand slightly. The whole assembly of wavelets, all along the wavefront, set



Wave-tank illustrations show how plane waves spread, or diffract, beyond an aperture. Smaller apertures Δx result in a bigger lateral spread Δk_x of wavevectors, which tie to momentum.

out the new position of the wavefront by taking a line tangent to all the tiny spheres.



Huygen's construction predicts the new wavefront after propagating a little distance. Fresnel took it further, and literally: the new wave is the mathematical sum, the net interference, of all the little wavelets, *including phase*.

From this, it's immediately obvious that a concave-forward wavefront should converge to a focus, and a convex-forward wavefront should diverge.

Augustin-Jean Fresnel (1788-1827) carried Huygens's geometrical construction one step further, taking not only the tangent to the spherical wavelets to identify the curve of the new wavefront, but the mathematical sum of wavelets, including their phase of oscillation, to find the evolution of the wave completely. This is the principle of superposition and interference which successfully proved how an opaque disk can nonetheless form a very bright spot in the middle of its shadow. (Though first observed by Delisle, and by Maraldi, a century beforehand, this spot is called by two names: *Poisson's spot*, and the *spot of Arago*. Poisson was a theoretician who disbelieved Fresnel's theory, and showed that it must predict a bright spot amid the dark shadow cast by a disk – which presumably proved Fresnel was wrong, by *reductio ad absurdum.* Arago was the head of the prize committee for the competition in which Fresnel presented his new theory – he took the question to his laboratory, and with a 2mm metallic disk he showed the absurd spot in fact existed. Though this was not in fact the turning point supposed in legend, and though much discussion about the fundamental nature of light and the meaning of Fresnel's theory continued, the committee agreed on enough to award Fresnel the Grand Prix of 1819.)

Why waves don't travel as rays

Only a *plane wave* possesses a perfectly well-defined (zero uncertainty) wave vector k, and the associated cost is that the wave must necessarily have infinite extent (complete uncertainty in position x). If we make a barrier, as to create a slit, to eliminate much of the transverse extent, we remove an infinite number of little Huygen's wavelets along the wavefronts, left and right, which previously interfered to continually keep reconstructing a perfect plane wave going forward. The consequence is that the smaller we make such a slit, the more nearly the transmitted wave on the other side looks like a spherical wave coming from a single point without neighbours.

Turn this around, in order to understand what is required, to focus light down to a spot: we can simply run time backwards for the light leaving a tiny slit, to see that to make a smaller and smaller focal spot takes a wider and wider cone of waves converging. The no-longer-unique vectors k point in a bigger and bigger range of directions Δk, in order to define a smaller and smaller spot Δx.

In fact, diffraction theory teaches that the field pattern formed at a distance one focal length beyond a lens is the Fourier transform of the amplitude and phase of the light incident on the lens. You may already know that the Fourier transform of a gaussian function is another gaussian function. All these observations together lead us directly to *gaussian beam optics*.

Before going farther, read the attached Appendix on Gaussian Beams as solutions to the wave equation in the case of propagation of light nearly parallel to an axis — the *paraxial wave equation*.

Objectives

The study of the optical physics of gaussian beams, by:

a) imaging the intensity distribution of a gaussian beam for a number of positions before and through its focal spot
b) determining the relationship $R(z)$ between wavefront radius of curvature R and axial position z, by reflecting the beam exactly back on itself from surfaces of different curvature R.
c) investigating near-field diffraction, the optical physics that is in play for a gaussian beam near its smallest spot-size at focus. Real beams cannot focus to zero size, for the same reason that light passed through a fine slit naturally spreads out after passing.

New techniques:

i. how the ABCD matrix method you may have learned for geometrical optics without diffraction is adapted and still valid for gaussian optics, where diffraction is a governing fact
ii. methods to align optics precisely

Associated software:

- *ImageJ* (free) — https://imagej.nih.gov/ij/
-

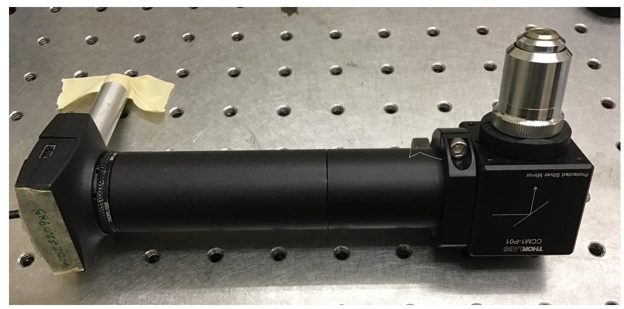
EQUIPMENT PROVIDED FOR YOUR INVESTIGATIONS

Laser — HeNe laser at 632nm with smooth transverse profile, key-operated. Output power is 5mW and is safe without laser goggles, but you must never look directly into the beam, and should always protect against unintended reflections and other stray beams.

Beam expander — A two-lens 10x beam telescope expands the beam to about 1cm and collimates it to have minimum divergence.

Lenses — Several lenses of different focal lengths are mounted.

XYZ Positioner — Take the horizontal axis of the laser optical path to be *z*, the transverse horizontal axis *x*, and the tranverse vertical axis to be *y*. The 3-axis translation stage will give you measurements of its displacements in *x, y, z*.

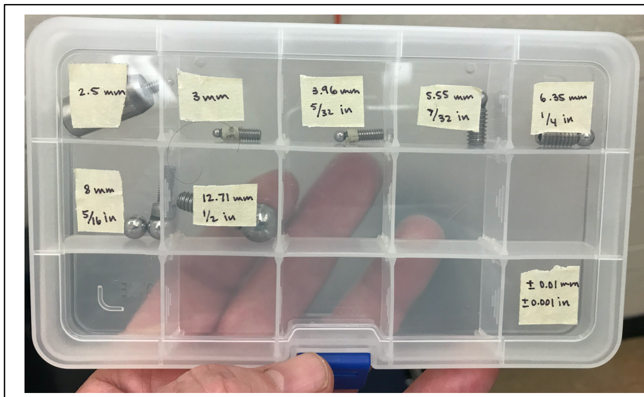


Beam microscope. The USB camera connected can be controlled and analysed using ImageJ software. Calibrate by moving the microscope sideways a distance measured on micrometers, and comparing to movement of the image on the camera.

Beam microscope (beam profiler) — A microscope objective coupled to a 1" lens tube, with a 90° mirror to make a compact setup, ending in a digital camera at the image plane. Magnification is the ratio of distances (objective-lens to camera sensor)/(object plane to objective-lens). The beam profiler can be mounted on the XYZ positioner.

Reflectors — A flat mirror, a number of steel or glass balls of different sizes are provided. The curved surfaces of different lenses also reflect light from the interfaces, where the index of refraction changes (Fresnel reflection).

If you have plans that require something else, please speak first to the supervising professor. Lab staff are happy to help, if components are available on the shelf or can be easily and quickly obtained.



Mounted steel balls: spherical reflectors.
Consider also the mounted glass balls, especially also high-quality (better than λ/10) spherical-surface glass lenses of different focal length (provided).

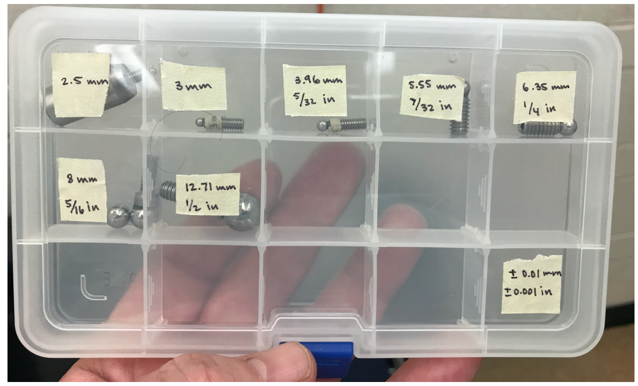
GAUSSIAN BEAM CHARACTERIZATION

Beam Profiling — Use the beam profiling setup (microscope objective, 45° mirror, 1" lens tube, digital camera) to image the beam and plot $w(z)$. from the geometrical limit, through the Rayleigh range around the beam-waist, and into the geometrical limit on the other side.

Obvious questions:

- how will you calibrate your camera, to know real sizes in your image?
- focused beams are much more intense than spread-out beams. What is the best way to manage this change of a few orders of magnitude, on an 8-bit or 10-bit digitizing camera?

Radius of curvature — As beams expand, the radius of curvature of the phase-fronts get larger and larger – they become flatter, and in the limit are plane-waves. However, gaussian beams have flat wave-fronts at their beam-waist. You can find the beam-waist by mounting a flat mirror on the XYZ translation stage, and finding the position that sends the beam backward on itself, exactly as it came in: same direction, same size at each point. This is *retrocollimation*.



Mounted steel balls: spherical reflectors.
Consider also the mounted glass balls, also high-quality (better than λ/10) spherical-surface glass lenses of different focal length (provided).

At points away from the waist, the beam has a spherical radius of curvature *R(z)*. For different spherical reflecting-surfaces, find the location at which the beam is precisely retrocollimated. Bear in mind that the wavefronts are flat at the beam-waist, and in the limit of z large they again approach flat. Therefore any for any smaller radius acquired, except the unique beam-waist, that same radius must appear again at some other location in the gaussian beam.

Plot *R(z)* from your measurements. Plot this on a double-ordinate graph, to put your *R(z)*alongside your *w(z)*characterization.

Obvious questions:

- you will replace the flat mirror with a steel ball, and that with a glass lens spherical surface. How will you know the z-value for each? The flat mirror lets you find z=0 for the gaussian beam-waist, but how can you reference that position in space once you remove the flat mirror, and keep it to use with the balls?
- retrocollimation sends a beam exactly back on itself. How then can you visualize the returning beam, if a card or camera to view it would block the returning beam?
- How will you determine your error-bars, your confidence about the best position z to fit a given R of a ball or lens?

GENERAL TIPS

- It's generally helpful first to establish the optical axis from laser to XYZ stage, perhaps parallel to holes in the breadboards, and then add optics you want while preserving this axis
- The ABCD matrix method for ray-tracing will let you calculate beams through optics in the geometrical limit. The same matrices, applied differently, let you do the same optics but for gaussian beams. See the Appendix attached, for both ray-tracing and gaussian-beam methods
- The open-source code ImageJ available online for free for many platforms from the US National Institutes of Health, provides excellent tools for image analysis, including plotting profiles across images of gaussian beams, saving profiles to text files for curve-fitting, and enhanced visualisation like 3D plots and false-colour images.

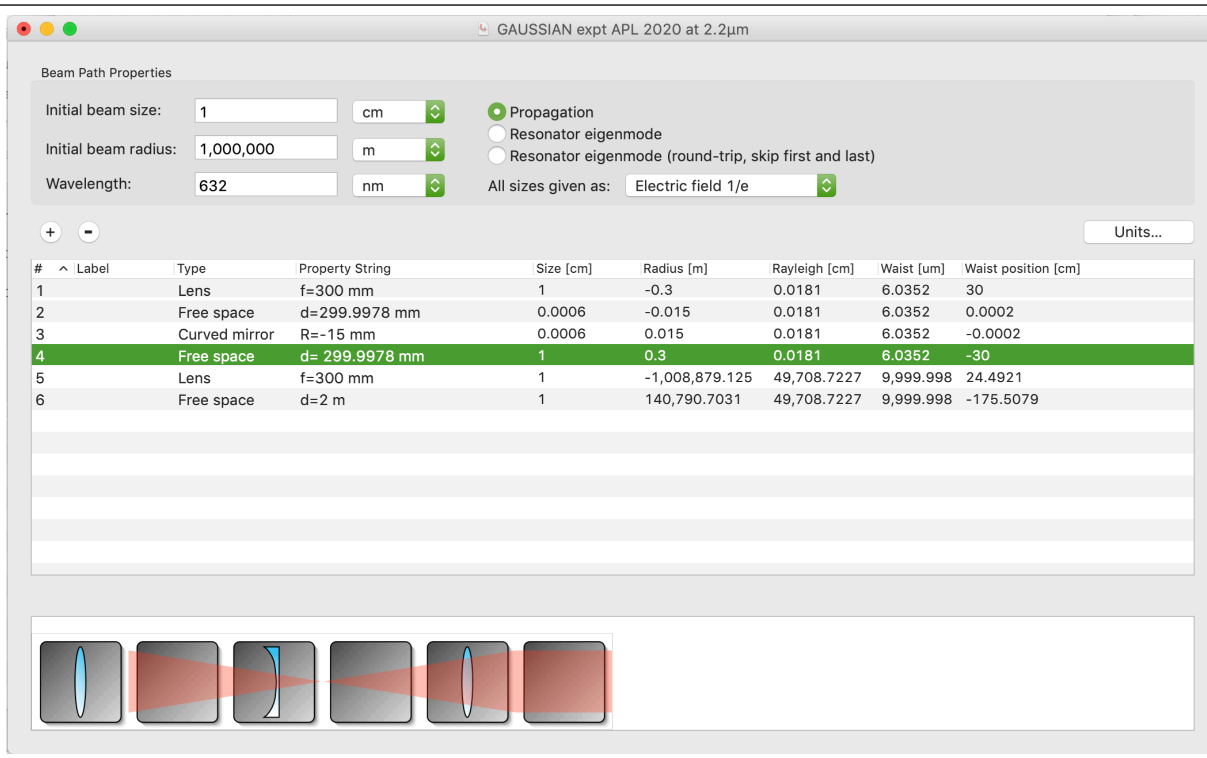
- Videos teaching lab skills for optics are available, see the APL webpage for this experiment for links.
- Index-cards, business cards, of different sizes can be very helpful in alignment — punch a hole through the card with a ballpoint pen, fine scissors, hole-punch, etc..

MODELLING AND THEORY

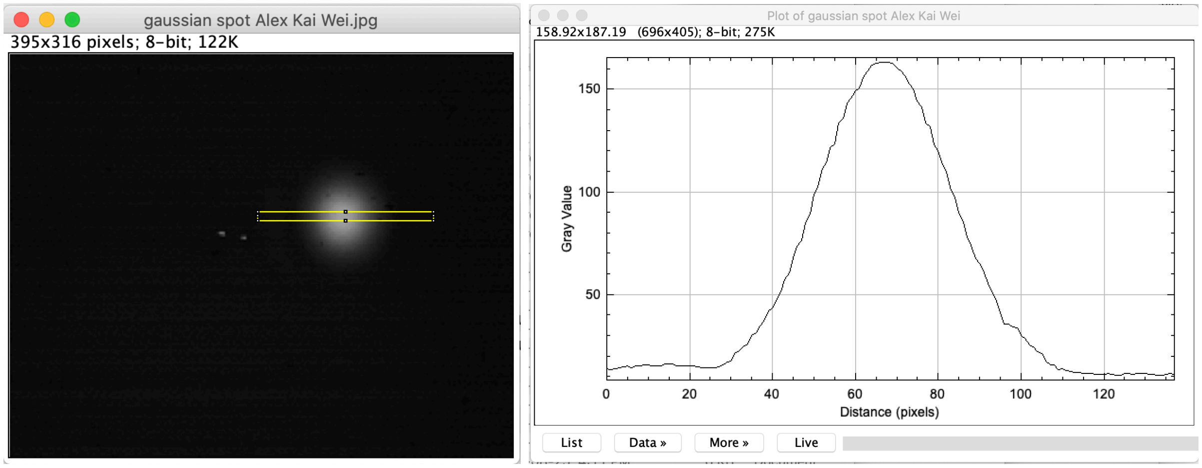
For the ray-trace picture, show that the ABCD matrices for your setup produce exactly one location where the ball might reflect and collimate the beam going backwards.

Repeat now for the same matrix elements, but using the fractional-linear transformation, the way that the $q(z)$ gaussian beam parameter transforms at each point in your setup and leads to two different places where any ball may be placed. Is it true that this is possible for any size of ball?

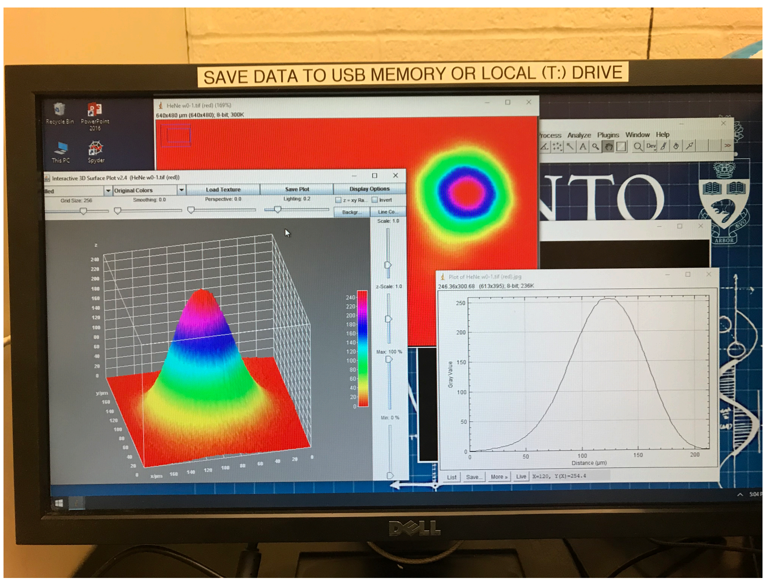
Look online for apps that will let you lay out an optical system and transform a Gaussian beam, element by element, to simplify the modelling you need, and to reproduce your results.



ABCD modelling routine from Professor Daniel Côté at Université Laval, a past graduate student in Physics at the University of Toronto. This code v3.1for Mac is for OSX 10.1 (buggy in 10.14), but similar software exists online in open-source and commercial software, as well as Python, MATLAB and Mathematica libraries.



