

A NEW MATHEMATICAL MODEL FOR A  
PROPAGATING GAUSSIAN BEAM

by

Barbara Tehan Landesman

---

A Dissertation Submitted to the Faculty of the  
COMMITTEE on OPTICAL SCIENCES (Graduate)  
In Partial Fulfillment of the Requirements  
For the Degree of  
DOCTOR OF PHILOSOPHY  
In the Graduate College  
THE UNIVERSITY OF ARIZONA

1 9 8 8

THE UNIVERSITY OF ARIZONA  
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read  
the dissertation prepared by Barbara Tehan Landesman

entitled A New Mathematical Model for a Propagating Gaussian Beam

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

and recommend that it be accepted as fulfilling the dissertation requirement  
for the Degree of Doctor of Philosophy

H. N. Barnett

6/27/88  
Date

Roland V. Shack

6/27/88  
Date

Norm N. Lane

6/27/88  
Date

\_\_\_\_\_

Date

\_\_\_\_\_

Date

Final approval and acceptance of this dissertation is contingent upon the  
candidate's submission of the final copy of the dissertation to the Graduate  
College.

I hereby certify that I have read this dissertation prepared under my  
direction and recommend that it be accepted as fulfilling the dissertation  
requirement.

H. N. Barnett  
Dissertation Director

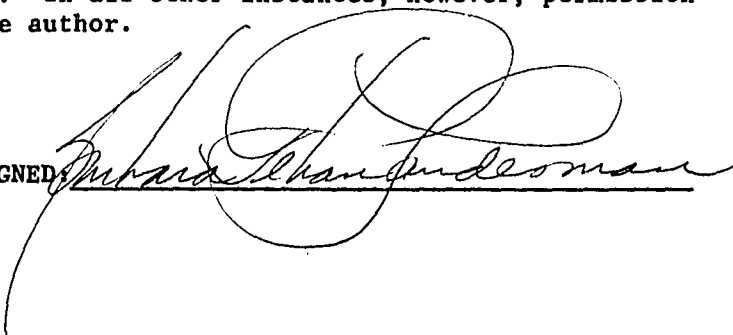
11/10/88  
Date

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED

  
Richard L. Anderson

**Dedicated to the Immaculate Heart of Mary.**

#### ACKNOWLEDGEMENTS

I would like to thank the members of the Final Exam Committee for their efforts on my behalf. In particular, I wish to acknowledge the patience and guidance of my advisor, Harrison Barrett, in bringing this research to fruition. I also wish to thank John Aarsvold for his helpful suggestions and Roland Shack for his role in the development of the ideas contained herein. The many excellent illustrations in this dissertation are the work of Wallace Varner of the Audiovisual Department and James Abbot of the Physics Department. Norma Emptage typed the manuscript, for which I am sincerely grateful.

Finally, and most importantly, I would like to thank my husband, Alan, without whose love, support, and unwavering belief in my abilities this research would never have been completed. I also wish to thank my very good friend, Anita Mullins, for asking the question, "Why not?".

This research was supported in part by the Joint Armed Services Optical Program and the Eastman-Kodak Company.

## TABLE OF CONTENTS

	Page
LIST OF ILLUSTRATIONS . . . . .	7
LIST OF TABLES. . . . .	9
ABSTRACT. . . . .	10
1. INTRODUCTION. . . . .	12
2. HISTORICAL DEVELOPMENT OF A MATHEMATICAL MODEL FOR GAUSSIAN BEAMS. . . . .	21
Approximations, Interpretations and Limitations. . . . .	36
Historical References to Ellipsoidal Wavefronts. . . . .	41
Summary. . . . .	43
3. AN EXACT ZERO-ORDER SOLUTION TO THE SCALAR WAVE EQUATION IN OBLATE SPHEROIDAL COORDINATES . . . . .	44
Spheroidal Coordinate Systems . . . . .	46
Zero-Order Solutions to the Scalar Wave Equation . . . . .	55
A New Mathematical Model for the Fundamental Mode of a Propagating Gaussian Beam . . . . .	62
Summary. . . . .	68
4. A GEOMETRICAL MODEL FOR THE FUNDAMENTAL MODE OF A PROPAGATING GAUSSIAN BEAM . . . . .	69
Skew-Line Generator of a Ruled Surface . . . . .	71
A Geometrical Interpretation of $\exp[-i \tan^{-1} \xi / \eta]$ . . . . .	90
The Skew Line as a Ray . . . . .	95
Nonorthogonal Coordinate System. . . . .	98
Summary. . . . .	105
5. A COMPLETE FAMILY OF SOLUTIONS TO THE SCALAR WAVE EQUATION. . . . .	108
Solution of the Scalar Helmholtz Equation in Oblate Spheroidal Coordinates. . . . .	114
Legendre Polynomials and Associated Legendre Functions. . . . .	119
Spherical Hankel Functions . . . . .	122
Summary. . . . .	124
6. AN HISTORICAL PERSPECTIVE ON GAUSSIAN BEAM PROPAGATION AND COMPLEX FUNCTIONS . . . . .	126
Complex Arguments. . . . .	127
Complex Point Sources. . . . .	130
Complex Rays . . . . .	133

TABLE OF CONTENTS (Continued)

	Page
7. IDEAS FOR FUTURE RESEARCH. . . . .	143
Vector Wave Equation. . . . .	143
Resonator Analysis and Beam Origin. . . . .	145
Diffraction . . . . .	147
Gaussian Beams with Elliptical Cross Section. . . . .	150
Optical Design and Aberration Theory. . . . .	151
Summary . . . . .	153
8. CONCLUSIONS. . . . .	155
APPENDIX A: FIRST ORDER GAUSSIAN BEAM PROPAGATION USING SKEW LINE MODEL. . . . .	
	158
APPENDIX B: PROOF THAT AN ARC OF AN OBLATE ELLIPSE IS PERPENDICULAR TO ITS ATTENDANT SKEW-LINE FAN. . . . .	
	179
APPENDIX C: REFLECTION OF A SKEW-LINE RAY FROM AN ELLIPTICAL MIRROR . . . . .	
	184
REFERENCES. . . . .	188

## LIST OF ILLUSTRATIONS

Figure		Page
2.1	Confocal resonator geometry. . . . .	24
2.2	Gaussian amplitude distribution. . . . .	29
2.3	Hyperbolic contour of constant amplitude in Gaussian beam showing surfaces of constant phase. . . . .	30
2.4	Contour of the fundamental mode of a Gaussian beam . . . . .	37
3.1	Prolate spheroidal coordinate system . . . . .	47
3.2	Oblate spheroidal coordinate system. . . . .	49
3.3	Three-dimensional view of prolate spheroidal coordinate system showing focal points. . . . .	52
3.4	Three-dimensional view of oblate spheroidal coordinate system showing focal ring . . . . .	53
4.1	The orientation of skew line $l$ with respect to the $z$ axis for two possible deviation angles, $\theta$ and $\delta$ . . . . .	72
4.2	Cross section of the hyperboloid generated by the rotation of a skew line about the $z$ axis . . . . .	74
4.3	The skew line $l$ projected onto an $x$ - $y$ plane. . . . .	75
4.4	Different ruled surfaces for varying amounts of twist $\alpha$ . . . . .	76
4.5	Two possible twist orientations for the skew line $l$ . . . . .	79
4.6	A clockwise fan of skew lines ending in elliptical arc PA . . . . .	84
4.7	Cross section of an oblate ellipse and hyperbola with common foci at $\pm d$ . . . . .	86
4.8	Differential change in twist angle, $\Delta\alpha$ . . . . .	92
4.9	Clockwise and counterclockwise skew-line fans. . . . .	97



## LIST OF ILLUSTRATIONS (Continued)

Figure	Page
4.10 Left-handed (clockwise) and right-handed (counter-clockwise) right conoids. . . . .	100
4.11 The unit base vectors $\hat{u}$ , $\hat{v}$ , and $\hat{\theta}$ for a nonorthogonal coordinate system . . . . .	101
5.1 Geometry for a polar coordinate system whose origin has been shifted to one of the foci of the prolate spheroidal coordinate system. . . . .	111

## LIST OF TABLES

Table		Page
5.1	Legendre polynomials. . . . .	120
5.2	Associated Legendre functions . . . . .	121
5.3	Higher-order values of the function, $e^{-kd} h_n^{(1)}[kd(\xi - i\eta)]$	123

## ABSTRACT

A new mathematical model for the fundamental mode of a propagating Gaussian beam is presented. The model is two-fold, consisting of a mathematical expression and a corresponding geometrical representation which interprets the expression in the light of geometrical optics. The mathematical description arises from the  $(0,0)$  order of a new family of exact, closed-form solutions to the scalar Helmholtz equation. The family consists of nonseparable functions in the oblate spheroidal coordinate system and can easily be transformed to a different set of solutions in the prolate spheroidal coordinate system, where the  $(0,0)$  order is a spherical wave. This transformation consists of two substitutions in the coordinate system parameters and represents a more general method of obtaining a Gaussian beam from a spherical wave than assuming a complex point source on axis. Further, each higher-order member of the family of solutions possesses an amplitude consisting of a finite number of higher-order terms with a zero-order term that is Gaussian.

The geometrical interpretation employs the skew-line generator of a hyperboloid of one sheet as a ray-like element on a contour of constant amplitude in the Gaussian beam. The geometrical characteristics of the skew line and the consequences of treating it as a ray are explored in depth. The skew line is ultimately used to build a nonorthogonal coordinate system which allows straight-line propagation of a Gaussian beam in three-dimensional space.

Highlights of the research into other methods used to model a propagating Gaussian beam--such as complex rays, complex point sources and complex argument functions--are reviewed and compared with this work.

## CHAPTER 1

## INTRODUCTION

Amplification of electromagnetic radiation by stimulated emission at optical or near optical wavelengths was proposed and demonstrated around 1960 by Schawlow and Townes [1958], Maiman [1960], and Javan, Bennett and Herriott [1961]. This amplifier or "laser", an acronym for "light amplification by stimulated emission of radiation", was an optical resonator consisting of two mirrors enclosing a gain medium. In an optical resonator, the mirrors reflect energy back and forth through the gain medium building up large field intensities at specified or resonance frequencies until a steady-state oscillation condition is reached. The resonator also acts as a spatial and frequency filter. Fields with certain prescribed spatial variation and frequency will experience a large loss of energy on each round-trip, guaranteeing that low-loss modes will predominate in the steady-state condition.

Following the development of the laser, other researchers, such as Fox and Li [1961], Boyd and Gordon [1961], Boyd and Kogelnik [1962], and Kogelnik and Li [1966], constructed mathematical expressions to describe and predict the behavior of the optical radiation generated by these devices. Their work, and that of others since them, concentrated on characterizing the fields or modes which

predominate after the laser oscillation has reached a steady-state. Furthermore, these fields have generally been expressed as solutions to the scalar Helmholtz or wave equation. These steady-state, low-loss modes will propagate beyond the resonator into a medium or optical system where a model of their behavior is fundamental to understanding and predicting their interaction with the medium or optical system.

A mode of a laser resonator is a wave function that reproduces itself in shape, amplitude and phase after each roundtrip. Fox and Li [1961] expressed this requirement as an integral equation in which the field or wave function across one mirror differs from the field across the opposite mirror by only a complex constant. This quantity determines the energy loss and phase shift the field experiences per round trip. The wave function describes the spatial variation of the field in three dimensions; that is, the modes are functions of the transverse coordinates at a specified position along the resonator.