Example of image analysis of gaussian beam, using profiling in ImageJ. Further information available in separate instructions for Beam-Profiling using ImageJ.



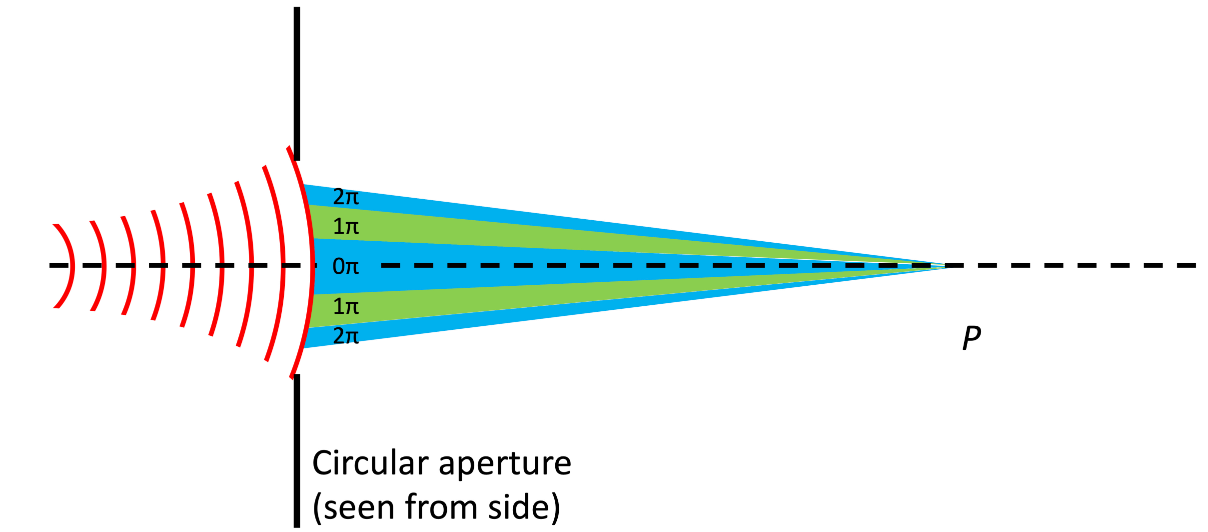
Example of 3D analysis of gaussian beam, using ImageJ

OPTIONAL RELATED QUESTIONS ON DIFFRACTION

NEAR-FIELD DIFFRACTION AND THE SPOT OF ARAGO

In the far-field limit (Fraunhofer diffraction), distances are such that from source to aperture, and onward from aperture to observation point, wavefronts are plane waves. Summing up all these plane waves $\exp\{i\vec{k}\cdot\vec{x}\}$, leads to a Fourier transform of the amplitude at the aperture.

In the near-field limit (Fresnel diffraction), wavefronts from source to aperture, and aperture to observation point, are curved (i.e., spherical wavelets, of which gaussian beams are the paraxial approximation). Consider a circular aperture, centred on the optical axis, which can be increased in size (e.g., an iris diaphragm as inside a DSLR camera). Because the spherical wavefronts at the aperture are domed, and let's say convex domes, the part of the wavefront which is on axis is closer to some distant on-axis point that the parts of the wavefront at the edge of the aperture. Therefore there are imaginary annular *zones* on the dome, for which the propagation phase difference can increase by an *average* amount π for each subsequently larger zone. These are called Fresnel zones.



A spherically curved wavefront, at a circular aperture centred on-axis, makes a dome for surfaces of constant phase. All points on this dome start off in phase with each other. However, different points travel different distances when travelling to the axis. This defines annular zones, on the dome – rings – each ring travels to the axis with roughly the same phase. But the next ring will have an average phase-change of π: light from that ring interferes destructively with light from the first ring. Thus in the figure above, blue zones add nearly constructively with each other at the observation point on axis. An aperture which passes only the innermost blue zone above will give light at P, but a slightly larger aperture that admits the green zone as well will produce *less* light at P.

Using the expanded HeNe beam, show empirically that for the right setup, you can have a beam pass through an open circular aperture and yet have a completely dark spot in the centre of the apertured beam. Make the calculations necessary to explain what you see.

Special hole-punches are provided that create holes of different shapes. Using the expanded HeNe beam, record the intensity pattern at different distances from the apertures. At what distance does the pattern become that expected for Fraunhofer diffraction? What happens beyond that point?

Obvious questions:

- for the circular aperture, like an iris diaphragm, what happens at the observation point as you steadily increase the diameter?
- *distance* as a bare number means nothing, since one can choose different measurement units. However, the distance of one characteristic of your setup compared to another characteristic distance is dimensionless, and is about relationship. All meaning in physics comes from relationship. What is/are the relationship(s) that change(s), to go from one diffraction regime to another? Therefore, what's another way to change the same natural relationship, besides changing the aperture size?

Tips:

- ImageJ lets you take an image you have recorded and produce from the intensity pattern an FFT which will correspond to Fraunhofer diffraction for perfect plane-wave illumination. In fact, you want first to convert to amplitude, then FFT, then afterwards convert back to intensity.
- When you use an imaging system, any focussed point in the image plane has collected rays that all travel the same optical distance from their one origin point in the object plane (Fermat's principle). This means there's no change in relative phase for the wavelets, no change in interference, no diffraction. If you use a lens and make an image of your different apertures, you'll be able to collect diffraction patterns that start with in effect *zero propagation distance* from the aperture. As you back off the lens and camera, you image planes in space which are located at successively distant propagation distances, starting from zero.

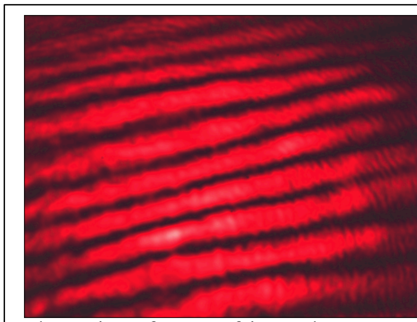
PINHOLE INTERFEROMETER (UNDER DEVELOPMENT, ASK APL COORDINATOR)

A relatively simple interferometer can let you precisely determine $R(z)$ for gaussian beams. You've seen that any wavefront passing through a small aperture will spread into spherical waves on the other side. This can make a reference beam, to be interfered against the original laser beam – for instance, a laser beam with flat wavefronts will interfere with the spherical waves to make a series of rings.

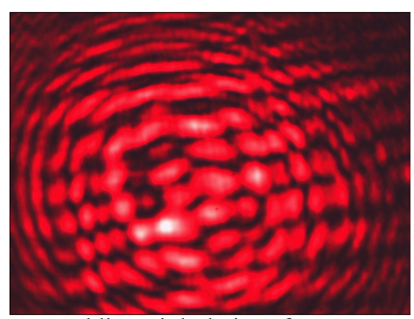
A thin gelatin-film plastic neutral-density filter (Kodak Wratten filter) has been prepared with a small hole (~30µm diameter) drilled through it with a focused pulsed laser. The attenuation value (about ND3) of the filter is chosen so that the intensity of the whole laser beam passed through, directly, ends up about the same as the intensity of the spherical wave that has expanded through the small hole (about the ratio of areas, expanded beam to drilled hole).

Various variations of this scheme are possible to explore.

Theory and Application of Point-Diffraction Interferometers
R. N. Smartt and W. H. Steel 1975 Jpn. J. Appl. Phys. 14 351



Planar interference fringes between two slightly angled HeNe beams



Now adding pinhole interferometer, to measure wavefront curvature

Images from the experiment, in different iterations

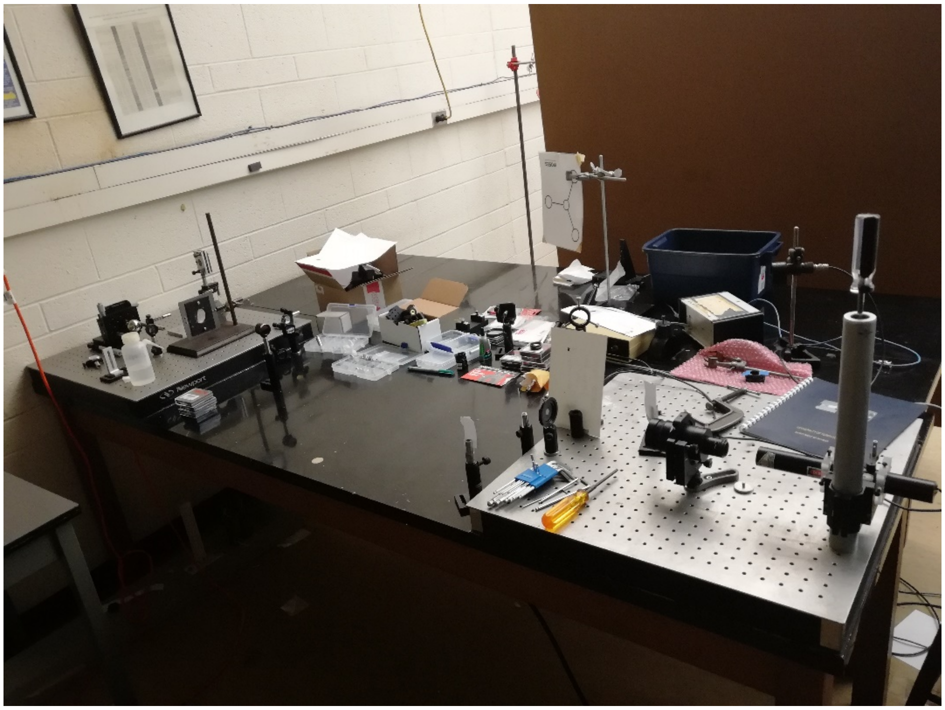


Fig. 1. An image of the overall original setup for the Gaussian Beam Experiment

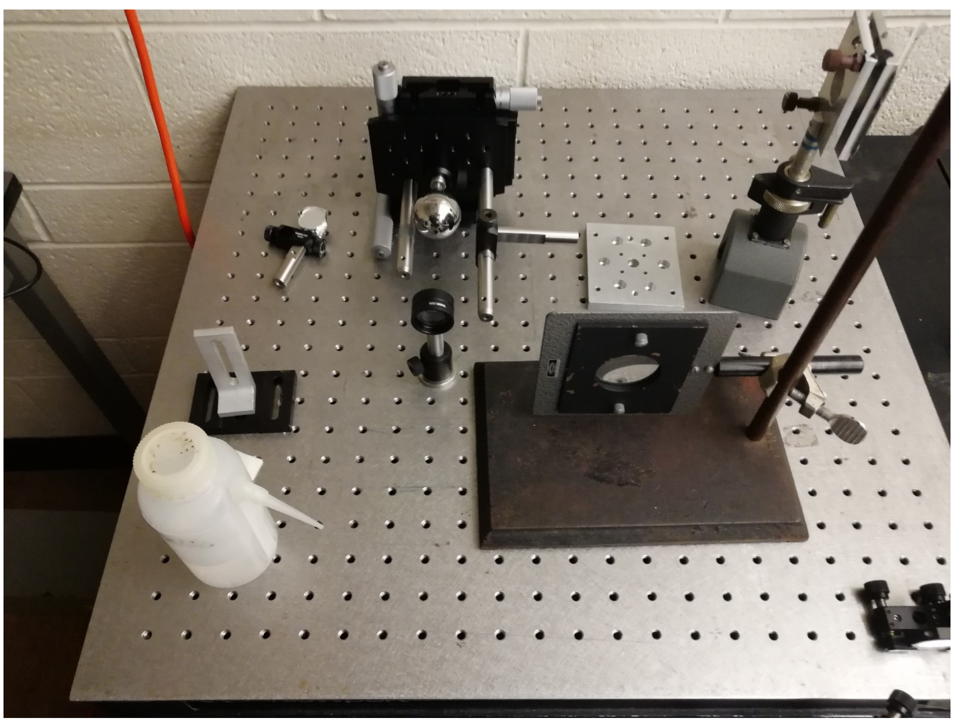


Fig. 2. An image of the steel balls attached to the translation stage setup with various lenses. Three separate translation stages allow for full control in the x,y,z planes at the micro scale.

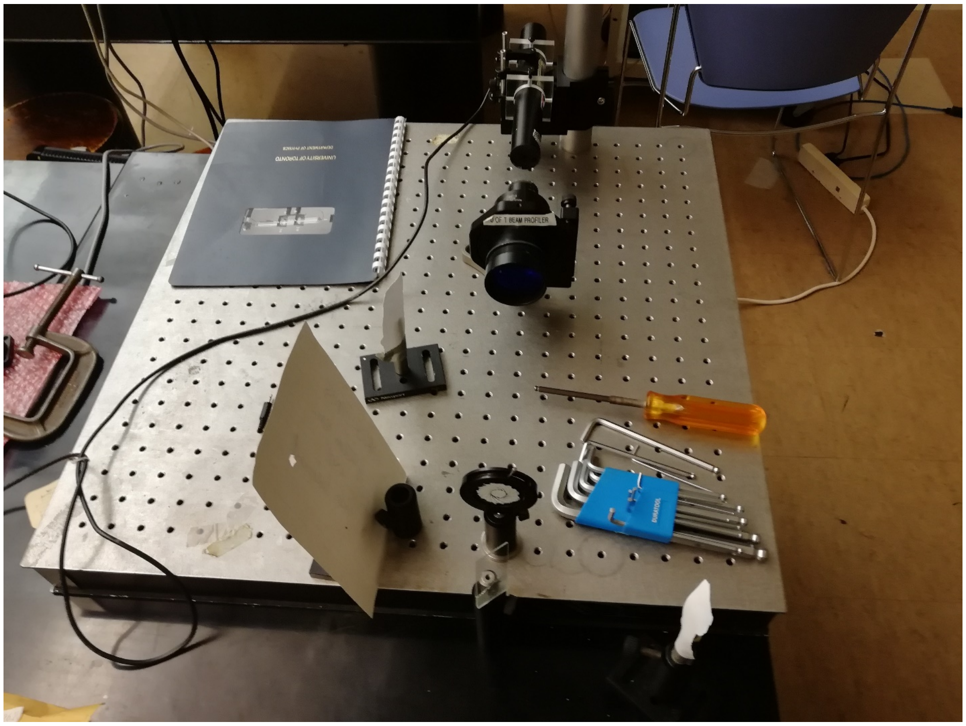


Fig. 3. An image of the HeNe laser into a beam expander and then into an iris. The painted iris allows for users to determine the position of the collimated beam, with the beamsplitter (microscope slide) in front of the iris, also contributing to user's ease in determining the collimated beam size.

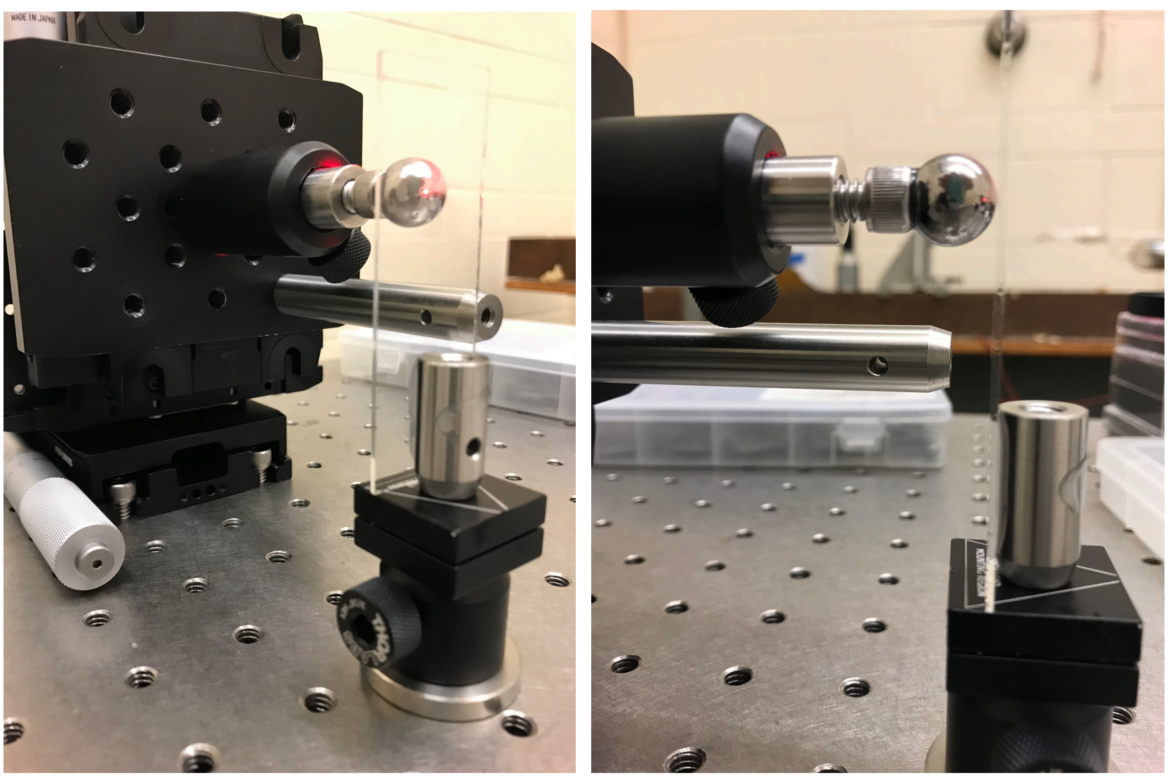


Fig. 4. One method for referencing absolute z-position in space: special mount with microscope slide – the reflected laser beam is very sensitive to the moment the ball touches the glass.