The mathematical expression used most frequently to represent the modes of a laser is the one introduced by Kogelnik and Li [1966]. In this representation, the fundamental mode of the beam consists of the product of an exponential phase term that is constant over the surface of a parabola, a Gaussian amplitude distribution, an additional amplitude modulation term that describes the beam's loss of energy as it traverses the resonator and an additional phase term that helps determine the resonant frequency. They expressed the

higher-order modes as the product of Hermite polynomials times this fundamental for resonators with rectangular symmetry, and they used Laguerre polynomials for resonators with circular cylindrical symmetry.

Kogelnik and Li began their derivation for both the Hermite-Gaussian and Laguerre-Gaussian polynomials by postulating solutions to the scalar wave equation. This required the introduction of two approximations. First, the solutions were assumed to vary slowly in the direction of propagation ( $z$ ) such that the second derivative with respect to  $z$  was dropped from the differential equation. Second, the solutions were assumed to be valid in only a very limited region near the optic axis, the paraxial region. This is the reason the phase contours are rigorously parabolic rather than spherical; a parabola is a first-order approximation to a sphere and is equivalent to a sphere within the paraxial region. With few exceptions, the overwhelming majority of authors in this area of research also employed these two approximations in developing mathematical expressions for the modes generated by a laser resonator.

Depictions of the fundamental mode of a propagating Gaussian beam have not been limited to mathematical descriptions of the wave function but have also included geometrical interpretations of these wave functions. The reasons for a geometrical model are two-fold: first, to provide some physical insight into the actual wave phenomena and second, to predict the first-order properties of a Gaussian beam

as it propagates through a medium or an optical system. These first-order properties include waist dimensions and locations, wavefront radii and waist sizes at lenses, in pupil planes, and in image planes. To a large extent, the geometrical model of a Gaussian beam has been based on the traditional Kogelnik and Li mathematical expression. This description consists of a spherical wavefront, whose center varies in a nonlinear fashion along the axis, and contours of constant amplitude that are represented as hyperboloids of revolution of one sheet. Various methods have been proposed to trace the transformation of the beam size and wavefront curvature as given by this model through an optical system.

Deschamps and Mast [1964] observed that the transformation of the diameter and wavefront curvature of Gaussian beams through a sequence of lenses is equivalent to an impedance transformation through a reciprocal two-port network. In this case, refraction by a lens and transfer through sections of free space are analogous to parallel and series reactances, respectively. Equating Gaussian beam parameters to electrical impedance, a complex quantity, carried through to the complex beam parameter  $q$  of Kogelnik and Li [1966] as well as to the Gaussian Beam Chart introduced by Collins [1964] and expanded on by Li [1964]. The chart consists of two orthogonal families of circles and is similar to the Smith Chart used for calculating the characteristics of electromagnetic fields along a transmission line. Deschamps and Mast also outlined a technique for calculating the first-order properties of a Gaussian beam by tracing the lateral



foci of the amplitude hyperbola. This technique was later formalized by Laues [1967]. The transformation of the beam diameter and wave-front curvature can also be expressed in terms of the elements of the ray matrix relating ray positions and slopes at the input and output planes of a system. Known as the ABCD law for transforming a Gaussian beam, this method was introduced by Kogelnik [1965] and remains quite popular. Another method for tracing Gaussian beam properties in an optical system, proposed by Steier [1966], consisted of applying geometrical optics to the amplitude hyperbola as an envelope of rays. In this equivalent-ray-packet technique, the envelope gives the Gaussian beam spot size, and the curves perpendicular to the average ray slope provide the Gaussian beam phase fronts. Finally, Arnaud [1969] introduced the very useful concept of a complex ray to represent Gaussian beams. He presented a very convenient beam tracing method that portrayed a complex ray as two real rays that can be traced through an optical system by ordinary ray trace methods. The theory behind this method will be examined in Chapter 6. Herloski, Marshall, and Antos [1983] utilized this method in a conventional optical design program.

The mathematical model for a propagating Gaussian beam, which will be presented here, will utilize two simplifications found in all of the previous work on this subject. First, we shall confine the discussion to solutions of the scalar wave equation and second, we will concentrate on the fields produced by a resonator after it has reached a steady state. Since it is only these steady-state fields that propagate beyond the resonator, this latter restriction is

trivial. Basically, we do not concern ourselves with the resonator at all, only with the fields it generates. Limiting this presentation to the domain of scalar theory is a more serious restriction. It means that electromagnetic radiation will be treated as a scalar phenomenon; that is, that only the scalar amplitude of one transverse component of the electric or magnetic field will be considered. It is assumed that all other components can be treated independently in the same fashion. This approach ignores the fact that the electric and magnetic fields are coupled through Maxwell's equations and cannot be treated independently. In general, scalar theory yields very accurate results if the diffracting aperture (in this case, the resonator mirror) is large compared to the wavelength and if the fields of interest are not studied too close to the aperture. Both of these conditions will be met here but scalar theory runs into yet another difficulty when used to describe beams with a Gaussian amplitude distribution. Under the presumption of scalar theory, we assume that the electric field is polarized in the same direction everywhere. However, according to Maxwell's equations, the field must have a zero divergence in a charge-free region. As a result, there can be no spatial modulation of the field in its direction of polarization. In other words, as Mukunda, Simon and Sudarshan [1984] and Lax, Louisell and McKnight [1975] have shown, the spatial modulation of the field affects its polarization. Therefore, a scalar field with a Gaussian amplitude distribution cannot exist, let alone propagate. Nevertheless, we will

carry on with the pleasant fiction that such a beam does indeed exist as well as propagates since the development of a scalar theory is the best and most methodical starting point for developing Gaussian amplitude solutions to the vector wave equation.

Before beginning this development of a new mathematical model, we shall take an in-depth look at the historical developments that led up to Kogelnik and Li's traditional Gaussian beam model--both the mathematical expression and the geometrical interpretation. We shall examine the approximations, apart from the assumption of scalar theory, that went into their description. We shall also discuss the contradictions their geometrical model presents in the light of geometrical optics.

Chapter 3 will begin the derivation of the new mathematical model. This model will ultimately contain two parts--a mathematical expression for the fundamental mode of a propagating Gaussian beam and a geometrical interpretation of that expression. The wave function presented here is the zero-order mode of an entire family of exact, closed-form solutions to the scalar wave equation. Chapter 3 will concentrate on the development of this zero-order mode and compare it, term by term, to the traditional beam description of Kogelnik and Li. In Chapter 5 we shall extend the derivation to all of the higher-order terms of the family. Along the way, we shall point out a very simple but elegant means of obtaining Gaussian beam harmonics from spherical waves and spherical harmonics.

In Chapter 4 we present the geometrical interpretation of this new wave function. The model is couched in the oblate spheroidal coordinate system because of the ease with which wavefronts and amplitude contours can be identified with the coordinate surfaces. The oblate spheroidal coordinate system consists of a family of confocal ellipses and hyperbolas rotated about the semi-minor axis of the ellipse. The wavefront of the Gaussian beam is identified with a section of this oblate ellipsoid while the hyperboloid of revolution of one sheet becomes a contour of constant amplitude. Further, we use the skew-line generator of this hyperboloid as a ray-like element to facilitate the geometrical interpretation of the new mathematical expression. Finally, we shall use this skew line to build a non-orthogonal coordinate system in which to study straight-line propagation of Gaussian beams. The usefulness of this nonorthogonal coordinate system cannot be found in the design of optical systems using Gaussian beams. Rather, it will provide the necessary formalism with which to study Gaussian beam propagation in inhomogeneous or stratified media as well as Gaussian solutions to the vector wave equation.

The work presented here bears some similarity to recent work on complex point sources, complex rays and complex-argument solutions to the scalar wave equation. This previous work has much appeal in its ability to apply the power and simplicity of spherical waves to the Gaussian beam problem. We shall therefore devote Chapter 6 to a

comparison between previous research and the results given here.

Finally, we shall look at some ideas for future research in

Chapter 7. The conclusions we shall reach in this paper are notable as much for the answers they provide as the questions they pose.

CHAPTER 2  
 HISTORICAL DEVELOPMENT OF A MATHEMATICAL MODEL  
 FOR GAUSSIAN BEAMS

The traditional description of the fundamental mode of a Gaussian beam derives from the resonator mode analysis of Fox and Li [1961]. This analysis was a numerical calculation of the scalar field across the mirrors of a Fabry-Perot resonant cavity where an initially launched plane wave is reflected back and forth many times between the mirrors. Specifically, they used the Rayleigh-Sommerfeld diffraction integral to compute the Fresnel field  $u_p$  at the mirror  $p$  due to the illuminated aperture  $A$  of the other mirror  $a$ .

$$u_p = \frac{ik}{4\pi} \iint_A u_a \frac{e^{+ikR}}{R} (1 + \cos\theta) dS. \quad (2.1)$$

In this formulation,  $u_a$  is the field distribution across the mirror  $a$ ,  $k$  is the propagation constant of the medium,  $R$  is the distance from a point on the mirror  $a$  to the observation point, and  $\theta$  is the angle that  $R$  makes with the unit normal to mirror  $a$ . This approximation assumes that the dimensions of the mirror are large in terms of wavelength ( $R \gg \lambda$ ) and that the field is very nearly transverse electromagnetic and uniformly polarized in one direction.

A plane-wave disturbance is launched at one of the mirrors and allowed to traverse the resonator many times. After many transits, a steady state is reached where the fields across both mirrors are identical except for a complex constant that is independent of position. The diffraction integral becomes:

$$v = \gamma \iint_A v \frac{e^{+ikR}}{R} (1 + \cos\theta) \frac{ik}{4} ds, \quad (2.2)$$

where  $v$  is the field distribution across both mirrors and can be regarded as a normal mode of the cavity. The constant  $\gamma$  specifies the attenuation and propagation constant associated with this normal mode.

The authors used the method of successive approximations to solve the integral Equation (2.2) numerically for three resonator configurations: (1) rectangular plane mirrors, (2) circular plane mirrors, and (3) confocal spherical or paraboloidal mirrors. In the case of the confocal spherical mirrors, the calculation indicated that the  $TEM_{00}$  mode possessed a Gaussian amplitude distribution and a phase distribution that was coincident with the spherical mirrors.

Concurrently, Boyd and Gordon [1961] were seeking an analytical solution for the mode patterns of a resonator with symmetrically placed concave spherical mirrors. They solved the problem initially for the specific case of a confocal resonator and then generalized their results for mirrors of identical radius but spaced a distance  $L$  apart. Here again, the normal modes or

eigenfunctions of the confocal resonator, are obtained by requiring that the field distribution over one mirror reproduce itself within a constant over the opposite mirror. The resulting integral equation is given by Equation (2.2) where Boyd and Gordon assume the field distribution  $v$  to be a product of functions separable in  $x$  and  $y$ . The authors suppose further a square reflector of dimension  $2a$  which is small compared to the mirror spacing  $L$ , and thus  $\theta$  is very nearly zero. Figure 2.1 illustrates a confocal resonator with the transverse scale enlarged to show detail. Equation (2.2) becomes

$$f_m(x)g_n(y) = \gamma \int_{-a}^{+a} \int_{-a}^{+a} \frac{ik}{2\pi R} e^{+ikR} f_m(x')g_n(y') dx' dy' . \quad (2.3)$$

To evaluate the distance  $R$  in the exponential phase term, they used the binomial expansion. The expression for  $R$  is given by

$$R = \sqrt{L^2 + (x-x')^2 + (y-y')^2} , \quad (2.4)$$

or approximately,

$$R = L \left\{ 1 + \frac{(x-x')^2 + (y-y')^2}{2L^2} - \frac{(x-x')^2 + (y-y')^2}{8L^4} + \dots \right\} . \quad (2.5)$$

When  $L^3\lambda \gg a^4$ , the third term of this expansion is much less than one and can be ignored. Parabolic wavelets now replace the Huygens



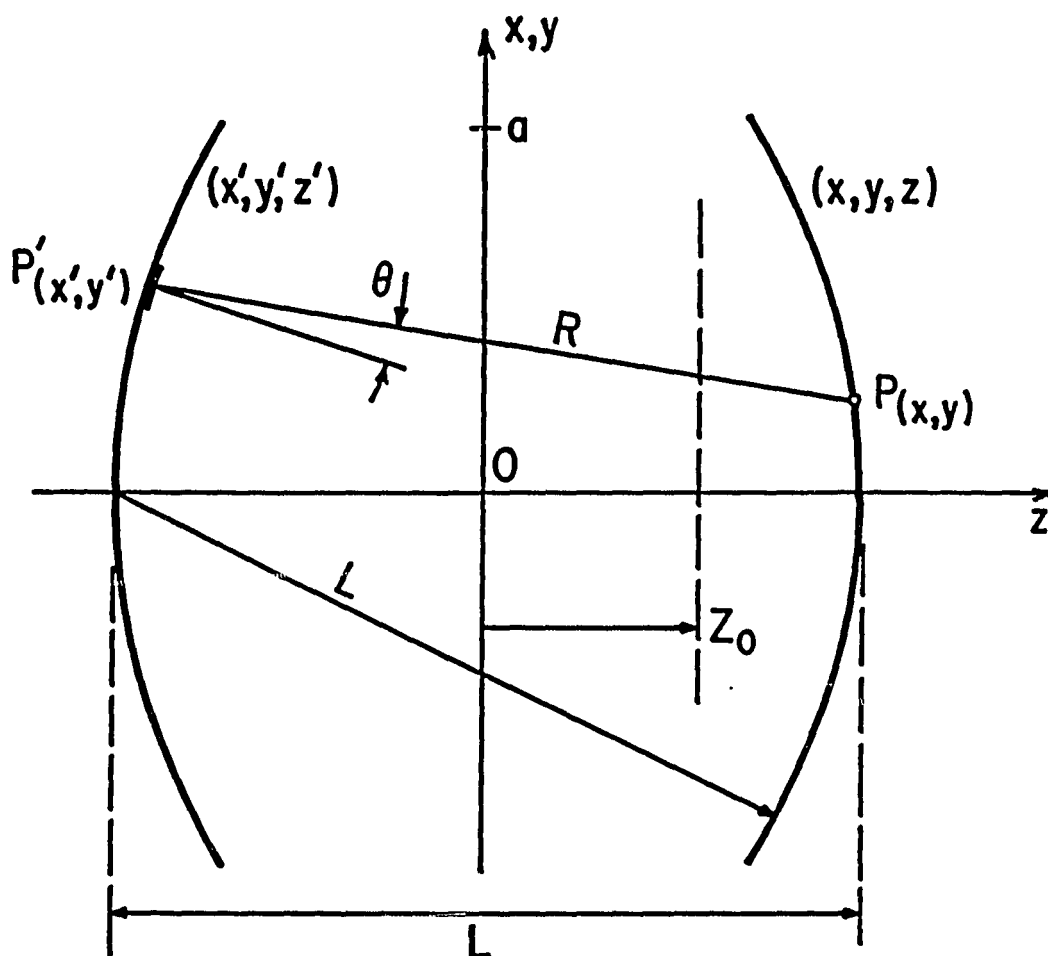


Figure 2.1 Confocal resonator geometry

(spherical) wavelets implicit in Equation (2.2). We can rewrite Equation (2.3) in terms of dimensionless variables where

$$c = \frac{a^2 k}{L} = 2\pi \frac{a^2}{L\lambda}; \quad X = \frac{x\sqrt{c}}{a}; \quad Y = \frac{y\sqrt{c}}{a}, \quad (2.6)$$

and

$$F_m(X) \triangleq f_m(x) \quad G_n(Y) \triangleq g_n(y) \quad (2.7)$$

Substituting these values into Equation (2.3) yields

$$F_m(X)G_n(Y) = \frac{\gamma i e^{+ikL}}{2\pi} \int_{-\sqrt{c}}^{+\sqrt{c}} F_m(X') e^{iXX'} dX' + \int_{-\sqrt{c}}^{+\sqrt{c}} G_n(Y') e^{iYY'} dY' \quad (2.8)$$

Thus, the field distribution across one mirror is proportional to the finite Fourier Transform of the distribution across the other. This is not a surprising result. Each mirror is a focusing element and imposes a quadratic phase factor on the incident field. Likewise, the Fresnel diffraction integral contains a quadratic phase factor due to each point on the diffracted wavefield propagating and expanding with distance  $z$ . Since the distance between the mirrors is exactly their radius, these phase factors cancel.

Slepian and Pollack [1961] considered solutions to the finite Fourier Transform,

$$F_m(X) = \frac{1}{X_m \sqrt{2\pi}} \int_{-\sqrt{c}}^{+\sqrt{c}} F_m(X') e^{+iXX'} dX' \quad (2.9)$$

They showed solutions to be

$$F_m(c, \eta) \propto S_{om}(c, \eta) \quad (2.10)$$

$$X_m = \frac{2c}{\pi} i^m R_{om}^{(1)}(c, 1), \quad m = 0, 1, 2, \quad (2.11)$$

where  $S_{om}(c, \eta)$  and  $R_{om}(c, 1)$  are, respectively, the angular and radial wave functions in prolate spheroidal coordinates as defined by Flammer [1957]. Further,  $\eta = x/a$  for  $F_m(X)$  and  $\eta = y/a$  for  $G_n(Y)$ . In the small angle approximation, that is  $\eta^2 \gg 1$ , Flammer shows that

$$\begin{aligned} F_m(X) &\approx \frac{\Gamma\left(\frac{m}{2} + 1\right)}{\Gamma(m+1)} H_m(X) e^{-\frac{1}{2} X^2} \\ &= \frac{\Gamma\left(\frac{m}{2} + 1\right)}{\Gamma(m+1)} (-1)^m e^{+\frac{1}{2} X^2} \frac{d^m}{dX^m} e^{-X^2} \end{aligned} \quad (2.12)$$

The mode shape is then approximately a Gaussian times a Hermite polynomial  $H_m(X)$ . The first three modes are then

$$\begin{aligned}
F_0(c, \eta) &= \frac{1}{2} e^{-\frac{1}{2} c \eta^2} = \frac{1}{2} e^{-\frac{\pi x^2}{L\lambda}} \\
F_1(c, \eta) &= \eta \sqrt{\pi c} e^{-\frac{1}{2} c \eta^2} = \pi x \sqrt{\frac{2}{\lambda L}} e^{-\frac{\pi x^2}{L\lambda}}, \quad (2.13) \\
F_2(c, \eta) &= (2c\eta^2 - 1) e^{-\frac{1}{2} c \eta^2} = \left(4 \frac{\pi x^2}{L\lambda} - 1\right) e^{-\frac{\pi x^2}{L\lambda}}.
\end{aligned}$$

In two dimensions, the eigenfunction solutions to Equation (2.8), and therefore the resonator mode patterns, are the spheroidal wave functions,  $S_{0m}(c, x/a)$  and  $S_{0m}(c, y/a)$ . These modes can be expressed in terms of the Hermite-Gaussian polynomials:

$$\begin{aligned}
F_m(X)G_n(Y) &= \frac{\Gamma\left(\frac{m}{2}+1\right)\Gamma\left(\frac{n}{2}+1\right)}{\Gamma(m+1)\Gamma(n+1)} H_m(X)H_n(Y) \cdot \\
&\exp\left(-\frac{\pi(x^2+y^2)}{L\lambda}\right) \quad (2.14)
\end{aligned}$$

These eigenfunction solutions exist on the reflecting surfaces. Since these solutions are real, the mirrors are surfaces of constant phase. This result agrees with the numerical calculations of Fox and Li.

Boyd and Gordon went on to expand their results for symmetrically located surfaces of constant phase inside and outside the resonator. Boyd and Kogelnik [1962] went further, defining mode patterns, beam shapes and sizes, and resonance conditions for

arbitrary resonator configurations, as well as identifying high-loss (unstable) and low-loss (stable) geometries. Regardless of the specifics, each analysis produced consistent conclusions for the fundamental mode pattern of a resonator: a Gaussian amplitude distribution across surfaces of constant phase that are approximately spherical. Expressed mathematically, the wave function at any point is given approximately by:

$$\exp \left( - \frac{x^2+y^2}{w_s^2} - \frac{ik(x^2+y^2)}{2R} \right) = \exp \left( - \frac{ik(x^2+y^2)}{2} \left[ \frac{1}{R} - \frac{2i}{kw_s^2} \right] \right) \quad (2.15)$$

where  $w_s$ , is the "spot size", or the distance from the axis at which the wave amplitude is  $1/e$  times the on-axis value, as illustrated in Figure 2.2, and  $R$  is the radius of curvature of the phase front. The form of the function in brackets remains the same for all surfaces, but  $R$  and  $w_s$  vary in such a way that a surface of constant amplitude describes a hyperboloid of one sheet about the  $z$  axis. A Gaussian beam cross section is illustrated in Figure 2.3.

Kogelnik and Li [1966] later developed a mathematical derivation to explain this behavior. The scalar wave equation can be written as

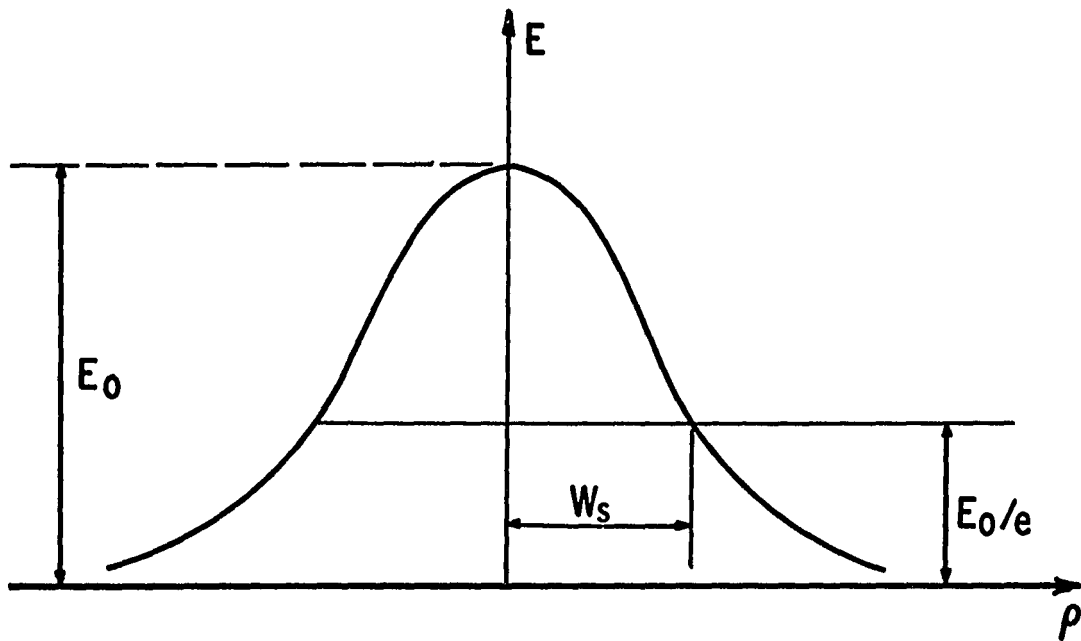


Figure 2.2 Gaussian amplitude distribution.