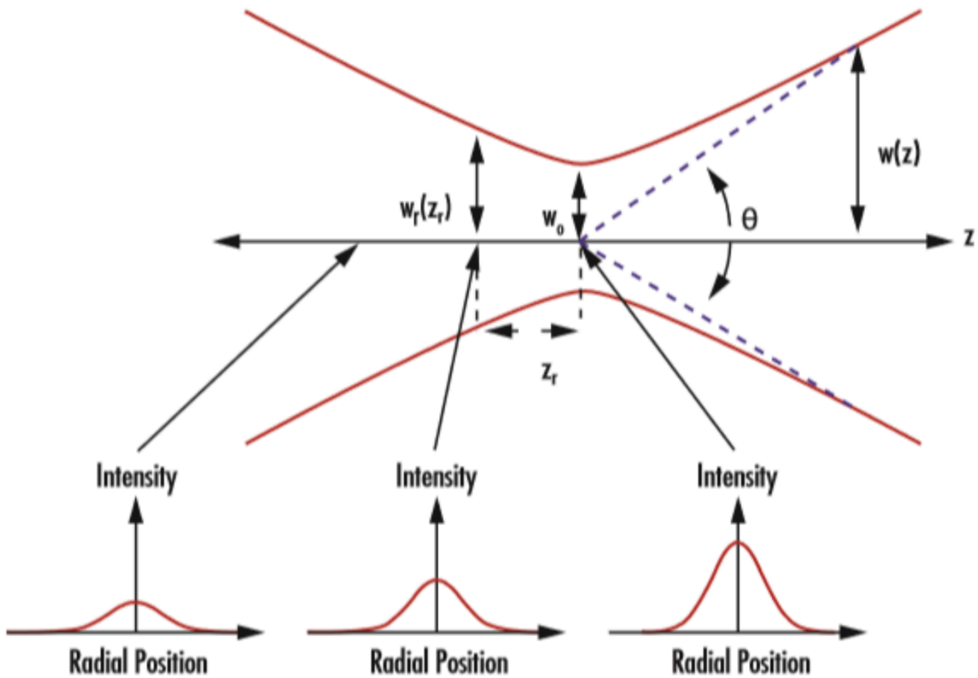


Fig. 5. Schematic of a gaussian beam: This figure makes clear that the hyperboloids which trace the 1/e-max surfaces are not surfaces of constant intensity, but just mark the width of the diminishing profile.

CHAPTER 6

Matrix Methods in Paraxial Optics

Our life is frittered away by detail...
Simplify, simplify
H.D. Thoreau

The previous chapter has indicated the usefulness of ray optics in analyzing the imaging properties of systems. Within the paraxial approximation particularly simple results were obtained for distinct optical elements which were homogeneous in a plane perpendicular to the optical axis. The problem of treating rays which pass through several optical elements could in principle be carried out with much algebraic manipulation. In this chapter we introduce a matrix method which accomplishes the same task but in a much more straightforward fashion. Transformation of rays on passage through a complex optical system then simply reduces to the multiplication of matrices associated with each optical element. At the end of the chapter we consider applying these methods to telescopes and microscopes.

6.1. Optical Rays and Transformations

Recall that a ray is defined, in the limit of geometrical optics, where $\lambda \to 0$, to be a beam of light of infinitesimal transverse extent. A ray, as a line in space, is completely defined by its distance from a given axis and the slope of that line relative to the axis. In the paraxial approximation, the slope is equivalent to the angle the ray makes with the optic axis. This is indicated in figure 6.1.1.

In the *paraxial approximation* the slope is equivalent to the angle between the ray and the axis. If we take the axis to be the z-axis, then the ray, at a certain reference plane $z = z_1$, is completely defined by $r(z_1)$which specifies the distance of the ray from the axis and $r'(z_1) = dr/dz|(z = z_1) = \theta(z_1)$. These together form what is called the *ray vector*

$$\vec{R} = \begin{bmatrix} r(z_1) \\ \left.\frac{dr}{dz}\right|_{z=z_1} = \theta(z_1) \end{bmatrix}.$$

Note that the two components of the vector do not have the same dimensionality.

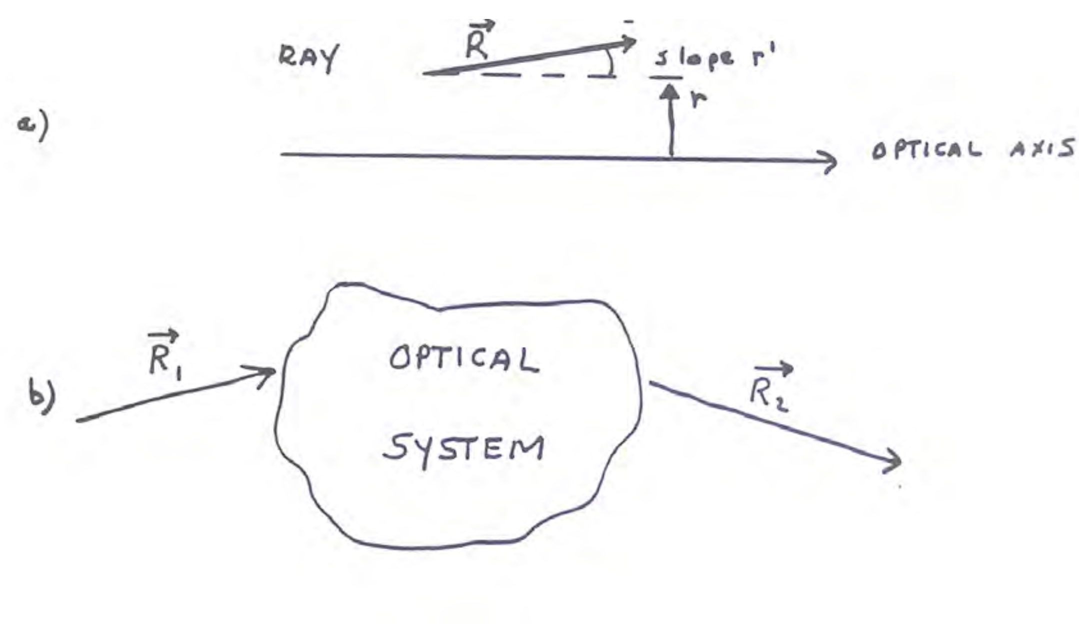


FIGURE 6.1.1. Transformation of a ray by an optical system.

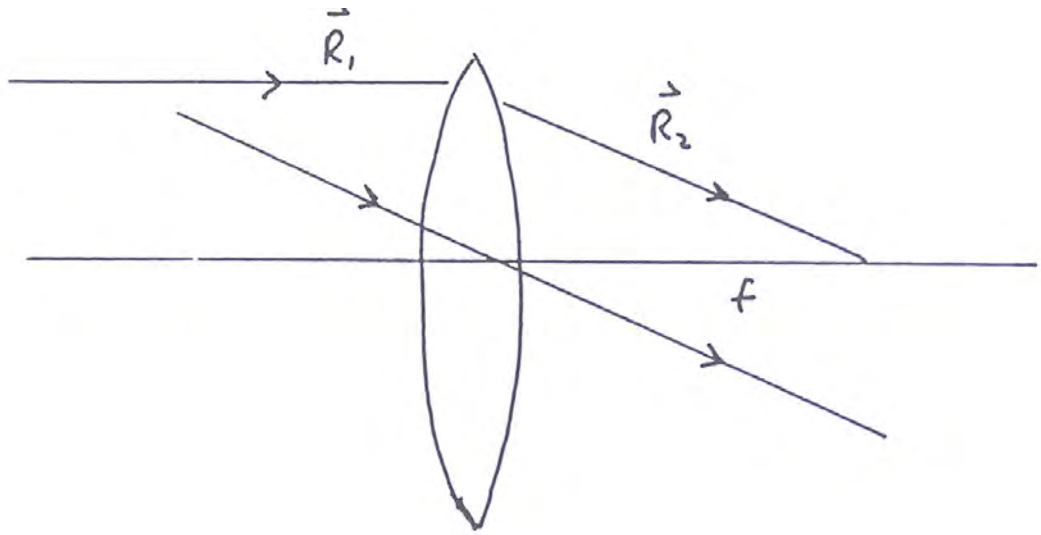


FIGURE 6.1.2. Ray transformation by a simple lens.

Upon passage through an optical system the ray vector is transformed into a new ray vector as indicated in figure 6.1.1. One expects, for a linear optical system, that there is a simple functional relationship between the incident and emerging ray vector components so that if $\vec{R}_1$ and $\vec{R}_2$ are the vectors describing the incident and emerging rays we would have

$$r_2 = f(r_1, \theta_1) \quad \text{and} \quad \theta_2 = g(r_1, \theta_1)$$

where f and g are two undetermined functions. For an imaging system (one which does not cause any transverse distortion in an incident light distribution) the transformation laws must be linear so that

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

or

$$\vec{R}_2 = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot \vec{R}_1 \qquad (6.1.1)$$

where the A , B, C and D are parameters to be determined for a particular system. They define a so-called *ABCD matrix*.

For a thin lens of focal length f we can determine the transformation matrix as follows. With respect to figure 6.1.2 it is clear that the transformation matrix must satisfy the following conditions:

1) For a thin lens we must have that $r_1 = r_2$ for all θ_1. This implies $A = 1$ and $B = 0$.

2) For a ray which passes through the center of the lens ($r_1 = 0$) we must have that the slope (angle) doesn't change so that $D = 1$.

3) An incident ray which is parallel to the axis of the lens must pass through the focal point, by definition of the focal point. Hence for $\theta_1 = 0$, as seen in the diagram, we must have

$$\theta_2' = -\frac{r_1}{f}$$

or

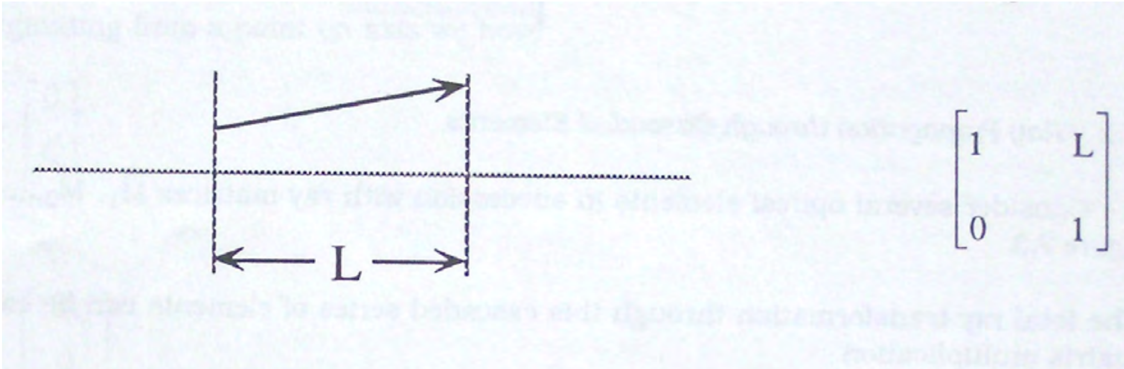
$$C = -\frac{1}{f}.$$

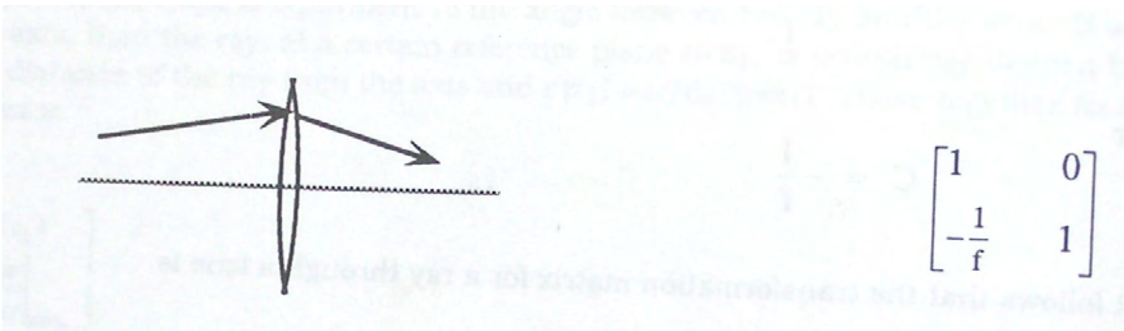
It follows that the *transformation matrix* for a ray through a lens is

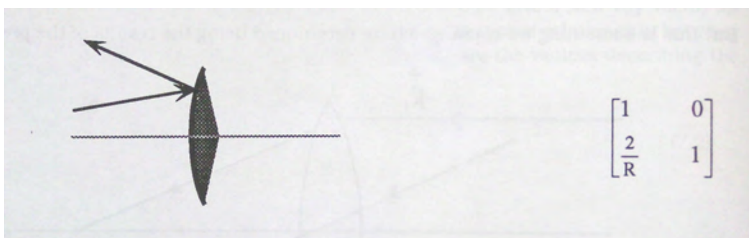
$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -f^{-1} & 1 \end{bmatrix}.$$

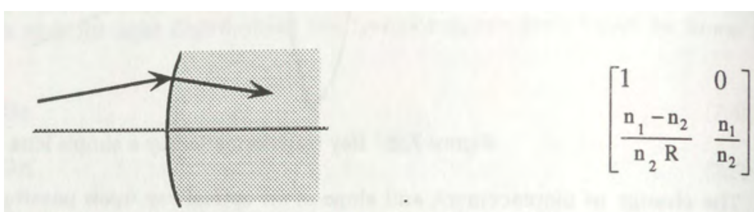
But this is something we could also have determined using the results of the previous chapter.

The change in displacement and slope of an optical ray upon passing through a wide variety of simple optical elements can be written in the same general form as equation 6.1.1. The matrix derived for a particular optical element is known as the *ray matrix* for the element. Most of these are derived simply by considering the results of









the previous chapter. In all cases the matrices have a determinant which is equal to n_1/n_2 where n_1 and n_2 are the refractive indices at the input and output planes.

We now proceed to list all the common ray matrices within the paraxial approximation.

a) Free space propagation (virtual rays propagate with negative distances, not to be confused with left & right):

b) Thin lens, focal length *f:*

c) Spherical mirror, radius R, (recall that virtual rays subsequently propagate with negative distance); $R > 0$ for convex incidence:

d) Curved dielectric interface: $R > 0$ for convex-surface incidence:

e) Refraction at a plane interface (from medium n_1 to medium n_2):

6.2. Ray Propagation Through Cascaded Elements

Consider several optical elements in succession with ray matrices $\overleftrightarrow{M}_1, \overleftrightarrow{M}_2, \ldots \overleftrightarrow{M}_n$.. .as shown in figure 6.2.1.

The total ray transformation through this cascaded series of elements can be calculated from simple matrix multiplication

$$\vec{R}_2 = \overleftrightarrow{M}_2 \cdot \vec{R}_1 = \overleftrightarrow{M}_2 \cdot \overleftrightarrow{M}_1 \cdot \vec{R}_0$$

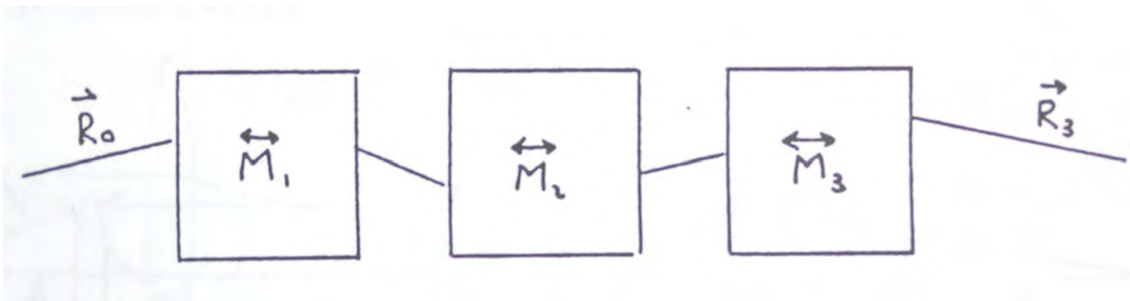


FIGURE 6.2.1. Ray matrix systems in cascade.

etc. so that the general result is

$$\vec{R}_n = \overleftrightarrow{M}_n \cdot \overleftrightarrow{M}_{n-1} ... \overleftrightarrow{M}_1 \cdot \vec{R}_1 = \overleftrightarrow{M}_{tot} \cdot \vec{R}_0$$

and the overall ray transformation matrix is given by

$$\overleftrightarrow{M}_{tot} = \overleftrightarrow{M}_n \cdot \overleftrightarrow{M}_{n-1} ... \overleftrightarrow{M}_1.$$

Note the order in which the matrices are multiplied and how this is related to the order in which the ray encounters the different optical elements.

As an illustration of how the matrix method works let us consider the following problem. A point source is located on axis 20 cm in front of a pair of lenses with focal length of 10 and -10 cm respectively and separated by 10 cm. Does the lens system lead to the formation of a real image on the far side of the second lens?

For a real image to form light must pass through the axis on the far side of the second lens for an arbitrary initial ray. This means it would have to be possible to find a location where $r = 0$ after the second lens. If it occurs let us call this distance z behind the second lens. The overall transformation matrix for the problem is seen to be

$$\overleftrightarrow{M}_{tot} = \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0.1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 10 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ -0.1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 20 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -0.1z & 10 \\ -0.1 & 0 \end{bmatrix}.$$

For an initial ray originating from a point on axis we have

$$\vec{R}_0 = \begin{bmatrix} 0 \\ \theta_0 \end{bmatrix}$$

and

$$\vec{R}_f = \begin{bmatrix} 10r_0' \\ 0 \end{bmatrix}.$$

Since this ray is independent of z and never has an $r = 0$, we conclude that it does not correspond to a real image (only the ray launched directly along the axis ($\theta_0 = 0$) remains on axis.

6.3. Telescopes and Microscopes

We close this short chapter with a discussion of two very important optical systems, the microscope and the telescope. A telescope is an instrument which images and magnifies objects far away while a microscope magnifies objects which are very small.

An astronomical *telescope*, for example, focuses for infinity and typically has a large light gathering optical element which may be a mirror or lens. The size of this element determines the brightness of the image since its area directly determines the amount of light gathered. As we shall see later in the chapter on Fraunhofer diffraction, the size of the "objective" also determines the ability of the telescope to resolve closely spaced distant objects. Figure 6.3.1illustrates a specific type of telescope which is known as the Kepler astronomical refractor.

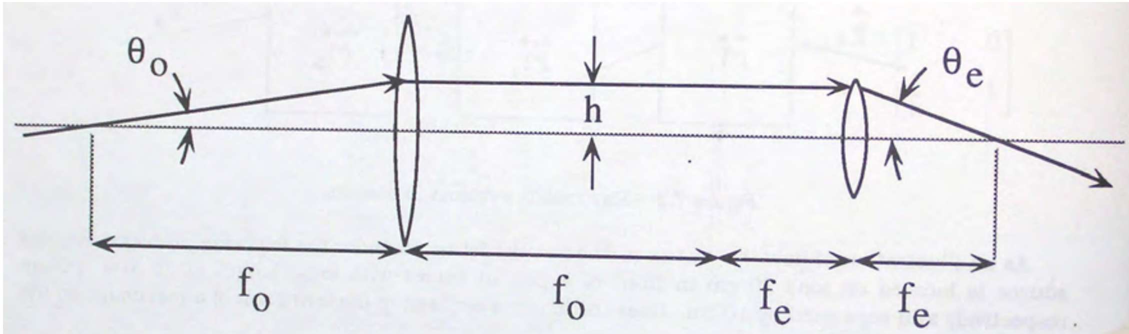


FIGURE 6.3.1. Astronomical telescope

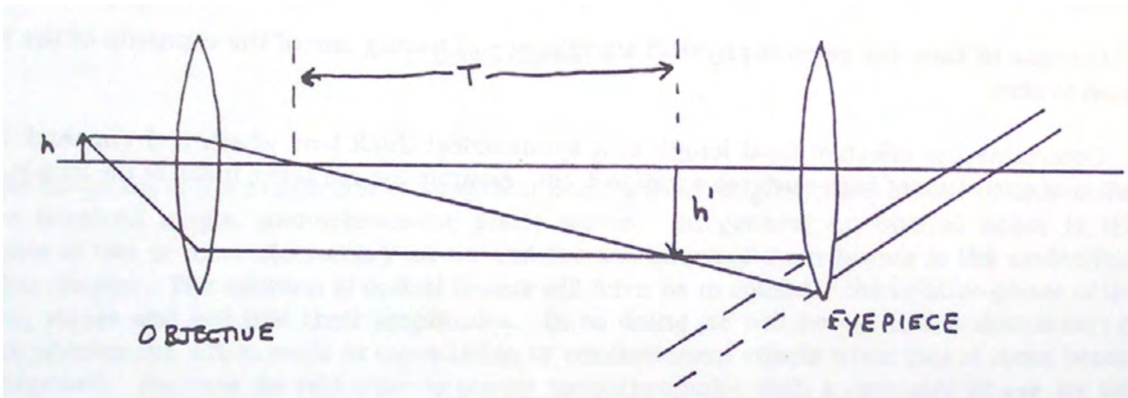


FIGURE 6.3.2. Magnification of a compound microscope.

The lenses are spaced so that the second focus of the first lens coincides with the first focus of the second lens. One of the most important characteristics of a telescope is its magnification which can be considered to be the ratio of the angular size of the image to the angular size of the object. The magnification can be determined easily using ray matrix techniques but can be determined more trivially as follows. From the figure we see that this *magnification* is given as

$$M = \frac{\theta_e}{\theta_o} = \frac{\left(\frac{h}{f_e}\right)}{\left(\frac{h}{f_o}\right)} = \frac{f_o}{f_e}.$$

We now turn our attention to microscopes. One can form a *simple microscope* from one lens of particularly short focal length but we shall not discuss this trivial case here. *Compound microscopes* like telescopes come in many different configurations and in simplest form, like a telescope, they consist of two lenses. The big difference, of course, is that telescopes try to magnify things far away while microscopes try to magnify close objects. Figure 6.3.2 shows the typical lens configuration for a compound microscope.

It is seen that the objective has a very short focal length so as to generate a large image in the focal plane of the second lens which is located a "tube length", T, away from the focal plane of the objective lens. The *magnification* of the object by the objective is easily seen to be

$$M_o = \frac{T}{f_o}.$$

The eyepiece acts as a magnifier. Since for most humans the distance of most distinct vision is 25 cm, eyepieces are designed so as to yield a virtual image 25 cm away from the eye. The magnification of the eyepiece is therefore approximately given by $25/f_e$ giving as a magnification for the whole system

$$M = \frac{25T}{f_e f_o}.$$

Microscope eyepieces are usually designated by their magnification but their focal length can be determined from the formula above.

References

J.R. Meyer-Arendt, *Introduction to Classical and Modern Optics*, Prentice-Hall, 1984.
A.E. Siegman, *Lasers*, Oxford, 1987.

Problems

1. a) Show that the transformation matrix associated with passage through a spherical dielectric interface which separates a medium of refractive index n_1 from one of refractive index n_2 is

$$\begin{bmatrix} 1 & 0 \\ \frac{n_1 - n_2}{n_2 R} & \frac{n_1}{n_2} \end{bmatrix}$$

where R is the radius of curvature of the surface.

b) Show that the focal length of a thin lens made of glass of refractive index n and with surfaces possessing radius of curvature R_1 and R_2 is given by

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

if the lens is in vacuum.

2. Using the ABCD matrices determine the (in focus) transverse magnification of a telescope which has an objective of focal length f_1 separated from an eyepiece of focal length f_2.

3. An optical system when traversed by an optical ray in one direction has a certain ABCD matrix. What is the corresponding matrix if the ray enters from the opposite direction?

4. A lens guide consists of a large number of identical lenses of focal length f separated from each other by a distance d. What must the relation between d and f be such that a ray, launched into the lens system at a small angle be confined to the lens system?

5. Discuss in turn the general physical significance of having one of the elements of the ray matrix equal to zero.

6. Determine the effective focal length of a symmetrical thick lens which is 1 cm thick and has a radius of curvature of both surfaces equal to 5 cm. Assume the refractive index of the lens is 1.5.

CHAPTER 6

Matrix Methods in Paraxial Optics

Our life is frittered away by detail...
Simplify, simplify
H.D. Thoreau

The previous chapter has indicated the usefulness of ray optics in analyzing the imaging properties of systems. Within the paraxial approximation particularly simple results were obtained for distinct optical elements which were homogeneous in a plane perpendicular to the optical axis. The problem of treating rays which pass through several optical elements could in principle be carried out with much algebraic manipulation. In this chapter we introduce a matrix method which accomplishes the same task but in a much more straightforward fashion. Transformation of rays on passage through a complex optical system then simply reduces to the multiplication of matrices associated with each optical element. At the end of the chapter we consider applying these methods to telescopes and microscopes.

6.1. Optical Rays and Transformations

Recall that a ray is defined, in the limit of geometrical optics, where $\lambda \to 0$, to be a beam of light of infinitesimal transverse extent. A ray, as a line in space, is completely defined by its distance from a given axis and the slope of that line relative to the axis. In the paraxial approximation, the slope is equivalent to the angle the ray makes with the optic axis. This is indicated in figure 6.1.1.

In the *paraxial approximation* the slope is equivalent to the angle between the ray and the axis. If we take the axis to be the z-axis, then the ray, at a certain reference plane $z = z_1$, is completely defined by $r(z_1)$which specifies the distance of the ray from the axis and $r'(z_1) = dr/dz|(z = z_1) = \theta(z_1)$. These together form what is called the *ray vector*

$$\vec{R} = \begin{bmatrix} r(z_1) \\ \frac{dr}{dz}\Big|_{z=z_1} = \theta(z_1) \end{bmatrix}.$$

Note that the two components of the vector do not have the same dimensionality.

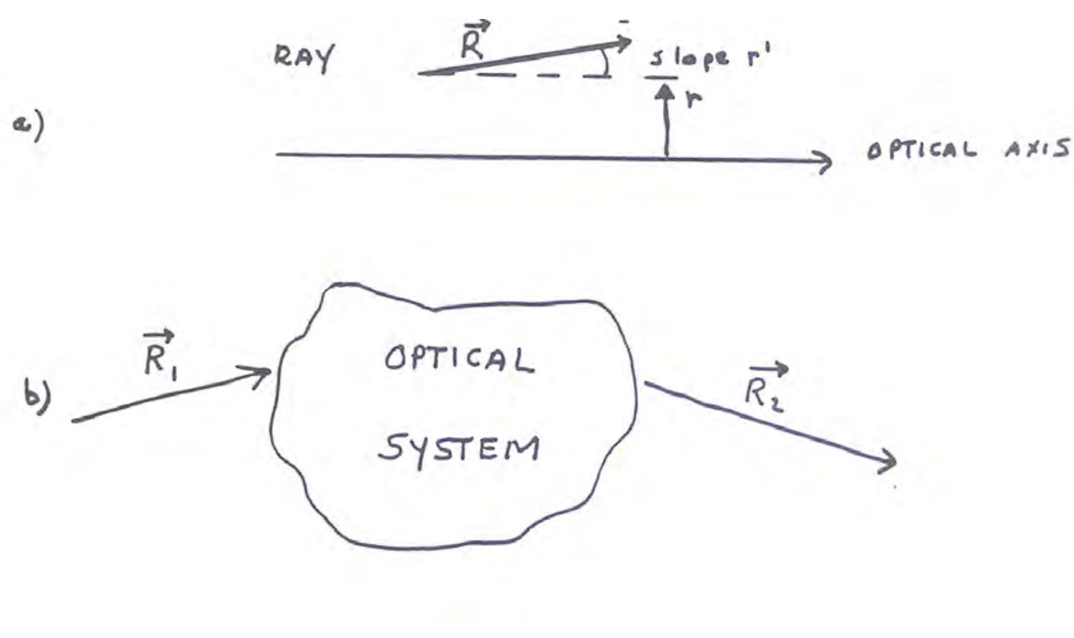


FIGURE 6.1.1. Transformation of a ray by an optical system.

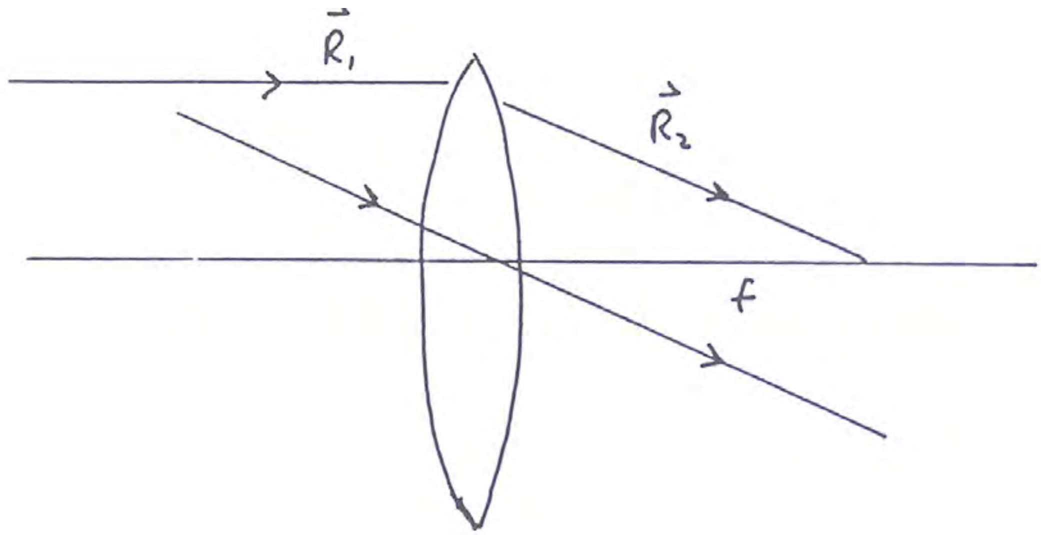


FIGURE 6.1.2. Ray transformation by a simple lens.

Upon passage through an optical system the ray vector is transformed into a new ray vector as indicated in figure 6.1.1. One expects, for a linear optical system, that there is a simple functional relationship between the incident and emerging ray vector components so that if $\vec{R}_1$ and $\vec{R}_2$ are the vectors describing the incident and emerging rays we would have

$$r_2 = f(r_1, \theta_1) \quad \text{and} \quad \theta_2 = g(r_1, \theta_1)$$

where f and g are two undetermined functions. For an imaging system (one which does not cause any transverse distortion in an incident light distribution) the transformation laws must be linear so that

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

or

$$\vec{R}_2 = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot \vec{R}_1 \tag{6.1.1}$$

where the A , B, C and D are parameters to be determined for a particular system. They define a so-called *ABCD matrix*.

For a thin lens of focal length f we can determine the transformation matrix as follows. With respect to figure 6.1.2 it is clear that the transformation matrix must satisfy the following conditions:

1) For a thin lens we must have that $r_1 = r_2$ for all θ_1. This implies $A = 1$ and $B = 0$.

2) For a ray which passes through the center of the lens ($r_1 = 0$) we must have that the slope (angle) doesn't change so that $D = 1$.

3) An incident ray which is parallel to the axis of the lens must pass through the focal point, by definition of the focal point. Hence for $\theta_1 = 0$, as seen in the diagram, we must have

$$\theta_2' = -\frac{r_1}{f}$$

or

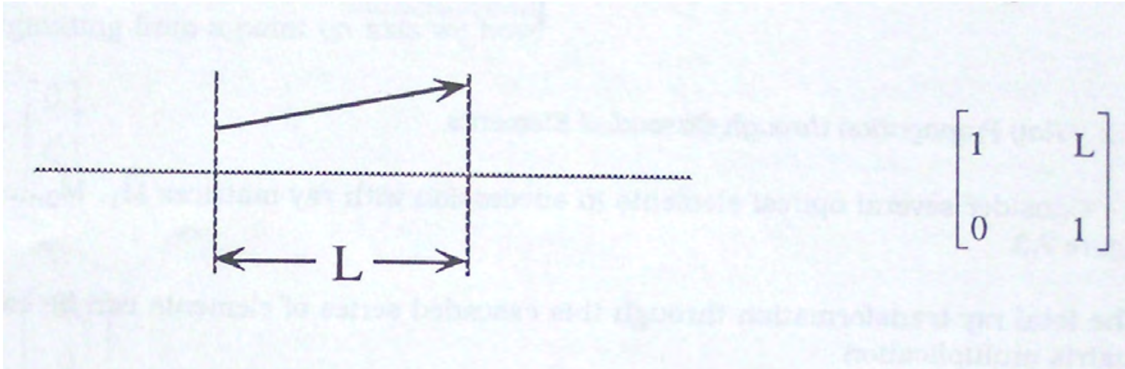
$$C = -\frac{1}{f}.$$

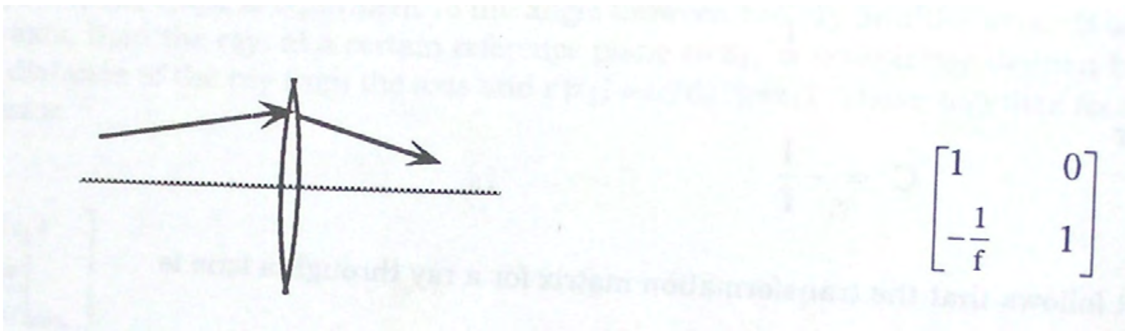
It follows that the *transformation matrix* for a ray through a lens is

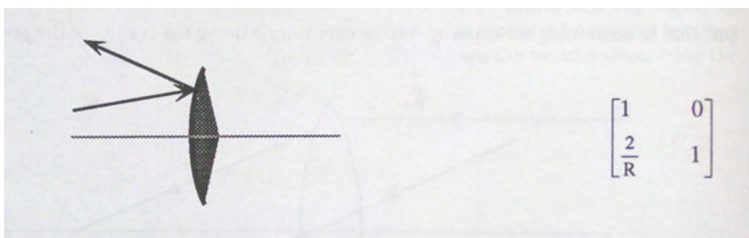
$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -f^{-1} & 1 \end{bmatrix}.$$

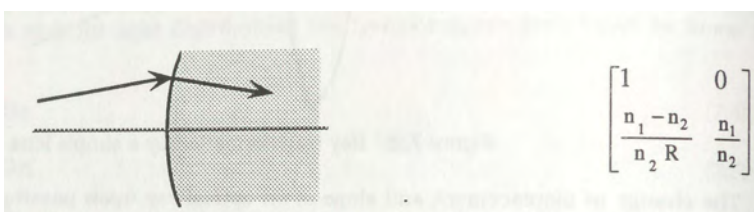
But this is something we could also have determined using the results of the previous chapter.