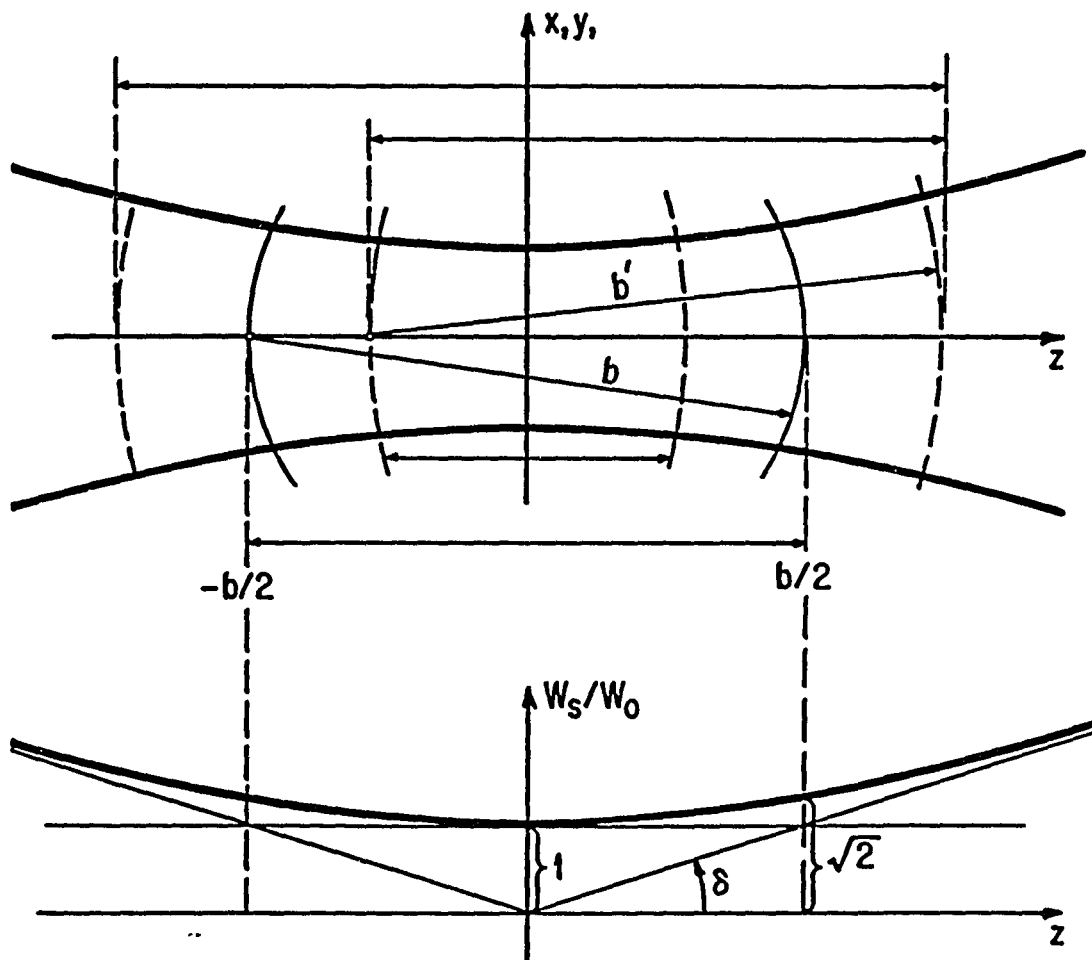


Figure 2.3 Hyperbolic contour of constant amplitude in Gaussian beam showing surfaces of constant phase

$$\nabla^2 u + k^2 u = 0 \quad , \quad (2.16)$$

where the field distribution  $u$ , traveling in the  $z$  direction, is

$$u = \psi(x, y, z) \exp(-ikz) \quad . \quad (2.17)$$

Kogelnik and Li postulated that the complex function  $\psi$  has a form similar to Equation (2.15), that is, a curved wavefront and a non-uniform intensity distribution along with complex phase shifts associated with propagation. They further assumed that  $\psi$  varies so slowly with  $z$  that its second derivative  $\partial^2 \psi / \partial z^2$  can be neglected. Therefore insertion of Equation (2.17) into (2.16) yields

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} - 2ik \frac{\partial \psi}{\partial z} = 0 \quad . \quad (2.18)$$

A solution to Equation (2.18) possessing the aforementioned properties can be written explicitly as

$$(x, y, z) = \exp \left\{ -i \left( P(z) + \frac{k(x^2 + y^2)}{2q(z)} \right) \right\} \quad . \quad (2.19)$$



The parameter  $P(z)$  represents the complex phase shift associated with beam propagation and  $q(z)$  is a complex beam parameter as identified by the function in brackets in Equation (2.15). Specifically,

$$\frac{1}{q(z)} = \frac{1}{R(z)} - \frac{2i}{kw(z)^2} , \quad (2.20)$$

where  $R(z)$  is the radius of wavefront curvature and  $w(z)$  is the spot size. Now we shall derive the explicit dependence upon  $z$  of both  $P(z)$  and  $q(z)$ .

After insertion of Equation (2.19) into (2.18) and comparing terms of equal powers of  $x^2 + y^2$ , one obtains the relations

$$\frac{\partial P(z)}{\partial z} = \frac{-i}{q(z)} , \quad (2.21)$$

and

$$\frac{\partial q(z)}{\partial z} = 1 . \quad (2.22)$$

Equation (2.22) indicates a linear rate of change of the complex beam parameter with respect to  $z$ , which can be written as

$$q_2(z) = q_1(z) + z , \quad (2.23)$$

where the beam parameter in an output plane  $q_2$  is related to the parameter in an input plane  $q_1$  by the separation  $z$  between the planes.

Using expressions (2.20) and (2.23), one can derive characteristic equations for  $R(z)$  and  $w(z)$ . The Gaussian beam contracts to a minimum beam diameter,  $2w_0$ , at a "beam waist" where the phase front is a plane. At this point,  $R(z) \rightarrow \infty$  and

$$\frac{1}{q_0(z)} = 0 - \frac{2i}{kw_0^2} \quad (2.24)$$

$$q_0(z) = \frac{ikw_0^2}{2} .$$

At a distance  $z$  from the plane of the waist, the equation for the complex beam parameter becomes

$$\frac{1}{q(z)} = \frac{1}{q_0(z)+z} = \frac{1}{R(z)} - \frac{2i}{kw(z)^2} . \quad (2.25)$$

Equating the real and imaginary terms yields

$$w^2(z) = w_0^2 \left[ 1 + \left( \frac{2z}{kw_0^2} \right)^2 \right] , \quad (2.26)$$

and

$$R(z) = z \left( 1 + \frac{kw_0^2}{2z} \right)^2 . \quad (2.27)$$

The beam contour  $w(z)$  is a hyperbola with asymptotes inclined to the axis at the far field diffraction angle for the fundamental mode,

$$\delta = \frac{2}{kw_0} . \quad (2.28)$$

Next, we examine the behavior of  $P(z)$ , the complex phase shift associated with propagation of the beam, with respect to  $z$ . Using the plane of the waist as the input plane and inserting (2.23) into (2.21) yields

$$\frac{P(z)}{z} = - \frac{i}{q(z)} = - \frac{i}{z + ikw_0^2/2} . \quad (2.29)$$

Integration of (2.29) generates an expression for  $P(z)$ ,

$$\begin{aligned} iP(z) &= \ell n \sqrt{1 + (2z/kw_0^2)^2} \\ &\quad - i \tan^{-1} \left( \frac{2z}{kw_0^2} \right) \\ &= \ell n \left( \frac{w(z)}{w_0} \right) - i \tan^{-1} \left( \frac{2z}{kw_0^2} \right) . \end{aligned} \quad (2.30)$$

Substituting (2.30) into (2.19) produces the result

$$\psi(x,y,z) = \frac{w_0}{w(z)} e^{i\Phi} e^{-ik\frac{(x^2+y^2)}{2q(z)}} \quad (2.31)$$

It now becomes apparent that the real part of  $P(z)$  represents a phase difference  $\Phi$  between a Gaussian beam wavefront and a plane wave, while the imaginary part produces an amplitude factor  $w_0/w(z)$ , giving the expected on-axis amplitude decrease due to expansion of the beam.

With these results, the fundamental field distribution  $u$  can be written

$$u(x,y,z) = \frac{w_0}{w(z)} \exp \left\{ -i \left( kz - \Phi + \frac{k(x^2+y^2)}{2R(z)} - \frac{(x^2+y^2)}{w^2(z)} \right) \right\} \quad (2.32)$$

where

$$\Phi = \tan^{-1} \left( \frac{2z}{kw_0^2} \right) \quad (2.33)$$

Figure 2.4 illustrates the field just described.

Finally, many authors (for example Harris, Tavener and Mitchell [1969] and Gaskill [1978]) identify a further beam characteristic called the Rayleigh range. This is defined to be the distance from the waist,  $z_0$ , at which the beam width has increased by a factor of  $\sqrt{2}$  and the radius of curvature has a magnitude of  $2z_0$ . Therefore, at  $z = z_0$ , we can write the beam width and radius as

$$w(z_0) = \sqrt{2}w_0 \quad (2.34)$$

$$R(z_0) = 2z_0 \quad ,$$

where

$$z_0 = \frac{kw_0^2}{2} \quad (2.35)$$

It is a simple matter to show that the characteristic function for the radius of curvature  $R(z)$  has a minimum at the Rayleigh range.

#### Approximations, Interpretations and Limitations

Equation (2.32) describes a scalar field possessing a Gaussian amplitude distribution, a paraboloidal wavefront and a phase draw-up or slip in the position of the phase front as the beam propagates. While this may be far from the ideal behavior of an electromagnetic wave, it should be kept in mind that all of the analyses hitherto used to derive a mathematical description of a Gaussian beam use the Fresnel approximations to part of the Rayleigh-Sommerfeld diffraction kernel. These approximations are generally employed in diffraction problems to make the integral computations and mathematical manipulations easier. Specifically, they include the paraxial approximation, a truncation of the binomial expansion and the usual scalar diffraction approximation where aperture dimensions are large compared to wavelength.