The change in displacement and slope of an optical ray upon passing through a wide variety of simple optical elements can be written in the same general form as equation 6.1.1. The matrix derived for a particular optical element is known as the *ray matrix* for the element. Most of these are derived simply by considering the results of









the previous chapter. In all cases the matrices have a determinant which is equal to n_1/n_2 where n_1 and n_2 are the refractive indices at the input and output planes.

We now proceed to list all the common ray matrices within the paraxial approximation.

a) Free space propagation (virtual rays propagate with negative distances, not to be confused with left & right):

b) Thin lens, focal length *f:*

c) Spherical mirror, radius R, (recall that virtual rays subsequently propagate with negative distance); $R > 0$ for convex incidence:

d) Curved dielectric interface: $R > 0$ for convex-surface incidence:

e) Refraction at a plane interface (from medium n_1 to medium n_2):

6.2. Ray Propagation Through Cascaded Elements

Consider several optical elements in succession with ray matrices $\overleftrightarrow{M}_1, \overleftrightarrow{M}_2, \ldots \overleftrightarrow{M}_n$.. .as shown in figure 6.2.1.

The total ray transformation through this cascaded series of elements can be calculated from simple matrix multiplication

$$\vec{R}_2 = \overleftrightarrow{M}_2 \cdot \vec{R}_1 = \overleftrightarrow{M}_2 \cdot \overleftrightarrow{M}_1 \cdot \vec{R}_0$$

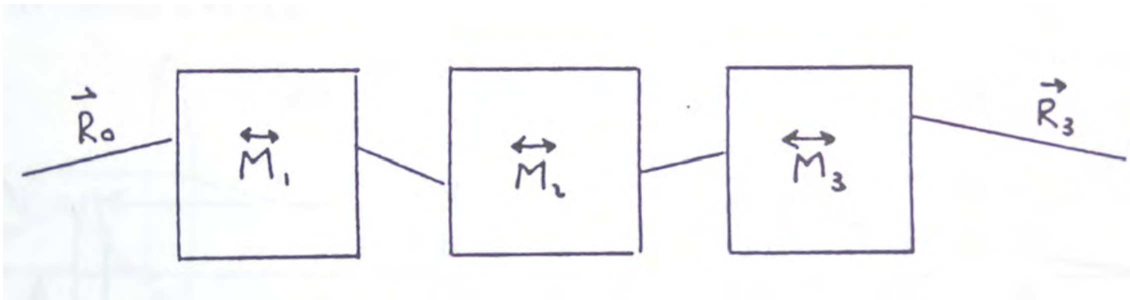


FIGURE 6.2.1. Ray matrix systems in cascade.

etc. so that the general result is

$$\vec{R}_n = \overleftrightarrow{M}_n \cdot \overleftrightarrow{M}_{n-1} ... \overleftrightarrow{M}_1 \cdot \vec{R}_1 = \overleftrightarrow{M}_{tot} \cdot \vec{R}_0$$

and the overall ray transformation matrix is given by

$$\overleftrightarrow{M}_{tot} = \overleftrightarrow{M}_n \cdot \overleftrightarrow{M}_{n-1} ... \overleftrightarrow{M}_1.$$

Note the order in which the matrices are multiplied and how this is related to the order in which the ray encounters the different optical elements.

As an illustration of how the matrix method works let us consider the following problem. A point source is located on axis 20 cm in front of a pair of lenses with focal length of 10 and -10 cm respectively and separated by 10 cm. Does the lens system lead to the formation of a real image on the far side of the second lens?

For a real image to form light must pass through the axis on the far side of the second lens for an arbitrary initial ray. This means it would have to be possible to find a location where $r = 0$ after the second lens. If it occurs let us call this distance z behind the second lens. The overall transformation matrix for the problem is seen to be

$$\overleftrightarrow{M}_{tot} = \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0.1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 10 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ -0.1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 20 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -0.1z & 10 \\ -0.1 & 0 \end{bmatrix}.$$

For an initial ray originating from a point on axis we have

$$\vec{R}_0 = \begin{bmatrix} 0 \\ \theta_0 \end{bmatrix}$$

and

$$\vec{R}_f = \begin{bmatrix} 10r_0' \\ 0 \end{bmatrix}.$$

Since this ray is independent of z and never has an $r = 0$, we conclude that it does not correspond to a real image (only the ray launched directly along the axis ($\theta_0 = 0$) remains on axis.

6.3. Telescopes and Microscopes

We close this short chapter with a discussion of two very important optical systems, the microscope and the telescope. A telescope is an instrument which images and magnifies objects far away while a microscope magnifies objects which are very small.

An astronomical *telescope*, for example, focuses for infinity and typically has a large light gathering optical element which may be a mirror or lens. The size of this element determines the brightness of the image since its area directly determines the amount of light gathered. As we shall see later in the chapter on Fraunhofer diffraction, the size of the "objective" also determines the ability of the telescope to resolve closely spaced distant objects. Figure 6.3.1illustrates a specific type of telescope which is known as the Kepler astronomical refractor.

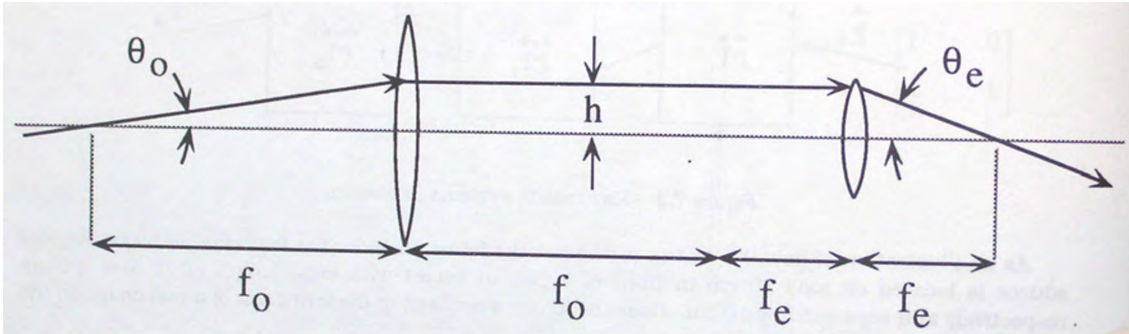


FIGURE 6.3.1. Astronomical telescope

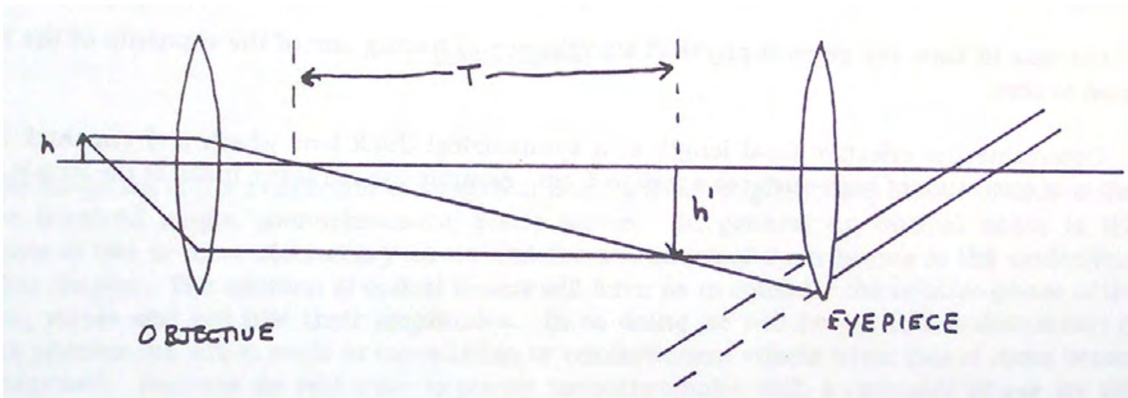


FIGURE 6.3.2. Magnification of a compound microscope.

The lenses are spaced so that the second focus of the first lens coincides with the first focus of the second lens. One of the most important characteristics of a telescope is its magnification which can be considered to be the ratio of the angular size of the image to the angular size of the object. The magnification can be determined easily using ray matrix techniques but can be determined more trivially as follows. From the figure we see that this *magnification* is given as

$$M = \frac{\theta_e}{\theta_o} = \frac{\left(\frac{h}{f_e}\right)}{\left(\frac{h}{f_o}\right)} = \frac{f_o}{f_e}.$$

We now turn our attention to microscopes. One can form a *simple microscope* from one lens of particularly short focal length but we shall not discuss this trivial case here. *Compound microscopes* like telescopes come in many different configurations and in simplest form, like a telescope, they consist of two lenses. The big difference, of course, is that telescopes try to magnify things far away while microscopes try to magnify close objects. Figure 6.3.2 shows the typical lens configuration for a compound microscope.

It is seen that the objective has a very short focal length so as to generate a large image in the focal plane of the second lens which is located a "tube length", T, away from the focal plane of the objective lens. The *magnification* of the object by the objective is easily seen to be

$$M_o = \frac{T}{f_o}.$$

The eyepiece acts as a magnifier. Since for most humans the distance of most distinct vision is 25 cm, eyepieces are designed so as to yield a virtual image 25 cm away from the eye. The magnification of the eyepiece is therefore approximately given by $25/f_e$ giving as a magnification for the whole system

$$M = \frac{25T}{f_e f_o}.$$

Microscope eyepieces are usually designated by their magnification but their focal length can be determined from the formula above.

References

J.R. Meyer-Arendt, *Introduction to Classical and Modern Optics*, Prentice-Hall, 1984.
A.E. Siegman, *Lasers*, Oxford, 1987.

Problems

1. a) Show that the transformation matrix associated with passage through a spherical dielectric interface which separates a medium of refractive index n_1 from one of refractive index n_2 is

$$\begin{bmatrix} 1 & 0 \\ \frac{n_1-n_2}{n_2R} & \frac{n_1}{n_2} \end{bmatrix}$$

where R is the radius of curvature of the surface.

b) Show that the focal length of a thin lens made of glass of refractive index n and with surfaces possessing radius of curvature R_1 and R_2 is given by

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

if the lens is in vacuum.

2. Using the ABCD matrices determine the (in focus) transverse magnification of a telescope which has an objective of focal length f_1 separated from an eyepiece of focal length f_2.

3. An optical system when traversed by an optical ray in one direction has a certain ABCD matrix. What is the corresponding matrix if the ray enters from the opposite direction?

4. A lens guide consists of a large number of identical lenses of focal length f separated from each other by a distance d. What must the relation between d and f be such that a ray, launched into the lens system at a small angle be confined to the lens system?

5. Discuss in turn the general physical significance of having one of the elements of the ray matrix equal to zero.

6. Determine the effective focal length of a symmetrical thick lens which is 1 cm thick and has a radius of curvature of both surfaces equal to 5 cm. Assume the refractive index of the lens is 1.5.

1 Paraxial Optics

In earlier chapters plane waves were found to be useful for discussions of many elementary optical effects. At the opposite extreme of plane waves, we have the optical rays, which are ideal "pencils" of light with no width and which form the basis for geometrical optics, in which we can ignore the wavelength of light. Most realistic optical beams have a finite transverse extent and for most practical situations are important to consider.

Of particular interest are those modes or waves which have finite transverse extent or relatively small phase variations along directions perpendicular to the overall direction of propagation, but which can also form a complete set for describing any optical beam. We have already seen one example of such a wave in the chapters on Diffraction and Fourier optics where we considered the paraxial section of a spherical wave. We saw that we could replace

$$S(r) = \frac{e^{ikr}}{r} \quad \text{by} \quad h(x,y,z) = \frac{1}{z} e^{ikz} \; e^{ik(x^2+y^2)/2z} \tag{1}$$

for values of x, y small compared to z. The functional form basically represents the impulse function associated with Fresnel diffraction. This is a particular type of *paraxial* wave. In general, we will consider solutions to the wave equation, $\Psi(x,y,z)$, to be of a paraxial nature if their phase variations in the x,y direction are small compared to their phase variations in the direction of propagation (z). An alternative way of saying the same thing is to say that the $\vec{k}$ vectors associated with the plane wave expansion of the optical wave make small angles with respect to the z-axis. Notice that the paraxial waves do not necessarily imply finite transverse extent of the waves. Indeed with the definition we have for paraxial waves the plane wave e^{ikz} would be considered paraxial. For the paraxial waves we could consider writing

$$\Psi(x,y,z) = u(x,y,z)e^{ikz} \tag{2}$$

When this assumed form is substituted in the Helmholtz equation,

$$\bar{\nabla}^2 \Psi + k^2 \Psi = 0 \tag{3}$$

we obtain by direct substitution

$$\bar{\nabla}_T{}^2 u + \frac{\partial^2 u}{\partial z^2} + 2ik\frac{\partial u}{\partial z} = 0 \tag{4}$$

where

$$\bar{\nabla}_T = \hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} \tag{5}$$

We can consider the class of envelope functions, u, which vary only slowly along the direction of propagation. Particularly, take the change in the function over a wavelength to be small, i.e.,

$$\left|\frac{\partial u}{\partial z}\right| << k|u| \tag{6}$$

and also take the function to be smooth on the same scale, i.e.,

$$\left|\frac{\partial^2 u}{\partial z^2}\right| << k\left|\frac{\partial u}{\partial z}\right| \tag{7}$$

With this *slowly varying envelope approximation.* (SVEA), we arrive at what is known as the *paraxial wave equation:*

$$\bar{\nabla}_T{}^2 u + 2ik\frac{\partial u}{\partial z} = 0 \tag{8}$$

which is an approximate form of the wave equation. It can be verified that the function h(x,y,z) is an exact solution of the paraxial wave equation.

One of the most important types of paraxial waves is a *Gaussian beam* which is a particular solution of the paraxial wave equation. In this chapter we develop the theory of Gaussian beams and consider their properties in free space and in optical resonators. We will also consider a family of solutions to the paraxial equations, the Hermite-Gaussian beams, of which the Gaussian beam is a special member. These beams in general are very important in the discussion of light field distributions emerging from laser systems and Fabry-Perot resonators. There are many ways to introduce such beams, none of which is particularly insightful and most of which are mathematically cumbersome. For example, since it has been hinted that the Hermite-Gaussian beams are associated with, among other things, Fabry-Perot resonators, we might solve the 2wave equation for such resonators with appropriate boundary conditions. The natural modes of the resonator are Hermite-Gaussian beams but they can only be identified by extensive, self-consistent mathematical (computer) calculation. Our approach will be much more pragmatic. We will introduce a particular solution to the paraxial wave equation and later show that this particular solution satisfies the requirement of a mode of a resonator.

2 Gaussian Beams

Quite simply, to introduce the Gaussian beam we note that the paraxial wave equation is invariant with respect to a translation of the co-ordinate z to z- z_c where z_c is a constant. In particular, a very interesting solution of the paraxial wave equation occurs if we consider the function h(x,y,z) translated by the amount iz_o where z_o is a real constant. The function, h(x,y,z-iz_o), which obviously satisfies the paraxial wave equation, has an envelope function with the singularity on the z-axis (at z= 0) removed. For reasons to be explained later we will label this function u'_{oo} . It is given by

$$u'_{oo}(x,y,z) = \frac{1}{z-iz_o} \exp\left(\frac{ik[x^2+y^2]}{2(z-iz_o)}\right) \tag{9}$$

Like h(x,y,z), the function u'_{oo} is cylindrically symmetric about the z axis. It is convenient to normalize u'_{oo} (to give u_{oo}) through multiplication by a constant so that

$$\int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dy|u'_{oo}(x,y,z)|^2 = 1 \tag{10}$$

Normalization at one cross section, say at z=0, assures that the normalization will be the same at other values of z, by conservation of power. When the integral is carried out it is found that

$$u_{oo}(x,y,z) = \sqrt{\frac{kz_o}{\pi}}\ \frac{1}{z-iz_o} \exp\left(\frac{ik[x^2+y^2]}{2(z-iz_o)}\right) \tag{11}$$

Apart from a constant phase factor this can be put in the form

$$u_{oo}(x,y,z) = \sqrt{\frac{2}{\pi}}\ \frac{1}{w}\ e^{-i\phi} \exp\left(-\frac{[x^2+y^2]}{w^2}\right) \exp\left(\frac{ik(x^2+y^2)}{2R}\right) \tag{12}$$

where

$$w^2(z) = \left(\frac{\lambda z_0}{\pi}\right)\left(1+\frac{z^2}{z_o^2}\right) = w_0^2\left(1+\frac{z^2}{z_o^2}\right) \tag{13}$$

$$R(z)^{-1} = \frac{z}{z^2+z_o^2} \tag{14}$$

and finally

$$\tan\phi = \frac{z}{z_o} \tag{15}$$

This particular solution of the paraxial wave equation is the *fundamental Gaussian beam solution*. Note that apart from the wavelength λ and the location of the origin ($z=0$) a single parameter (e.g. z_o) completely defines the form of the beam. Before proceeding to use this solution we should understand the various factors which constitute this expression.

The properties of the Gaussian beam solution are:

1) The beam has a field and intensity profile which are a Gaussian function of the transverse variable $r = \sqrt{x^2 + y^2}$. The parameter w represents the value of r at which the field drops to e^{-1} of its value on axis. The parameter is sometimes referred to as the 3spot size since it is a measure of the transverse extent of the beam. The constant w_o is the minimum spot size and occurs at $z = 0$. Conversely we might wish to say that the choice of the displacement of the origin by the imaginary distance iz_o has fixed the minimum spot size. The parameter w_o is sometimes called the fundamental spot size or the beam waist. The distance z_o, known as the *confocal parameter*, is the distance over which the spot size increases by a factor of $\sqrt{2}$.