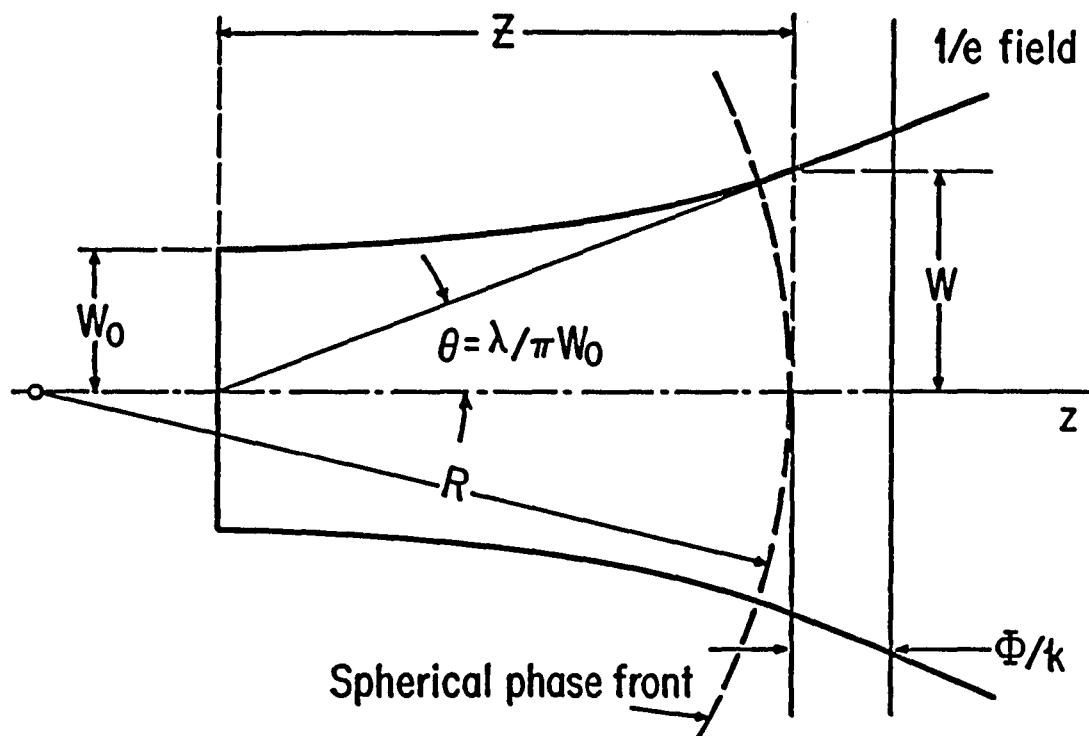


Figure 2.4 Contour of the fundamental mode of a Gaussian beam

The first two approximations limit the validity of any solution to a small region near the axis. In this area, sines and tangents of angles are very nearly indistinguishable from each other as well as from the angles themselves. This leads to certain ambiguities in the classical beam description which will be discussed in Chapter 3. In addition, truncation of the binomial expansion further restricts the functional description of the beam shape by replacing spheres with parabolas. The first-order approximation to a sphere is a parabola, and therefore the mathematical model as written in Equation (2.32) cannot differentiate between spheres and parabolas nor, for that matter, between spheres and a family of ellipsoids.

In the Rayleigh-Sommerfeld diffraction formula, the field disturbance at an output plane is found by convolving the disturbance at an input plane or aperture with the point spread function. This function is the normal derivative of a spherical wave and has the form:

$$\hat{n} \cdot \nabla \left( \frac{e^{ikr}}{kr} \right) = (\hat{n} \cdot \hat{r}) k \left[ i - \frac{1}{kr} \right] \frac{e^{ikr}}{kr}, \quad (2.36)$$

which is an elementary wavelet generated by an oscillating quadrupole. The Rayleigh-Sommerfeld diffraction integral expands the field across an aperture into a summation of elementary wavelets generated by quadrupoles.

At this point, it is usually assumed that the distance between the output plane and the aperture is much greater than one wavelength,  $r \gg \lambda$ , and the spread function becomes,

$$(\hat{n} \cdot \hat{r})k \left[ i - \frac{1}{kr} \right] \frac{e^{ikr}}{kr} \approx (\hat{n} \cdot \hat{r}) \frac{ie^{ikr}}{r} . \quad (2.37)$$

The resulting function has the form of a spherical wave with a  $90^\circ$  phase shift. By making the above assumption, we approximate the derivative of a function with the function itself. Further, we substitute an oscillating dipole, which generates spherical waves, for the quadrupole as the basis source in the diffraction integral. Kuper [1983] has shown that the quadrupole wavelet, not the dipole, reduces to a delta function in the plane of the aperture and therefore reproduces the assumed boundary conditions. It has also been shown by Shack [1983] that making this approximation eliminates some of the propagating terms, as well as the evanescent ones, from the diffracted wavefield. Clearly, the approximations inherent in the diffraction analysis as it has been performed historically restrict how precisely the classical model of a Gaussian beam predicts its propagation.

At best, then, the current description can be termed an approximation to the electromagnetic wave phenomena of diffraction. Therefore, we can explore how well the model agrees with another approximation, that of geometrical optics, as a first-order approach to predicting its characteristics upon propagation as well as the beam's behavior in an optical system. Doing so, however, reveals further inconsistencies in the beam description. One can infer that equal amplitude points on successive wavefronts lie along paths or trajectories that coincide with the direction of radiant energy.



Geometrical optics stipulates that in a homogeneous medium this direction of radiant energy is a straight line, since by symmetry, the path cannot be bent in any preferred direction, there being none. Unfortunately, this path is a hyperboloid of one sheet as a contour of constant amplitude is defined by the traditional model. Geometrical optics also maintains that a phase front will propagate along orthogonal trajectories. Again, the current model falls short since neither a sphere nor a parabola, as the wavefront is described, constitutes an orthogonal function to a hyperboloid. A set of confocal ellipses forms the orthogonal family to such a figure. For a hyperboloid of one sheet rotated about the  $z$  axis with respect to foci symmetrically located at  $\rho = \pm d$ , the perpendicular family will be oblate ellipsoids with foci also at  $\rho = \pm d$ . An oblate ellipsoid is an ellipse rotated about its semi-minor axis, in this case the  $z$  axis.

In addition to the contradictions posed by the nonorthogonal wavefronts and curved trajectories, the center of curvature of the expanding wavefronts displays some odd properties for a geometrical model. The position of the center of curvature, that is the location of the point source, varies in a nonlinear fashion along the axis with the asymptotic wavefront curvature centered at the waist. Likewise the radius of curvature changes nonlinearly, decreasing to a minimum at a distance equal to the Rayleigh range from the waist and then increasing again. Indeed, this behavior is consistent with the

radius of vertex curvature for an oblate ellipsoid with foci equal to the Rayleigh range. Coupled with the aforementioned peculiarities of the present model, this fact lends credence to the theory of an ellipsoidal wavefront for a Gaussian beam.

#### Historical References to Ellipsoidal Wavefronts

Deschamps [1971] suggested a family of ellipses confocal with the amplitude hyperbola as the exact expression for the Gaussian beam wavefront. He did so in a paper relating the field of a Gaussian beam to a spherical wave having a complex source. Arnaud [1969] originally proposed complex rays resulting from a point source whose position is complex as a means of describing the propagation of Gaussian beams. Since then, Arnaud [1985] and others (Shin and Felsen [1977], Keller and Streifer [1971]) have developed an extensive theory pertaining to the use of complex-point-source rays to describe a geometrical optics model for a propagating Gaussian beam. Indeed, this idea has received so much attention in the literature of the last ten years that its relationship to this work will be discussed in Chapter 6.

Vainshtein [1964] was the first to consider a wave function having an oblate ellipsoidal wavefront in a resonator. After the manner of Boyd and Gordon, he was studying the electromagnetic oscillations within an open resonator formed by identical circular or

rectangular spherical mirrors. He obtained formulas for the oscillation frequency and field distribution, interpreting each of the oscillations as a set of rays alternately reflected by the mirrors and restricted by a caustic surface. By expressing the wavefront as an ellipse, the formulas for the field distribution within the resonator are general enough to allow for changes in mirror curvature.

Vainshtein began with plane mirrors and gradually increased the curvature, leaving their dimensions and separation distance constant, while studying the effects of the change on the caustic surfaces and on the radiative losses within the resonator. Increasing the mirror curvature from a plane matches the variation of foci separation in an ellipse from infinity. In this situation, a concentric resonator results when the foci separation equals zero. Vainshtein was using the oblate ellipse as a general spheroid for his purposes and was not investigating an elliptical wavefront model.

Ito [1973] extends this idea further to discuss Gaussian beams in the context of ellipsoidal wave functions which include prolate and oblate spheroidal wave functions and elliptic cylinder functions. He suggests that since oblate spheroidal wave functions tend to the Laguerre-Gaussian polynomials in the short wavelength limit, then ellipsoidal waves might tend to a more general form of a class of beams in the short wavelength limit, including the Gaussian beam as a special case. Like Vainshtein, Ito does not investigate the significance of an elliptical wavefront but uses the ellipsoidal coordinates

as a more general system in which to express a broad class of beams emanating from dipoles and quadrupoles as well as aggregates of these and including beams with a Gaussian amplitude distribution as a special case.

#### Summary

We have seen how the mathematical model currently used to describe a beam with a Gaussian amplitude distribution developed. In particular, we have discussed the approximations peculiar to the diffraction analysis and the limitations and contradictions in the model when it is interpreted in the light of geometrical optics. In the next chapter, we begin the development of a new scalar model, devoid of approximations, which consists of both a new mathematical expression for the fundamental mode of a Gaussian beam and a new geometrical interpretation of this formulation expressed in the oblate spheroidal coordinate system.

## CHAPTER 3

AN EXACT ZERO-ORDER SOLUTION TO THE SCALAR WAVE  
EQUATION IN OBLATE SPHEROIDAL COORDINATES

In the past, mathematical descriptions of the fundamental mode of a propagating Gaussian beam have been expressed in a variety of coordinate systems. The expression introduced by Kogelnik and Li [1966] and used most frequently utilizes the Cartesian coordinate system to depict beams with rotational symmetry. As discussed at the end of the previous chapter, other authors have used the oblate spheroidal coordinate system because of the simplicity of modeling a contour of constant amplitude in the beam as a hyperboloid of one sheet, which is one of the oblate spheroidal coordinate surfaces. It is also possible to obtain exact solutions to the scalar Helmholtz equation in this coordinate system, as shown by Einziger and Raz [1987]. The solutions to the wave equation that will be developed here are similar to the results of Einziger and Raz, but the physical interpretations are radically different and the results include an entire family of higher-order solutions which will be presented in Chapter 5.

Most methods of determining solutions to the scalar Helmholtz equation begin by assuming a general solution that is separable in the coordinate system of choice. In the prolate and oblate spheroidal coordinate systems in particular, the definitive work on solutions to the wave equation is contained in Carson Flammer's Spheroidal Wave

Functions [1957]. The solutions presented in this chapter and in Chapter 5, although expressed in these coordinate systems, are not separable in either system. While this does represent an additional level of complexity, the resultant solutions have analytic forms instead of an infinite series as for Flammer's separable solutions. Furthermore, Flammer's method of solving the wave equation requires different solutions for different values of the product  $kd$ , where  $k$  is the propagation constant of the medium and  $2d$  is the focus spacing for the prolate or oblate spheroidal coordinates. In contrast, the exact solutions presented here are valid regardless of the value of the  $kd$  product.

Two families of closed-form, nonseparable solutions to the scalar Helmholtz equation are presented here: one in the prolate spheroidal coordinate system and the other in the oblate spheroidal coordinate system. The families are related by a "transformation", introduced by Flammer, which consists of two parameter substitutions. The first is to substitute  $\pm i\xi$  in the oblate spheroidal system, for the prolate spheroidal coordinate  $\xi$ ; the second is the replacement of  $ikd$  in the prolate spheroidal system with  $\pm kd$  in the oblate spheroidal one. Note that this differs from a simple coordinate transformation; furthermore, we perform the substitution or transmutation on both the Helmholtz operator and the wave function. This generates an entirely new and different family of solutions to the wave equation in the oblate spheroidal coordinate system from a family of solutions in the prolate spheroidal system. Both families are exact solutions to the wave equation, but they are radically different in form.

Specifically, we transmute a spherical wave with a real point source on axis into a Gaussian beam.

### Spheroidal Coordinate Systems

The prolate spheroidal coordinate system is formed by rotating a system of confocal ellipses and hyperbolas about the major axis of the ellipse. The axis of rotation is the z axis, and the foci occur on this axis, symmetrically located about the origin with a spacing of  $2d$ . The prolate spheroidal coordinates, shown in Figure 3.1, are related to rectangular coordinates by a set of parametric equations given by:

$$\begin{aligned}x &= d \sinh \mu \sin \theta \cos \phi , \\y &= d \sinh \mu \sin \theta \sin \phi , \\z &= d \cosh \mu \cos \theta ,\end{aligned}\tag{3.1}$$

$$\text{where } 0 \leq \theta \leq \pi, 0 \leq \mu < \infty, 0 \leq \phi \leq 2\pi .\tag{3.2}$$

Following Flammer's notation, we let  $\xi = \cosh \mu$  and  $\eta = \cos \theta$

. The parametric equations then become

$$\begin{aligned}x &= d (\xi^2 - 1)^{1/2} (1 - \eta^2)^{1/2} \cos \phi , \\y &= d (\xi^2 - 1)^{1/2} (1 - \eta^2)^{1/2} \sin \phi , \\z &= d \xi \eta ,\end{aligned}\tag{3.3}$$

$$\text{where } -1 \leq \eta \leq 1, 1 \leq \xi < \infty, 0 \leq \phi \leq 2\pi .\tag{3.4}$$

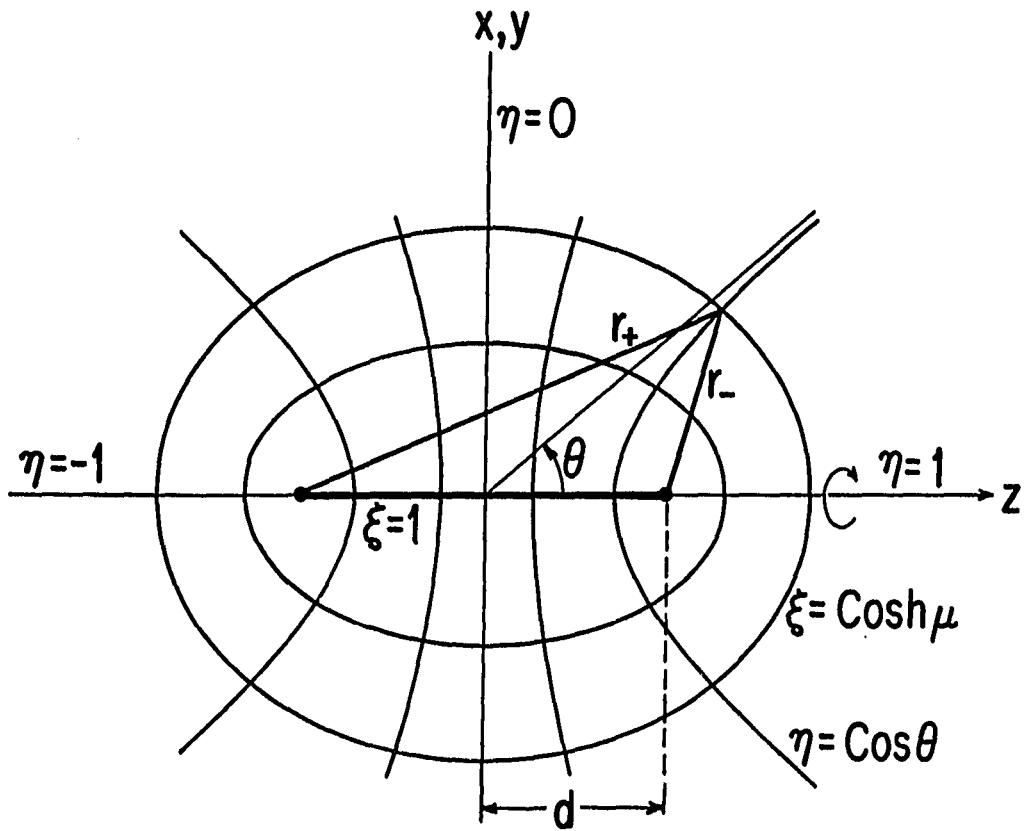


Figure 3.1 Prolate spheroidal coordinate system



The surface  $\xi = \text{constant} > 1$  is a prolate ellipsoid with a major axis of length  $2d\xi$  and a minor axis of length  $2d \sinh\mu$ . The degenerate surface  $\xi = 1$  is a straight line along the  $z$  axis from  $z = -d$  to  $z = +d$ . The surface  $|\eta| = \text{constant} < 1$  is a hyperboloid of revolution of two sheets whose asymptotes pass through the origin, inclined at an angle  $\theta = \cos^{-1}\eta$  to the  $z$  axis. The degenerate surface  $|\eta| = 1$  is that part of the  $z$  axis for which  $|z| > d$ . Finally, the surface  $\phi = \text{constant}$  is the azimuthal plane containing the  $z$  axis.

The oblate spheroidal coordinate system, shown in Figure 3.2, is formed by rotating a similar system of mutually orthogonal ellipses and hyperbolas about the minor axis of the ellipse. Again, the  $z$  axis is the axis of rotation, but now the focus is a ring of radius  $d$  in the  $x$ - $y$  plane. The parametric equations relating the oblate spheroidal coordinate system to Cartesian coordinates are

$$\begin{aligned} x &= d \cosh\mu \sin\theta \cos\phi , \\ y &= d \cosh\mu \sin\theta \sin\phi , \\ z &= d \sinh\mu \cos\theta . \end{aligned} \tag{3.5}$$

with either

$$0 \leq \theta \leq \pi, \quad 0 \leq \mu < \infty, \quad 0 \leq \phi \leq 2\pi , \tag{3.6a}$$

or

$$0 \leq \theta \leq 2\pi, \quad -\infty < \mu < \infty, \quad 0 \leq \phi \leq 2\pi . \tag{3.6b}$$

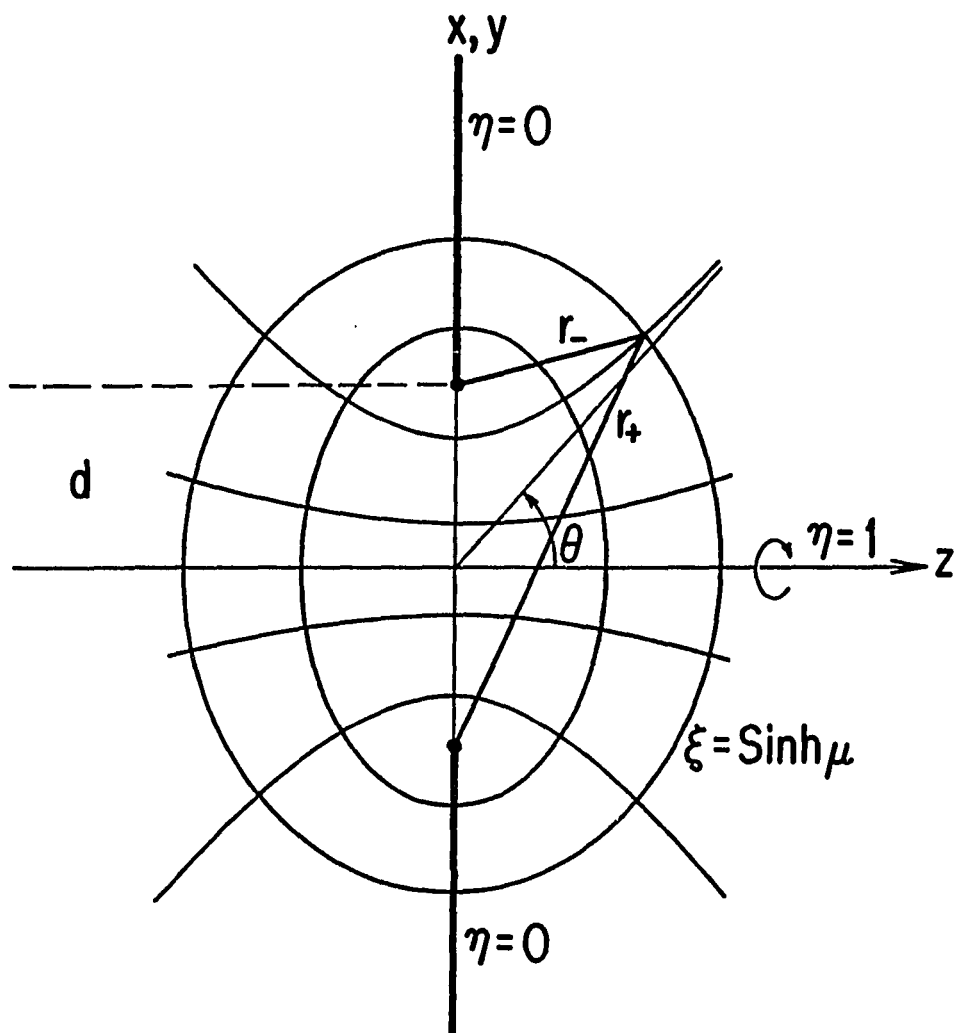


Figure 3.2 Oblate spheroidal coordinate system

In the oblate case, we let  $\xi = \sinh\mu$  and  $\eta = \cos\theta$ . The parametric equations then become

$$\begin{aligned}x &= d (1 + \xi^2)^{1/2} (1 - \eta^2)^{1/2} \cos\phi , \\y &= d (1 + \xi^2)^{1/2} (1 - \eta^2)^{1/2} \sin\phi , \\z &= d\xi\eta ,\end{aligned}\tag{3.7}$$

with either

$$-1 \leq \eta \leq 1 , 0 \leq \xi < \infty , 0 \leq \phi \leq 2\pi\tag{3.7a}$$

or

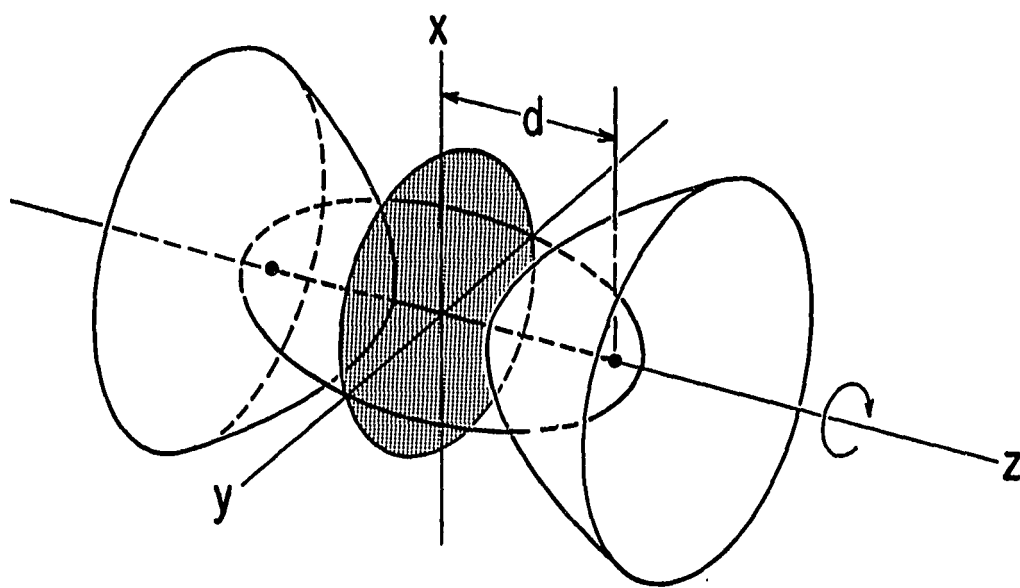
$$0 \leq \eta \leq 1 , -\infty < \xi < \infty , 0 \leq \phi \leq 2\pi .\tag{3.7b}$$