2) Surfaces for which the intensity is a constant fraction of the on-axis intensity (at the same value of z) are defined by the equation

$$\frac{r^2}{w_o^2\left(1 + \frac{z^2}{z_o^2}\right)} = C = \text{constant} \tag{16}$$

or

$$x^2 + y^2 - \frac{C w_o^2}{z_o^2} z^2 = C w_o^2 \tag{17}$$

These represent hyperboloids of revolution as illustrated in Fig. 1.

Note from the figure that the confocal parameter is a measure of the distance over which the beam is quasicollimated. It is akin to the *depth of focus or depth of field*, terms which are used by camera savants, hence the name, confocal parameter. The parameter z_o varies as w_o^2. Hence, for a more tightly focussed Gaussian beam, one will have a smaller depth of field over which the beam appears to be collimated. For example, if $\lambda = 1\mu m$ and $w_o = 1mm$ we obtain a depth of field of πm, but if $w_o = 10\ \mu m$ we obtain a depth of field of $\pi x 10^{-4}$ m!

3) In the far field where the hyperbolic surfaces approach asymptotes, we can calculate the uniform rate of divergence of the beam. For $z >> z_o$ we have that $w \propto z$. It follows that if θ is the full cone angle determined by the asymptotes, then

$$\tan\frac{\theta}{2} = \frac{w(z)}{z} \tag{18}$$

$$\cong \frac{w_o\left(\frac{z}{z_o}\right)}{z} = \frac{\lambda}{\pi w_o} \cong \frac{\theta}{2} \quad \text{for small} \quad \frac{\lambda}{\pi w_o} \tag{19}$$

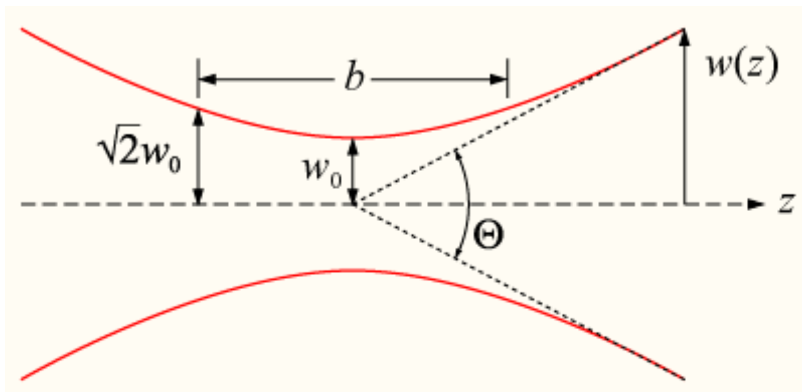


Fig. 1: Surfaces of constant spot size *w(z)* follow hyperboloids of revolution, according to Eqn. 17. Note that like a FWHM measurement, *w(z)* – the half-width 1/e-max of the *E*-field – is measured relative to the peak. But the peak decreases in either direction as one goes away from the beam-waist at *z*=0.

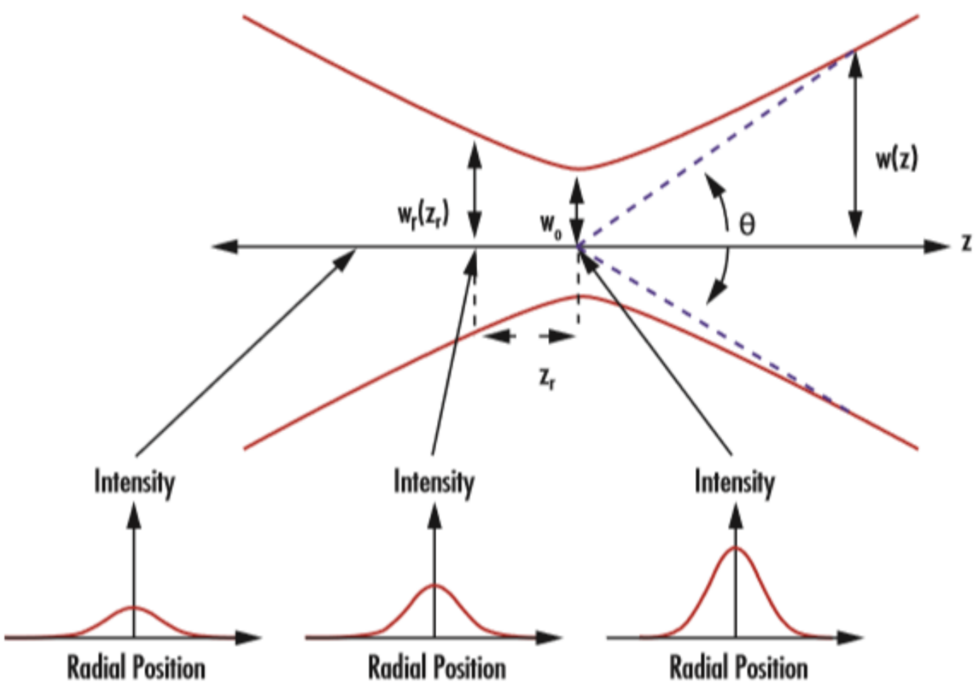


Fig. 2: This figure makes clear that the hyperboloids which trace the 1/e-max surfaces are not surfaces of constant intensity, but just the 'shoulders' of the profile.

For a 1mm fundamental spot size and λ = 1μm, we obtain a full angle of divergence (θ) of the beam of ≈ 10^{-3} radians (the spot increases in size by about 1mm for each metre of travel). For a 1 μm fundamental spot size the full angular divergence is greater than 1 radian. In this case one can question the Gaussian beam solution as being a valid solution of the paraxial wave equation.

4) The quantity R(z) is the radius of curvature of the surfaces of constant phase as shown in Figure 2.

At the beam waist the radius of curvature is infinite, as the defining equation for R(z) indicates. Alternatively, the plane for which R= □ could be used to define the location of the beam waist. For $z >> z_o$ we find that R(z) = z, which means that the beam in the far field is propagating like a portion of a spherical wave. This is consistent with our starting point since if $z >> z_o$ the Gaussian beam is essentially the same as h(x,y,z). Finally, note that the surfaces of constant phase are locally perpendicular to the surfaces of constant field or intensity, as must be for solutions to the homogeneous wave equation.

5) The phase factor ϕ determines the velocity of surfaces of constant phase. The phase speed of the Gaussian beam is *not* the speed of a plane wave in whatever medium the beam is propagating, which, in this case we have taken to be vacuum. We can determine the effective propagation constant from $\Psi(x,y,z)$ through the definition

$$\int_o^z k_{eff}\, dz = kz - \phi(z) \tag{20}$$

5

so that

$$k_{eff} = k - \frac{d\phi}{dz} = \frac{\omega}{c} - \frac{z_0}{z^2 + z_o^2} < \frac{\omega}{c} \tag{21}$$

The phase velocity is everywhere greater than the speed of light. At z = 0 in particular, the effective propagation constant is

$$k_{eff} = k - \frac{2}{k w_o^2} \tag{22}$$

The fact that the phase velocity is greater than c can be explained by the finite transverse extent of the beam. Such a beam can be written as a superposition of plane waves which have propagation vectors, $\vec{k}$, which are oriented at slightly different angles relative to the z- axis. If we consider the beam in the vicinity of z= 0 say, the typical x and y components of the propagation vector $\vec{k}$ of these waves are

$$k_x = k_y \cong \frac{\sqrt{2}}{w_o} \tag{23}$$

Thus with

$$k_z^2 + k_x^2 + k_y^2 = k^2 \tag{24}$$

we have

$$k_z = k - \frac{k_x^2 + k_y^2}{2k} = k - \frac{2}{k w_o^2} \tag{25}$$

in agreement with k_{eff}.

Before leaving these general comments on Gaussian beams it is interesting to examine the range of validity of the Gaussian beam as a solution to the paraxial wave equation. The key approximations we made in arriving at the paraxial wave equation is that

$$\left|\frac{\partial^2 u}{\partial z^2}\right| << k\left|\frac{\partial u}{\partial z}\right| \tag{26}$$

and

$$\left|\frac{\partial u}{\partial z}\right| << k\,|u| \tag{27}$$

For the Gaussian beam solution we have that

$$\frac{\partial u_{oo}}{\partial z} = -\left[\frac{1}{z - iz_o} + \frac{ik(x^2 + y^2)}{2(z - iz_o)}\right] u_{oo} \tag{28}$$

The omission of this term compared with $k|u|$ implies that

$$\frac{1}{z_o} << k \tag{29}$$

or that

6

$$\frac{1}{z_o k} = \frac{\lambda^2}{2\pi^2 w_o{}^2} << 1 \tag{30}$$

We therefore require that the beam waist must be large compared to the wavelength. Further, we must have

$$\frac{x^2 + y^2}{z^2 + z_o{}^2} << 1 \tag{31}$$

Since $x^2 + y^2$ is of the order of w^2, using the expression for w^2 in terms of z and z_o, we find that

$$\frac{w_o{}^2}{z_o{}^2} << 1 \tag{32}$$

which is the same condition as derived in equation 30.

As a final point it should be noted that the function u_{oo} is obviously a scalar quantity, but more importantly, it is an approximation for the electric field associated with the beam. It can't be the exact electric field because Gauss' law is not satisfied exactly. For a beam of finite cross section, in general the electric field must be a vector field which is not transverse in nature. Identifying u_{oo} with the electric field is valid to the same extent that the paraxial approximation is valid.

3 Transformation of Gaussian Beams

Gaussian beams not only represent one of the most fundamental solutions of the paraxial equation but they also represent one of the most common beams encountered, particularly when dealing with lasers. We have learned in some detail the properties of Gaussian beams and how they propagate in free space or a homogeneous medium. What happens to our description of these beams when they pass into or through a different medium such as a lens? Do we have to start from scratch and re-solve the paraxial wave equation with appropriate boundary conditions? Of course we could do that, but for many common situations it turns out we don't have to. It becomes easier to describe the transformation properties of Gaussian beams using matrix techniques.

To begin the discussion of the transformation properties recall that the parameter, $q = z - iz_o$, (known as the Gaussian beam parameter) completely specifies, apart from intensity, the Gaussian beam at position z. Indeed, we have that

$$\frac{1}{q(z)} = \frac{1}{R(z)} + \frac{i\lambda}{\pi n\, w(z)^2} \tag{33}$$

so that the real part of $1/q$ gives us the inverse radius of curvature of the beam while the imaginary part gives us the local spot size. If we can find how q transforms between two different points,

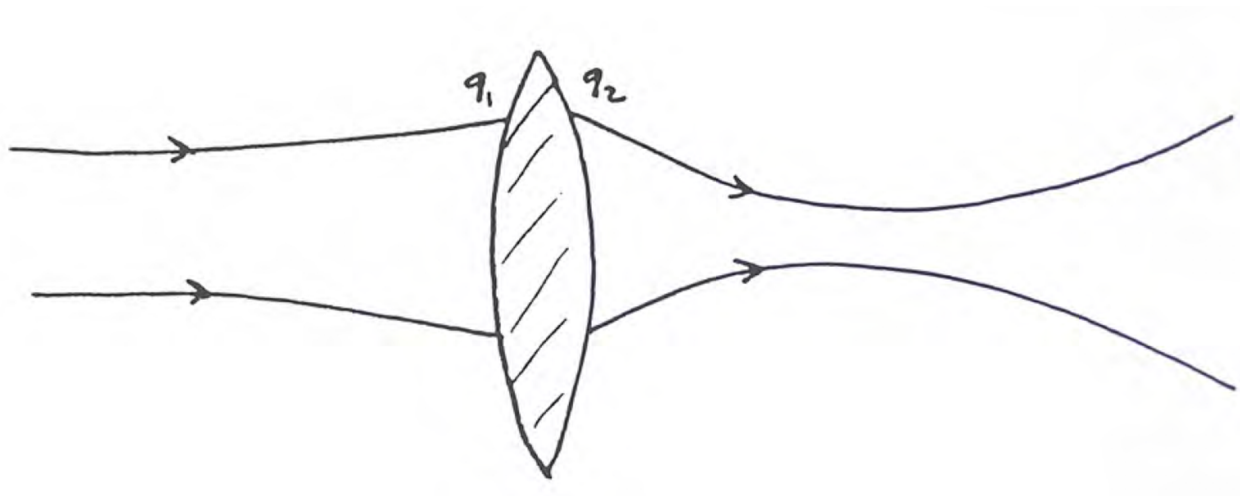


Fig. 3: Transformation of a Gaussian beam by a lens.

regardless of what lies in between the two points, we obviously can define the new Gaussian beam at the new position. For example, consider a Gaussian beam propagating in free space. At a plane defined by $z = z_1$ we have the Gaussian beam parameter

$$q_1 = z_1 - iz_o \tag{34}$$

while, at a different plane defined by $z = z_2$, we have

$$q_2 = z_2 - iz_o \tag{35}$$

The transformation properties of the Gaussian beam propagating in free space between the two planes is trivially given by

$$q_2 = q_1 + (z_2 - z_1) \tag{36}$$

Let's consider a more complicated example, say propagation through a lens. Consider a Gaussian beam propagating through a thin lens of focal length f such that the beam has a Gaussian beam parameter q_1 *immediately before the lens* and a new Gaussian beam parameter q_2 *immediately after the lens* as shown in Figure 3. For a thin lens, we know that the spot size of the Gaussian beam doesn't change so that

$$w_2(z) = w_1(z) \tag{37}$$

The lens, however, imposes a change on the phase front, since the total *optical* path length on axis is greater than away from the axis. The transmission function of a lens of focal length f can be taken as

$$T(x,y) = \exp\left(-\frac{ik[x^2 + y^2]}{2f}\right) \tag{38}$$

When we apply this to the Gaussian beam for fixed z, we have

$$T\,\Psi_{oo}(x,y,z) = \Psi'_{oo} \tag{39}$$

with the only difference between the two beams being the radius of curvature of the phase front. If the new radius of curvature is R', then

$$\frac{ik(x^2 + y^2)}{2R} - \frac{ik(x^2 + y^2)}{2f} = \frac{ik(x^2 + y^2)}{2R'} \tag{40}$$

or

$$\frac{1}{R} - \frac{1}{f} = \frac{1}{R'} \tag{41}$$

Because the spot size does not change we have that

$$\frac{1}{q} - \frac{1}{f} = \frac{1}{q'} \tag{42}$$

so that the transformation of the Gaussian beam is given by

$$q' = \frac{q}{\left(-\frac{1}{f}\right)q + 1} \tag{43}$$

8

Although it is difficult to prove in general, it turns out that the transformation of a Gaussian beam can be represented by an equation of the form

$$q' = \frac{Aq + B}{Cq + D} \tag{44}$$

which, in general, is known as a fractional linear, or Möbius transformation. For the Gaussian beam parameter this transformation is known colloquially as the "ABCD law". The four parameters, A,B,C,D define a transformation matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{45}$$

which although, for Gaussian beams, it is never used like a matrix, it has the same form as the ABCD matrices we considered for rays in chapter 6!

For propagation in free space through a distance z_2 - z_1 we have seen that the transformation matrix is given by

$$\begin{bmatrix} 1 & z_2 - z_1 \\ 0 & 1 \end{bmatrix} \tag{46}$$

while for a lens of focal length f we have the transformation matrix

$$\begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix} \tag{47}$$

where, of course, f is positive for a converging lens and negative for a diverging lens.

The other important transformation matrices for Gaussian beams can be shown to have forms identical to those for rays.

One of the benefits of using the Möbius transformation for the Gaussian beam parameter is that it becomes easy to treat multiple, successive transformations, e.g. by a lens, free-space propagation, an interface, other lenses, etc. Indeed, if a Gaussian beam is propagating through a series of N optical "elements" each of which has an associated transformation matrix M_i then the overall transformation matrix of the system is easily shown to be

$$\mathbf{M}_S = \prod_{i=1}^{N} \mathbf{M}_i \tag{48}$$

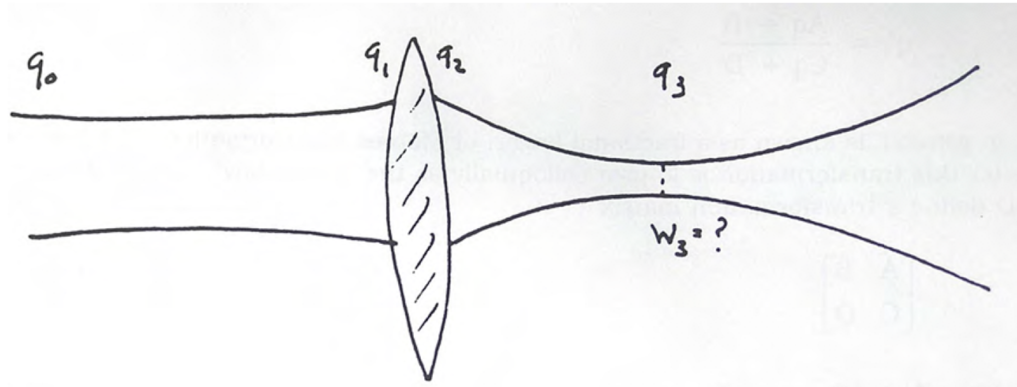


Fig. 2: Focussing a Gaussian laser beam

where the order of the matrices, from right to left is the order in which the Gaussian beam would encounter the associated elements, i.e.

$$\mathbf{M}_S = \ldots \mathbf{M}_{second}\,\mathbf{M}_{first} \tag{49}$$

To illustrate the simplicity of the use of the transformation matrices let's consider the following example. Say we have a He-Ne laser beam operating in a Gaussian mode with a divergence of 1mR and with a beam waist at the output of the laser of 0.4mm. What is the diffraction limited spot size we can achieve with a positive lens of focal length 2cm, located 1m from the beam waist? The situation is depicted in Fig. 2.

In considering the problem, we would start with a beam parameter q_o at the beam waist. This gets transformed into a parameter q_1 just before the lens, and a parameter q_2 just after the lens. Finally at the focal spot of the beam the parameter is q_3. If we can determine the imaginary part of q_3^{-1}, we will have the spot size at the focus. The overall transformation matrix of the system is the product of three matrices, namely those associated with propagation in free space through a distance of 1m, propagation through the lens and propagation through a distance which will bring us to the focal spot. This system matrix can then be used to relate q_o to q_3 from which we could find w_3. For illustration purposes however, let's break the problem up into its elementary constituents to see what actually happens to the Gaussian beam.

To determine q_o from the information given we recall that the divergence of a Gaussian beam is given by

$$\theta = \frac{2\lambda}{\pi w_o} \tag{50}$$

which, for the numbers given implies that λ =0.63 μm. It follows that

$$q_o = 0 - 0.8i; \quad z_o = 0.8\text{m} \tag{51}$$

and

$$q_1 = q_o + 1 \tag{52}$$

giving R_1 = 1.64m and w_1 = 0.64 mm. On passage through the lens we have

$$\begin{aligned} q_2^{-1} &= q_1^{-1} - f^{-1} \\ &= \frac{1}{R_1} - \frac{i\lambda}{\pi w_1^2} - \frac{1}{f} \\ &= (0.61 - 50) - 11.9 \text{ x } 10^{-3} i \end{aligned} \tag{53}$$

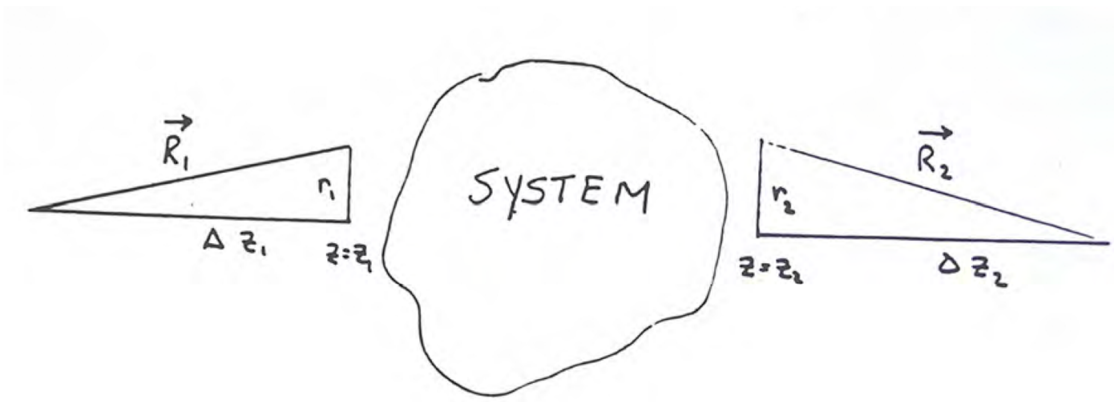


Fig. 3: Rays entering and leaving an optical system.

Note that the radius of curvature of the beam emerging from the lens is not 2cm, so the beam will not focus exactly 2cm behind the lens. Only an incident *plane wave* will focus at a distance f behind a lens of focal length f as we saw in the chapter on diffraction. To determine where the focal spot is in our case we note that

$$\begin{aligned} q_3 &= q_2 + l \\ &= (-.021 + l) + 2 \text{ x } 10^{-4} i \end{aligned} \tag{54}$$

where l is the distance to the focal point. Now the focal point is defined to be the position of the beam waist, which in turn is where the radius of curvature of the beam is infinite and the Gaussian beam parameter is purely imaginary. Hence l = 2.1 cm. We can then determine the beam waist from

$$q_3^{-1} = -i\left(\frac{\lambda}{\pi w_3^{\,2}}\right) = -5 \text{ x } 10^3 i \tag{55}$$

giving w_3 = 6.3 μm and also giving the depth of field, z_o(3) of the focussed beam to be 200 μm.

It's a remarkable fact that the same set of matrices apply to rays and Gaussian beams. It also applies to paraxial portions of spherical waves (z_o ∅ 0). This is remarkable for two reasons:

1) The transformation of Gaussian beams is governed by a fractional linear transformation while that of rays is governed by a true matrix transformation.

2) In dealing with rays one totally ignores the wave character of light while for Gaussian beams it is explicitly included.

We can remove some of the mystery of the similarity between the results if we rewrite the ray transformation law as

$$\frac{r_2}{r_2'} = \frac{A\left(\frac{r_1}{r'}\right) + B}{C\left(\frac{r_1}{r_1'}\right) + D} \tag{56}$$

Referring to Fig. 3, we can define a distance

$$\Delta z_1 = \left.\frac{r}{\left(\frac{dr}{dz}\right)}\right|_{z=z_1} = \frac{r_1}{r_1'} \tag{57}$$

which is the distance between the reference plane and the intersection point of the ray with the z-axis. The intersection point represents the effective source point from which all rays with the ray parameters r_1 r_1' seem to be emanating. Such rays of course lie on a cone. Similarly the distance

$$\Delta z_2 = \frac{r_2}{r_2'} \tag{58}$$

is the effective source or, possibly convergence or focus point, associated with all rays with parameters r_2 and r_2'.

A Gaussian beam may be considered to be the paraxial limit of a solution to a wave equation for a point source with the source shifted by the imaginary amount 11iz_o. Without the origin shift, recall that the paraxial solution of the wave equation is a portion of a spherical wave emanating from z = 0. The distance $z - iz_o = q$ measures the "complex distance" from the reference plane (location of the point source for spherical waves or beam waist for Gaussian waves) to the intersection point with the axis of the "complex ray" pertaining to the Gaussian mode. This is why q obeys the same transformation law as $r/r' = \Delta z$.

References

A. Yariv, *Introduction of Optical Electronics*, (Holt, Rinehart, Winston, New York), 1976.

A.E. Siegman, *Introduction to Laser Physics*, (Prentice Hall, New York) 1971.

H.A. Haus, *Waves and Fields in Optoelectronics*, (Prentice Hall, New York), 1981.

Exercises

1. Determine the approximate error made in associating u_{oo} with the magnitude of the electric field for a Gaussian beam at different points on the beam. At what point on the beam is the error likely to be largest?

2. A Gaussian beam with w_o =0.05mm and λ=0.5μm has its waist located 20 cm from a lens of focal length 2 cm. Behind the lens is a semi-infinite slab of glass with n= 1.5. Where does the beam come to a focus.

Gaussian Beam Propagation

In most laser applications it is necessary to focus, modify, or shape the laser beam by using lenses and other optical elements. In general, laser-beam propagation can be approximated by assuming that the laser beam has an ideal Gaussian intensity profile, which corresponds to the theoretical TEM_{00} mode. Coherent Gaussian beams have peculiar transformation properties which require special consideration. In order to select the best optics for a particular laser application, it is important to understand the basic properties of Gaussian beams.

Unfortunately, the output from real-life lasers is not truly Gaussian (although helium neon lasers and argon-ion lasers are a very close approximation). To accommodate this variance, a quality factor, M^2 (called the "M-squared" factor), has been defined to describe the deviation of the laser beam from a theoretical Gaussian. For a theoretical Gaussian, $M^2 = 1$; for a real laser beam, $M^2 > 1$. The M^2 factor for helium neon lasers is typically less than 1.1; for ion lasers, the M^2 factor typically is between 1.1 and 1.3. Collimated TEM_{00} diode laser beams usually have an M^2 ranging from 1.1 to 1.7. For high-energy multimode lasers, the M^2 factor can be as high as 25 or 30. In all cases, the M^2 factor affects the characteristics of a laser beam and cannot be neglected in optical designs.

In the following section, Gaussian Beam Propagation, we will treat the characteristics of a theoretical Gaussian beam (M^2=1); then, in the section Real Beam Propagation we will show how these characteristics change as the beam deviates from the theoretical. In all cases, a circularly symmetric wavefront is assumed, as would be the case for a helium neon laser or an argon-ion laser. Diode laser beams are asymmetric and often astigmatic, which causes their transformation to be more complex.

Although in some respects component design and tolerancing for lasers is more critical than for conventional optical components, the designs often tend to be simpler since many of the constraints associated with imaging systems are not present. For instance, laser beams are nearly always used on axis, which eliminates the need to correct asymmetric aberration. Chromatic aberrations are of no concern in single-wavelength lasers, although they are critical for some tunable and multiline laser applications. In fact, the only significant aberration in most single-wavelength applications is primary (third-order) spherical aberration.

Scatter from surface defects, inclusions, dust, or damaged coatings is of greater concern in laser-based systems than in incoherent systems. Speckle content arising from surface texture and beam coherence can limit system performance.

Because laser light is generated coherently, it is not subject to some of the limitations normally associated with incoherent sources. All parts of the wavefront act as if they originate from the same point; consequently, the emergent wavefront can be precisely defined. Starting out with a well-defined wavefront permits more precise focusing and control of the beam than otherwise would be possible.

For virtually all laser cavities, the propagation of an electromagnetic field, $E^{(0)}$, through one round trip in an optical resonator can be described mathematically by a propagation integral, which has the general form

$$E^{(1)}(x,y) = e^{-jkp} \iint_{Input\,Plane} K(x,y,x_0,y_0)\,E^{(0)}(x_0,y_0)\,dx_0dy_0 \qquad (2.1)$$

where K is the propagation constant at the carrier frequency of the optical signal, p is the length of one period or round trip, and the integral is over the transverse coordinates at the reference or input plane. The function K is commonly called the propagation kernel since the field $E^{(1)}(x, y)$, after one propagation step, can be obtained from the initial field $E^{(0)}(x_0, y_0)$ through the operation of the linear kernel or "propagator" $K(x, y, x_0, y_0)$.

By setting the condition that the field, after one period, will have exactly the same transverse form, both in phase and profile (amplitude variation across the field), we get the equation

$$\gamma_{nm}E_{nm}(x,y) \equiv \iint_{Input\,Plane} K(x,y,x_0,y_0)\,E_{nm}(x_0,y_0)\,dx_0dy_0 \qquad (2.2)$$

where E_{nm} represents a set of mathematical eigenmodes, and γ_{nm} a corresponding set of eigenvalues. The eigenmodes are referred to as transverse cavity modes, and, for stable resonators, are closely approximated by Hermite-Gaussian functions, denoted by TEM_{nm}. (Anthony Siegman, Lasers)

The lowest order, or "fundamental" transverse mode, TEM_{00} has a Gaussian intensity profile, shown in figure 2.1, which has the form

$$I(x,y) \propto e^{-k\left(x^2+y^2\right)} \qquad (2.3)$$

In this section we will identify the propagation characteristics of this lowest-order solution to the propagation equation. In the next section, Real Beam Propagation, we will discuss the propagation characteristics of higher-order modes, as well as beams that have been distorted by diffraction or various anisotropic phenomena.

BEAM WAIST AND DIVERGENCE

In order to gain an appreciation of the principles and limitations of Gaussian beam optics, it is necessary to understand the nature of the laser output beam. In TEM_{00} mode, the beam emitted from a laser begins as a perfect plane wave with a Gaussian transverse irradiance profile as shown in figure 2.1. The Gaussian shape is truncated at some diameter either by the internal dimensions of the laser or by some limiting aperture in the optical train. To specify and discuss the propagation characteristics of a laser beam, we must define its diameter in some way. There are two commonly accepted definitions. One definition is the diameter at which

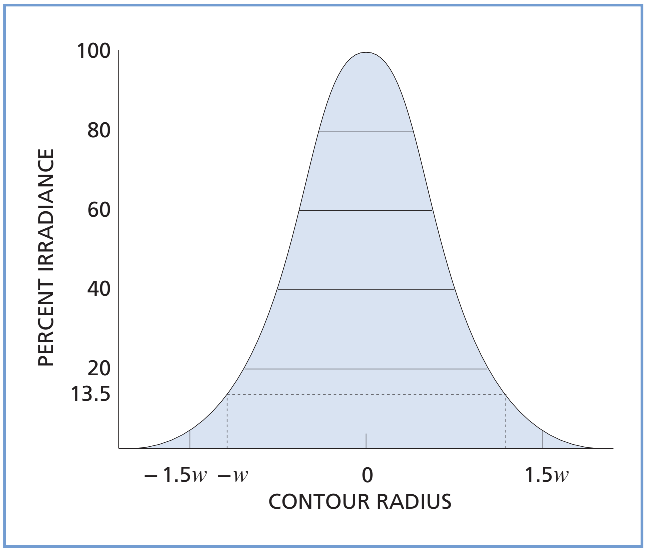


Figure 2.1 **Irradiance profile of a Gaussian TEM_{00} mode**

the beam irradiance (intensity) has fallen to $1/e^2$ (13.5 percent) of its peak, or axial value and the other is the diameter at which the beam irradiance (intensity) has fallen to 50 percent of its peak, or axial value, as shown in figure 2.2. This second definition is also referred to as FWHM, or full width at half maximum. For the remainder of this guide, we will be using the $1/e^2$ definition.

Diffraction causes light waves to spread transversely as they propagate, and it is therefore impossible to have a perfectly collimated beam. The spreading of a laser beam is in precise accord with the predictions of pure diffraction theory; aberration is totally insignificant in the present context. Under quite ordinary circumstances, the beam spreading can be so small it can go unnoticed. The following formulas accurately describe beam spreading, making it easy to see the capabilities and limitations of laser beams.

Even if a Gaussian TEM_{00} laser-beam wavefront were made perfectly flat at some plane, it would quickly acquire curvature and begin spreading in accordance with

$$R(z) = z\left[1+\left(\frac{\pi w_0^2}{\lambda z}\right)^2\right] \qquad (2.4)$$

and

$$w(z) = w_0\left[1+\left(\frac{\lambda z}{\pi w_0^2}\right)^2\right]^{1/2} \qquad (2.5)$$

where z is the distance propagated from the plane where the wavefront is flat, λ is the wavelength of light, w_0 is the radius of the $1/e^2$ irradiance contour at the plane where the wavefront is flat, $w(z)$ is the radius of the $1/e^2$ contour after the wave has propagated a distance z, and $R(z)$ is the wavefront radius of curvature after propagating a distance z. $R(z)$ is infinite at $z=0$, passes through a minimum at some finite z, and rises again toward infinity as z is further increased, asymptotically approaching the value of z itself. The plane $z=0$ marks the location of a Gaussian waist, or a place where the wavefront is flat, and w_0 is called the beam waist radius.

The irradiance distribution of the Gaussian TEM_{00} beam, namely,

$$I(r) = I_0 e^{-2r^2/w^2} = \frac{2P}{\pi w^2} e^{-2r^2/w^2}\ , \qquad (2.6)$$

where $w=w(z)$ and P is the total power in the beam, is the same at all cross sections of the beam.

The invariance of the form of the distribution is a special consequence of the presumed Gaussian distribution at $z=0$. If a uniform irradiance distribution had been presumed at $z=0$, the pattern at $z=\infty$ would have been the familiar Airy disc pattern given by a Bessel function, whereas the pattern at intermediate z values would have been enormously complicated.

Simultaneously, as $R(z)$ asymptotically approaches z for large z, $w(z)$ asymptotically approaches the value

$$w(z) = \frac{\lambda z}{\pi w_0} \qquad (2.7)$$

where z is presumed to be much larger than $\pi w_0/\lambda$ so that the $1/e^2$ irradiance contours asymptotically approach a cone of angular radius

$$\theta = \frac{w(z)}{z} = \frac{\lambda}{\pi w_0}. \qquad (2.8)$$

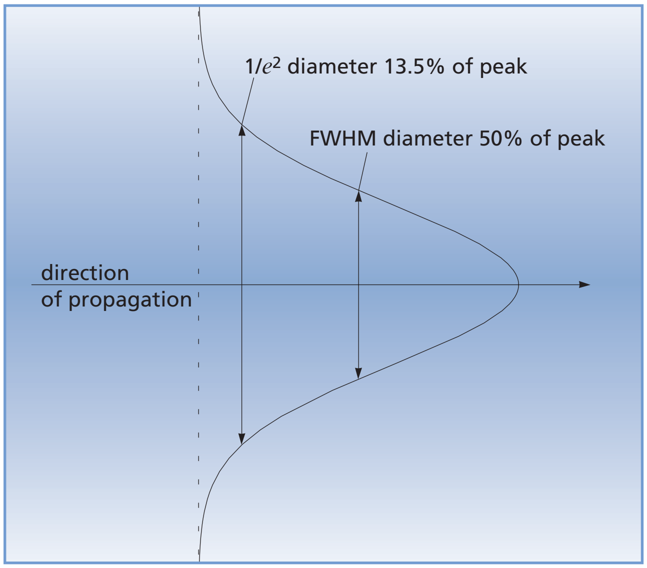


Figure 2.2 **Diameter of a Gaussian beam**

This value is the far-field angular radius (half-angle divergence) of the Gaussian TEM_{00} beam. The vertex of the cone lies at the center of the waist, as shown in figure 2.3.

It is important to note that, for a given value of λ, variations of beam diameter and divergence with distance z are functions of a single parameter, w_0, the beam waist radius.