In the oblate system, the surface  $|\xi| = \text{constant} > 0$  is an oblate ellipsoid with major axis of length  $2d \cosh\mu$  and minor axis of length  $2d |\sinh\mu|$ . The surface  $\xi = 0$  is a circular disk of radius  $d$  centered at the origin in the  $x$ - $y$  plane. The surface  $|\eta| = \text{constant} < 1$  is a hyperboloid of revolution of one sheet whose asymptotes pass through the origin inclined at an angle  $\theta = \cos^{-1}\eta$  with the  $z$  axis. The degenerate surface  $\eta = 1$  is the  $z$  axis. The surface  $\eta = 0$  is the  $x$ - $y$  plane except for the circular disk  $\xi = 0$ . Finally, the surface  $\phi = \text{constant}$  is again the azimuthal plane containing the  $z$  axis. The angle  $\phi$  is measured from the  $x$ - $z$  plane.

While it is tempting to think of the difference between the prolate and oblate systems as a simple rotation or flattening of the ellipsoids, such is not the case. As Figure 3.3 and Figure 3.4 demonstrate, the foci in the prolate spheroidal system are two distinct points on the  $z$  axis, whereas the focus in the oblate system is a ring in a plane perpendicular to the  $z$  axis. Furthermore, the hyperboloids of revolution consist of two separate sheets in the prolate case but is a single, continuous sheet in the oblate. Both, however, are systems of orthogonal curvilinear coordinates, i.e., the tangent planes of the three surfaces passing through any point in space are mutually perpendicular. The confocal ellipses and hyperbolas for both the prolate and oblate systems can be defined in terms of the usual definitions of ellipses and hyperbolas. That is, an ellipse is the locus of a point  $P$  which moves such that the sum of its distances to the foci remains constant. Similarly, a hyperbola is the locus of  $P$  which moves such that the difference of its distances to the foci remains constant. In Figures 3.1 and 3.2, these distances are shown as  $r_+$  and  $r_-$ . In the prolate case, these distances are

$$r_+ = [(z + d)^2 + x^2 + y^2]^{1/2}, \tag{3.8}$$

$$r_- = [(z - d)^2 + x^2 + y^2]^{1/2} .$$



**Figure 3.3** Three-dimensional view of prolate spheroidal coordinate system showing focal points

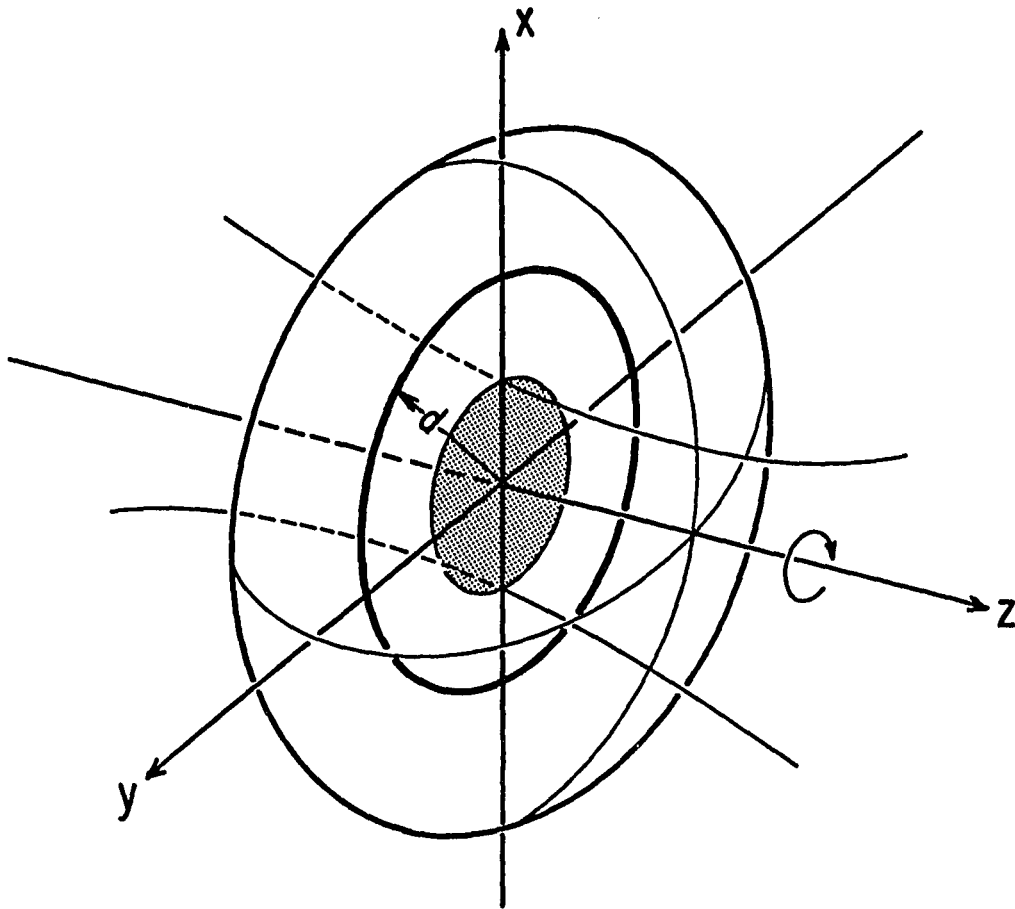


Figure 3.4 Three-dimensional view of oblate spheroidal coordinate systems showing focal ring

Substituting in the parametric equations given by Eq. (3.3) results in

$$r_+ = d[1 + \cosh^2\mu \cos^2\theta + 2 \cosh\mu \cos\theta + \sinh^2\mu \sin^2\theta]^{1/2} , \quad (3.9)$$

$$r_- = d[1 + \cosh^2\mu \cos^2\theta - 2 \cosh\mu \cos\theta + \sinh^2\mu \sin^2\theta]^{1/2} .$$

By use of the trigonometric identities,

$$\cosh^2\mu - \sinh^2\mu = 1 , \quad (3.10a)$$

$$\cos^2\theta + \sin^2\theta = 1 , \quad (3.10b)$$

Equation (3.9) reduces to

$$r_+ = d[\cosh\mu + \cos\theta] = d(\xi + \eta) , \quad (3.11)$$

$$r_- = d[\cosh\mu - \cos\theta] = d(\xi - \eta) .$$

The same procedure can be followed in the oblate spheroidal coordinate system, although, since the focus is a ring,  $r_+$  and  $r_-$  are drawn in the same azimuthal plane. Therefore,  $r_+$  and  $r_-$  are

$$r_+ = [(\rho + d)^2 + z^2]^{1/2} , \quad (3.12)$$

$$r_- = [(\rho - d)^2 + z^2]^{1/2} ,$$

where  $\rho = \sqrt{x^2 + y^2}$  and  $\phi$  is assumed constant. Substituting in the parametric equations for  $\rho$  and  $z$  given by Equation (3.5) leads to

$$r_+ = d[1 + \cosh^2\mu \sin^2\theta + 2\cosh\mu \sin\theta + \sinh^2\mu \cos^2\theta]^{1/2} , \quad (3.13)$$

$$r_- = d[1 + \cosh^2\mu \sin^2\theta - 2\cosh\mu \sin\theta + \sinh^2\mu \cos^2\theta]^{1/2} .$$

Again, by using the trigonometric identities in Equation (3.10),

Equation (3.13) becomes

$$r_+ = d[\cosh\mu + \sin\theta] ,$$

$$r_- = d[\cosh\mu - \sin\theta] . \quad (3.14)$$

Expressed in terms of  $\xi$  and  $\eta$ ,  $r_+$  and  $r_-$  are

$$r_+ = d \left( (1 + \xi^2)^{1/2} + (1 - \eta^2)^{1/2} \right) , \quad (3.15)$$

$$r_- = d \left( (1 + \xi^2)^{1/2} - (1 - \eta^2)^{1/2} \right) .$$

In both the prolate and oblate systems, the sum of  $r_+$  and  $r_-$  leads to a surface which is a function of  $\xi$  only--the ellipse. Likewise, the difference between  $r_+$  and  $r_-$  leads to a surface dependent on  $\eta$  alone--the hyperbola. In the prolate system these surfaces are given by

$$r_+ + r_- = 2d\xi , \quad (3.16a)$$

$$r_+ - r_- = 2d\eta ,$$

and in the oblate system they are expressed as

$$r_+ + r_- = 2d (1 + \xi^2)^{1/2} , \quad (3.16b)$$

$$r_+ - r_- = 2d (1 - \eta^2)^{1/2} .$$

#### Zero-Order Solutions to the Scalar Wave Equation

The scalar wave equation

$$(\nabla^2 + k^2)\psi = 0 , \quad (3.17)$$



is separable in eleven coordinate systems, including the prolate and oblate spheroidal coordinate systems. For an orthogonal coordinate system, the Laplacian is given by

$$\nabla^2 = \frac{1}{h_1 h_2 h_3} \left[ \frac{\partial}{\partial u_1} \left( \frac{h_2 h_3}{h_1} \frac{\partial}{\partial u_1} \right) + \frac{\partial}{\partial u_2} \left( \frac{h_1 h_3}{h_2} \frac{\partial}{\partial u_2} \right) + \frac{\partial}{\partial u_3} \left( \frac{h_2 h_1}{h_3} \frac{\partial}{\partial u_3} \right) \right], \quad (3.18)$$

The factors  $h_i$  are the metrical coefficients defined by

$$dx^2 + dy^2 + dz^2 = h_1^2 du_1^2 + h_2^2 du_2^2 + h_3^2 du_3^2. \quad (3.19)$$

In the prolate spheroidal coordinate system, these coefficients are

$$h_\xi = d \frac{(\xi^2 - \eta^2)^{1/2}}{(\xi^2 - 1)^{1/2}},$$

$$h_\eta = \frac{d(\xi^2 - \eta^2)^{1/2}}{(1 - \eta^2)^{1/2}}, \quad (3.20)$$

$$h_\phi = d (1 - \eta^2)^{1/2} (\xi^2 - 1)^{1/2}$$

After substituting these factors in the expression for the Laplacian, Helmholtz's equation becomes

$$\left[ \frac{\partial}{\partial \xi} (\xi^2 - 1) \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta} (1 - \eta^2) \frac{\partial}{\partial \eta} + \frac{\xi^2 - \eta^2}{(\xi^2 - 1)(1 - \eta^2)} \frac{\partial^2}{\partial \phi^2} + k^2 d^2 (\xi^2 - \eta^2) \right] \psi = 0 \quad (3.21)$$

As Flammer points out, it is possible to obtain Helmholtz's equation in oblate spheroidal coordinates from Equation (3.21) by making the transformation

$$\xi_{P/S} \rightarrow \pm i \xi_{O/S} \quad (3.22)$$

$$i k d_{P/S} \rightarrow \pm k d_{O/S}$$

from prolate to oblate coordinates respectively. The subscript (P/S) denotes prolate spheroidal coordinates and the subscript (O/S) denotes oblate spheroidal coordinates.