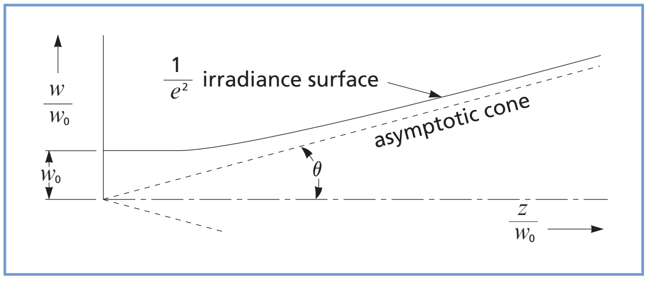


Figure 2.3 **Growth in 1/e² radius with distance propagated away from Gaussian waist**

Near-Field vs Far-Field Divergence

Unlike conventional light beams, Gaussian beams do not diverge linearly. Near the beam waist, which is typically close to the output of the laser, the divergence angle is extremely small; far from the waist, the divergence angle approaches the asymptotic limit described above. The Raleigh range (z_R), defined as the distance over which the beam radius spreads by a factor of $\sqrt{2}$, is given by

$$z_R = \frac{\pi w_0^2}{\lambda} . \qquad (2.9)$$

At the beam waist ($z=0$), the wavefront is planar [$R(0)=\infty$]. Likewise, at $z=\infty$, the wavefront is planar [$R(\infty)=\infty$]. As the beam propagates from the waist, the wavefront curvature, therefore, must increase to a maximum and then begin to decrease, as shown in figure 2.4. The Raleigh range, considered to be the dividing line between near-field divergence and mid-range divergence, is the distance from the waist at which the wavefront curvature is a maximum. Far-field divergence (the number quoted in laser specifications) must be measured at a distance much greater than z_R (usually $>10 \times z_R$ will suffice). This is a very important distinction because calculations for spot size and other parameters in an optical train will be inaccurate if near- or mid-field divergence values are used. For a tightly focused beam, the distance from the waist (the focal point) to the far field can be a few millimeters or less. For beams coming directly from the laser, the far-field distance can be measured in meters.

Typically, one has a fixed value for w_0 and uses the expression

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

to calculate $w(z)$ for an input value of z. However, one can also utilize this equation to see how final beam radius varies with starting beam radius at a fixed distance, z. Figure 2.5 shows the Gaussian beam propagation equation plotted as a function of w_0, with the particular values of $\lambda = 632.8$ nm and $z = 100$ m.

The beam radius at 100 m reaches a minimum value for a starting beam radius of about 4.5 mm. Therefore, if we wanted to achieve the best combination of minimum beam diameter and minimum beam spread (or best collimation) over a distance of 100 m, our optimum starting beam radius would be 4.5 mm. Any other starting value would result in a larger beam at $z = 100$ m.

We can find the general expression for the optimum starting beam radius for a given distance, z. Doing so yields

$$w_0(\text{optimum}) = \left(\frac{\lambda z}{\pi} \right)^{1/2} \qquad (2.10)$$

Using this optimum value of w_0 will provide the best combination of minimum starting beam diameter and minimum beam

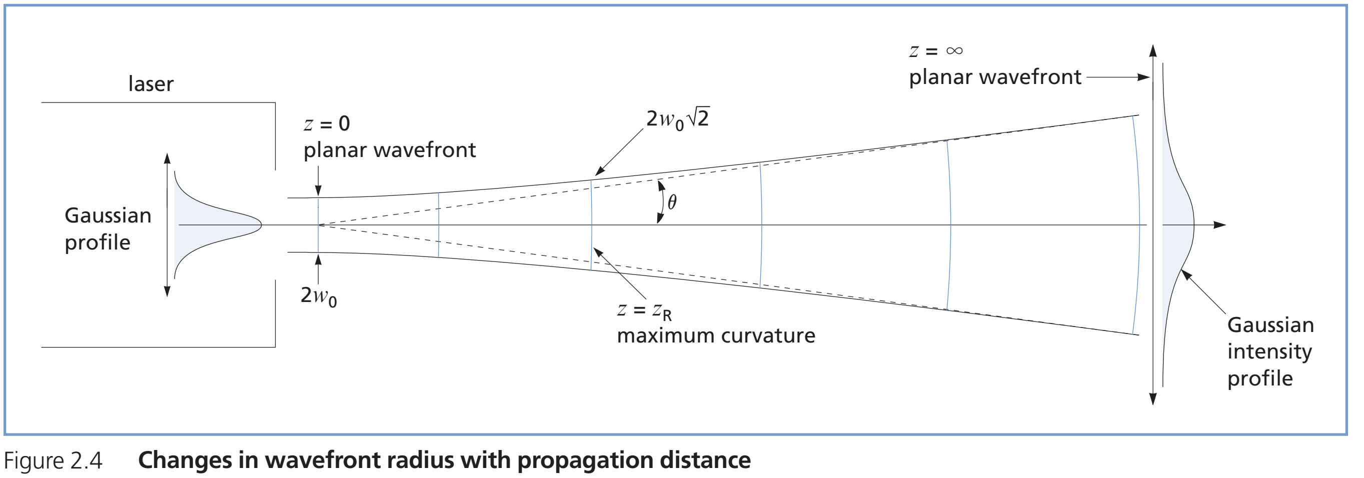


Figure 2.4 **Changes in wavefront radius with propagation distance**

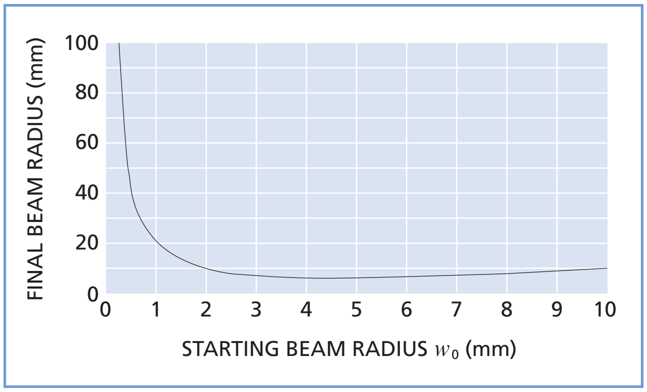


Figure 2.5 **Beam radius at 100 m as a function of starting beam radius for a HeNe laser at 632.8 nm**

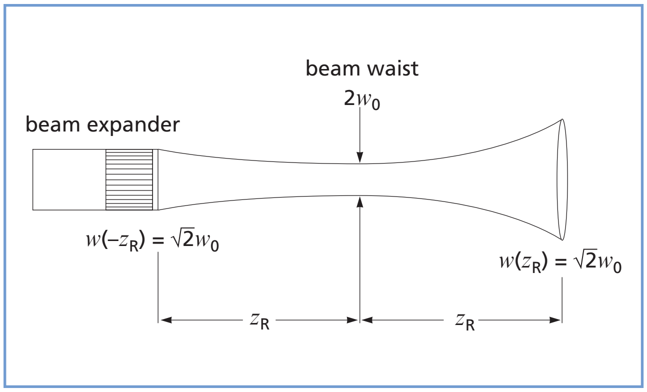


Figure 2.6 **Focusing a beam expander to minimize beam radius and spread over a specified distance**

spread [ratio of $w(z)$ to w_0] over the distance z. For $z = 100$ m and $\lambda = 632.8$ nm, w_0 (optimum) = 4.48 mm (see example above). If we put this value for w_0 (optimum) back into the expression for $w(z)$,

$$w(z) = \sqrt{2}(w_0) \tag{2.11}$$

Thus, for this example,

$$w(100) = \sqrt{2}(4.48) = 6.3 \text{ mm}$$

By turning this previous equation around, we find that we once again have the Rayleigh range (z_R), over which the beam radius spreads by a factor of $\sqrt{2}$ as

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

with

$$w(z_R) = \sqrt{2}w_0.$$

If we use beam-expanding optics that allow us to adjust the position of the beam waist, we can actually double the distance over which beam divergence is minimized, as illustrated in figure 2.6. By focusing the beam-expanding optics to place the beam waist at the midpoint, we can restrict beam spread to a factor of $\sqrt{2}$ over a distance of $2z_R$, as opposed to just z_R.

This result can now be used in the problem of finding the starting beam radius that yields the minimum beam diameter and beam spread over 100 m. Using $2(z_R) = 100$ m, or $z_R = 50$ m, and $\lambda = 632.8$ nm, we get a value of $w(z_R) = (2\lambda/\pi)^{1/2} = 4.5$ mm, and $w_0 = 3.2$ mm. Thus, the optimum starting beam radius is the same as previously calculated. However, by focusing the expander we achieve a final beam radius that is no larger than our starting beam radius, while still maintaining the $\sqrt{2}$ factor in overall variation.

Alternately, if we started off with a beam radius of 6.3 mm, we could focus the expander to provide a beam waist of $w_0 = 4.5$ mm at 100 m, and a final beam radius of 6.3 mm at 200 m.

APPLICATION NOTE

Location of the beam waist

The location of the beam waist is required for most Gaussian-beam calculations. Melles Griot lasers are typically designed to place the beam waist very close to the output surface of the laser. If a more accurate location than this is required, our applications engineers can furnish the precise location and tolerance for a particular laser model.

ImageJ Beam Profile Acquisition Procedure (Full). (edited from notes by Alex Kai Wei, 2020)

Installation of ImageJ

1. Download ImageJ from: https://imagej.nih.gov/ij/download.html. Ensure that the correct version for your computer system is installed.
2. Install a plugin for capturing images. It is suggested that you use webcam capture from ImageJ National Instruments found here: https://imagej.nih.gov/ij/plugins/webcam-capture/index.html.
    a. Extract the files into the plugins folder for ImageJ. Create a new folder named "WebcamCapture" to hold the extracted files.

Capturing the beam profile

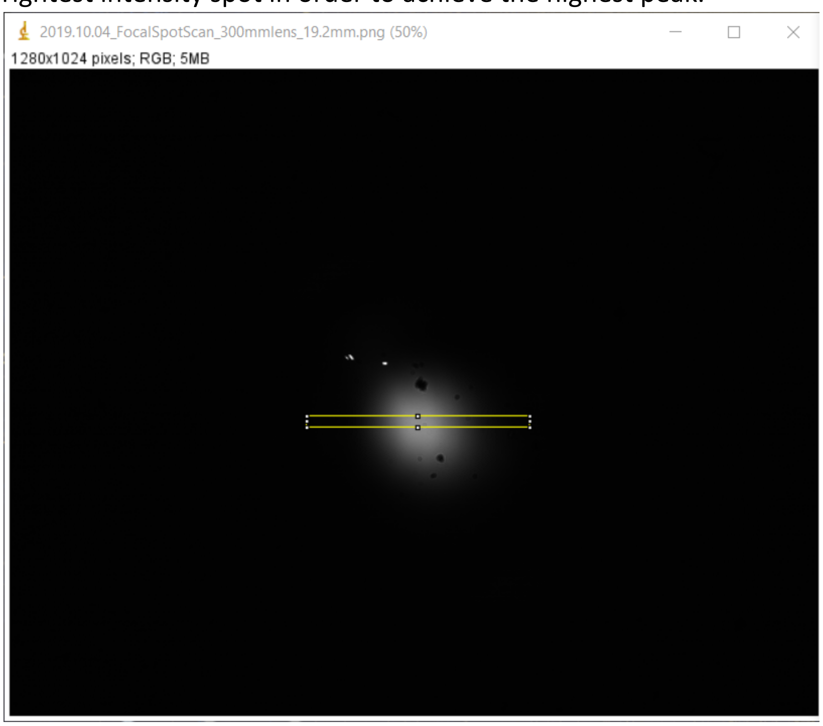
1. Connect an undergraduate lab camera to a lab computer or laptop and open ImageJ. Under the plugins tab, select "WebcamCapture" then "IJ webcam plugin".
2. A window will pop up showing the options for the camera. Change the "Camera name" to the desired camera using the dropdown menu. If the camera is not detected by ImageJ, check that the appropriate drivers are installed for the camera. For windows computers, this process should be automatic when it is plugged in, but if there is no driver, simply look up "camera name" + "driver" or check the drivers in the computer devices.
    a. On a windows computer you can check the drivers by going to Control Panel -> Hardware and Sound -> Devices and Printers: Device Manager. Find your device, (most likely under unknown or Universal Serial Bus controllers), right click on the device and select "Update Driver". Choose "Search automatically for updated driver software" and the necessary driver should be automatically installed for you. If not, you will need to download the driver from a 3rd party through downloading the driver from the internet.
    b. For now, the rest of the settings do not need to be touched as calibration of the pixels into µm is unknown. This process will be done manually. Press ok.
3. The output from the camera should pop up in a separate window. Center the beam on your screen using the micrometers on the translation stage.
4. Adjust the polarizer and filters in the laser system setup to ensure that the image is not too saturated. If you are unsure whether it is too saturated or not, refer to Processing the Beam Profile Step 1b. For the HeNe laser, an OD5 filter with a polarizer allows for easy correction of oversaturation.
5. To capture an image, simply save the current output of ImageJ. This can be done using CTRL+S or COMMAND+S (MAC). Alternatively, use File -> Save. The image output will be in .tif which is preferable. The output image type can be changed using File -> Save As. Create a new folder to hold the saved images for better organization.

Processing the Beam Profile

1. Most scientific cameras for use in optics are monochrome. If the camera used was colour (RGB), split the images into its separate colours for easier analysis. Check what type of image you have by selecting Image -> Type.

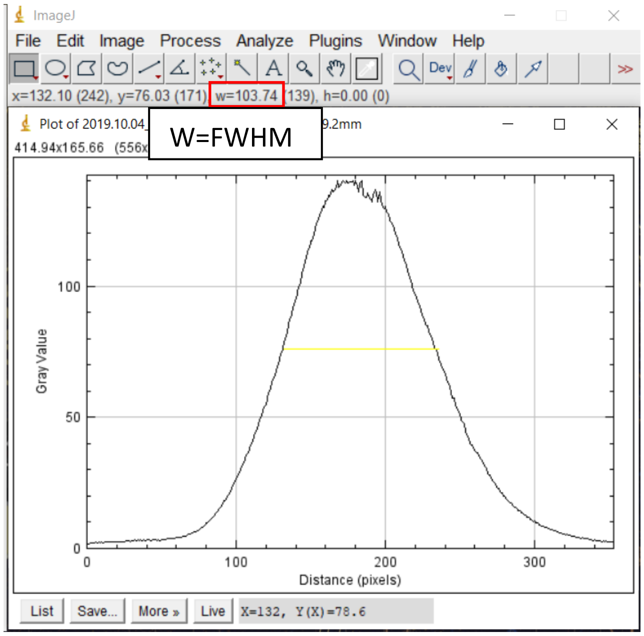
   a. Select Image -> Colour -> Split Channels. This should separate the image into the three RGB colours. This allows for the image of the beam spot not to be saturated and easier to view as it is now monochrome. Thus, from the three colours, choose the best image to analyze (if the red channel is saturated, the green or blue channel may have filtered down). The other windows for the other colours are not needed and can be closed at this time.
   b. Next to visualize the beam profile as clearly as possible, select Image -> Lookup Tables -> Spectrum. This table usually labels saturated parts of the image (pixel values greater than 255) in red. If the peak value is not correctly recorded, then the FWHM calculation will be false.
   c. Tip: When the beam is too large to fit onto the camera sensor, you can use your own phone camera to take an image of the beam profile projected onto a card or piece of paper and still achieve useful results. Calibration may be problematic, but consider using a card or paper that already has a graph-paper grid of known scale on it.

2. On your beam spot image, use the rectangle cursor to draw a thin horizontal rectangular box across the beam. Make sure there is enough “background” on either side of the box to determine both peak and background, for finding the FWHM. Cross the center of the beam spot or the brightest intensity spot in order to achieve the highest peak.



   a.

3. Press “CTRL+K” or Analyze -> Plot Profile. This will bring up the plot of the intensity of the beam profile against the pixels that were selected in the box drawn earlier. This is known as a ‘lineout’ through the profile.. Note that the x axis might appear in units of µm, but the default calibration is 1px:1µm, these values are simply the number of pixels. If the top of the peak is flat or near 255 bits, that means the image is saturated and the polarizers or filters used need to be optimized.

4. The FWHM of the profile can be found easily, if you now draw a box onto the plot itself between the half max locations, and read the length off for the line-length:
   a. 

5. Use: $I = I_o e^{\frac{-2r^2}{w(z)^2}}$ to find the width of the beam. r in this case will be the FWHM/2 when I/I_o is taken to be 0.5.

<u>Calibrating the camera distance scale</u>

1. To calibrate the camera and to figure out what the actual size of the beam is (not in pixels), take a calibration image.
   a. Option 1: With the open camera (no lens), take the initial image of the beam spot, blocking a portion of the beam spot with a sharp edge vertically (the edge of aluminum foil works well). Using the XYZ translation stage, move the camera a known distance horizontally, take another image and measure the distance change between the two images using the vertical edge as a reference between images. Ideally, you want to move the second image all the way across the screen, maximum convenient distance, to reduce error. Thus, the known distance moved in real life, and the number of pixels the image moved, give you the value of mm/pixel for the camera itself.
   b. Option 2: If you're using a lens, take an image of the beam spot on graph paper. Measure the distance between the gridlines in one direction and compare that to the size, measured in pixels, of the same gridlines on the computer image. This will calibrate mm/pixel not just for the camera, but for your whole imaging system.
   c. Option 3: for the beam microscope (beam profiler) you have a microscope lens, and far too much magnification to use graph paper. But you can image the focal spot of the laser, and still translate the camera and lens by a known amount. The pixel-shift between your two images will give number of pixels, the micrometer readings give the actual physical translation, and so the ratio of micrometer displacement to pixel-number shift will give you mm/pixel for the whole imaging system, including the magnification of the microscope.