The scalar wave equation in oblate spheroidal coordinates becomes

$$\left[ \frac{\partial}{\partial \xi} (\xi^2 + 1) \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta} (1 - \eta^2) \frac{\partial}{\partial \eta} + \frac{\xi^2 + \eta^2}{(\xi^2 + 1)(1 - \eta^2)} \frac{\partial^2}{\partial \phi^2} + k^2 d^2 (\xi^2 + \eta^2) \right] \psi = 0 \quad (3.23)$$

This result agrees, of course, with that obtained by substituting the metrical coefficients for the oblate spheroidal coordinate system into the expression for the Laplacian. For reference, these metrical coefficients are

$$\begin{aligned}
h_{\xi} &= d \frac{(\xi^2 + \eta^2)^{1/2}}{(1 + \xi^2)^{1/2}}, \\
h_{\eta} &= d \frac{(\xi^2 + \eta^2)^{1/2}}{(1 - \eta^2)^{1/2}}, \\
h_{\emptyset} &= d (1 + \xi^2)^{1/2} (1 - \eta^2)^{1/2}.
\end{aligned} \tag{3.24}$$

A spherical wave can be represented by a zero-order spherical Hankel function of the first kind, which is a well-known solution to the scalar Helmholtz equation. Let

$$\psi_{P/S} = e^{-ikd} \frac{e^{ikr_+}}{ikr_+}, \tag{3.25}$$

where  $r_+$  is the radial distance from the focal point  $z = -d$  in the prolate spheroidal coordinate system as given by Equation (3.11). In addition, we have added a constant factor,  $e^{-ikd}$ , to the wave function. The purpose for this will become clear, but for now it does not affect the nature of  $\psi_{P/S}$  as an exact solution to the wave equation. Finally, we can interpret Equation (3.25) as a spherical wave expanding outward from the focal point at  $z = -d$ , expressed in prolate spheroidal coordinates.

Let us now apply the transformation of Equation (3.22) to both the Helmholtz operator  $\nabla^2 + k^2$  and the wave function  $\psi_{P/S}$ . Equation (3.25) becomes

$$e^{-ikd} \frac{e^{ikd(\eta+\xi)}}{ikd(\eta+\xi)} \rightarrow \frac{e^{-kd} e^{kd(\eta+i\xi)}}{kd(\eta+i\xi)} \quad (3.26)$$

In addition, because of the transformation of the operator  $(\nabla^2 + k^2)$ , the scalar wave equation is that given by Equation (3.23) for the oblate spheroidal coordinate system. We have, in effect then, transformed the entire Helmholtz equation, operator and wave function, from the prolate to the oblate spheroidal coordinate system. In other words

$$\begin{aligned} (\nabla^2 + k^2)_{P/S} &\rightarrow (\nabla^2 + k^2)_{O/S} \\ \xi_{P/S} &\rightarrow i\xi_{O/S} \\ ikd_{P/S} &\rightarrow kd_{O/S} \end{aligned} \quad (3.27)$$

We can now interpret the wave function on the right in Equation (3.26) as a function in oblate spheroidal coordinates. We rewrite it as

$$\psi_{O/S} = e^{ikd\xi} \frac{e^{-kd(1-\eta)} e^{-i \tan^{-1} \xi/\eta}}{kd (\eta^2 + \xi^2)^{1/2}}, \quad (3.28)$$

where  $k$  is the propagation constant,  $d$  is the radius of the ring focus,  $\xi = \text{constant}$  is an oblate ellipse, and  $\eta = \text{constant}$  is a hyperboloid of revolution of one sheet.

The function  $\psi_{0/S}$  describes a wave with a wavefront that is nominally a section of an oblate ellipsoid ( $\xi = \text{constant}$ ). This wavefront is modified by the term  $\exp[-i \tan^{-1} \xi/\eta]$  as  $\eta$  varies from 1 to 0. In fact, this arctangent factor has a simple but intriguing geometrical interpretation that will be discussed in Chapter 4. The exponential amplitude  $\exp[-kd(1-\eta)]$  specifies the amplitude distribution on a surface perpendicular to the direction of propagation. In the paraxial limit, this term reduces to the traditional Gaussian amplitude distribution as shown in detail below. Further, the amplitude factor  $[\eta^2 + \xi^2]^{-1/2}$  ensures that the wave energy falls to zero in the limit as the wave propagates to infinity. That is, as  $z \rightarrow \infty$ ,  $\xi \rightarrow \infty$  and  $\psi_{0/S} \rightarrow 0$ . This factor equals  $(r_+r_-)^{1/2}/d$  where  $r_+$  and  $r_-$  are as given by Equation (3.15).

Equation (3.28) is essentially the same as Equation (21) in the paper by Einziger and Raz [1987]. However, Equation (21) in that paper contains an error. The sign of the term  $(\omega_0/\nu) \cdot a \cdot \cos \eta$  in the last line of the equation should be positive. Careful insertion of Equation (20) into Equation (18) will reveal this to be true. Furthermore, it is obvious that if the magnitude of  $u(\tau, r, r')$  is to have a maximum on axis,  $\eta = 0$ , the argument of the exponential amplitude must be zero. Additional differences between Equation (3.28) here and Einziger and Raz are in the notation used for the oblate spheroidal coordinates, the modifying coefficients, and the use here of complex exponential notation.

We now show why Flammer's substitution enables the same conversion of a spherical wave to a Gaussian beam as the complex-point-source approach. Consider the oblate spheroidal coordinate system with origin at  $(0,0,-id)$ . The parametric equations of Equation (3.7) become

$$\begin{aligned}x &= d_{O/S} (1 + \xi_{O/S}^2)^{1/2} (1 - \eta^2)^{1/2} \cos\phi , \\y &= d_{O/S} (1 + \xi_{O/S}^2)^{1/2} (1 - \eta^2)^{1/2} \sin\phi , \\z &= d_{O/S} \xi_{O/S} \eta - id_{O/S} .\end{aligned}\tag{3.29}$$

We now make the substitution

$$\begin{aligned}\xi_{O/S} &\rightarrow \mp i \xi_{P/S} , \\id_{O/S} &\rightarrow \mp d_{P/S} ,\end{aligned}\tag{3.30}$$

which is Equation (3.22) in reverse, and Equation (3.29) becomes

$$\begin{aligned}x &= d_{P/S} (\xi_{P/S}^2 - 1)^{1/2} (1 - \eta^2)^{1/2} \cos\phi . \\y &= d_{P/S} (\xi_{P/S}^2 - 1)^{1/2} (1 - \eta^2)^{1/2} \sin\phi , \\z &= d_{P/S} (\xi_{P/S} \eta - 1) .\end{aligned}\tag{3.31}$$

These are the parametric equations for the prolate spheroidal coordinate system whose origin has been shifted to the real point  $(0,0,-d)$ . This is the location of the prolate spheroidal focus at  $z = -d$ . We can now consider spherical waves of radius  $r_+$  given by Equation (3.11) and with a real point source on axis.

The transformation of Equation (3.22) has enabled us to convert a spherical wave to a wave with a Gaussian amplitude distribution without resorting to either the concept of complex point sources or the paraxial approximation. The function  $\psi_{0/s}$  is an exact solution to the scalar Helmholtz equation regardless of the value  $kd$  or distance from the  $z$  axis. In the paraxial limit, this new wave function exhibits all of the pertinent characteristics of the traditional scalar description of the fundamental mode of a Gaussian beam. This comparison will now be explored in depth.

A New Mathematical Model for the Fundamental Mode  
of a Propagating Gaussian Beam

To recapitulate from Chapter 2, the traditional scalar description of the fundamental mode of a Gaussian beam can be written

$$u(x,y,z) = \frac{w_0}{w(z)} \exp \left\{ -i \left[ kz - \Phi + \frac{k(x^2+y^2)}{2R(z)} \right] - \frac{x^2+y^2}{w^2(z)} \right\}, \quad (3.32)$$

where

$$\Phi = \tan^{-1} \left( \frac{z}{z_0} \right), \quad (3.33)$$

$$R(z) = \frac{z^2 + z_0^2}{z}, \quad (3.34)$$

$$w(z) = w_0 \left[ 1 + \left( \frac{z}{z_0} \right)^2 \right]^{1/2}. \quad (3.35)$$

The beam parameters  $z_0$  and  $w_0$  are known as the Rayleigh range and beam waist respectively, and are related by

$$z_0 = \frac{kw_0^2}{2}, \quad (3.36)$$

with  $k$  the propagation constant of the medium. The wave function described by Equation (3.32) is interpreted as having a spherical wavefront, with radius of curvature  $R(z)$ , a Gaussian amplitude distribution that expands according to the hyperbola defined by Equation (3.35), and a phase difference or draw-up  $\Phi$  between the Gaussian beam and a plane wave. The amplitude factor  $w_0/w$  gives the expected intensity decrease on axis attributable to expansion of the beam.

Equation (3.35) defines a hyperbola of constant amplitude  $1/e$  and divergence angle  $\delta$ . We can relate it to a single hyperboloid of revolution in the oblate spheroidal coordinate system, obtaining an



expression for  $z/z_0$  in oblate spheroidal coordinates that will facilitate the comparisons between the traditional Gaussian beam model and the one suggested in Equation (3.28). The value of the beam waist  $w_0$  is the value of  $\rho$  on the surface described by Equation (3.35) in the cross-sectional plane  $z = 0$ . Using the parametric equations in Equation (3.5), we note that when  $z = 0$ ,  $\mu = 0$ , and therefore  $\cosh\mu = 1$ . This leads to the expression for  $w_0^2$ , that is

$$w_0^2 = x^2 + y^2 = d^2 \cosh^2\mu \sin^2\delta = d^2 \sin^2\delta . \quad (3.39)$$

The value of  $w^2(z)$  equals  $\rho^2$  in any  $z \neq 0$  or  $\mu \neq 0$  plane.

Comparing these expressions with Equation (3.35), we can write

$$\frac{w^2(z)}{w_0^2} = \frac{d^2 \cosh^2\mu \sin^2\delta}{d^2 \sin^2\delta} = 1 + \frac{z^2}{z_0^2} . \quad (3.40)$$

Cancelling common terms leads to

$$\cosh^2\mu = 1 + \frac{z^2}{z_0^2} . \quad (3.41)$$

Once again, we employ the trigonometric identity of Equation (3.10a)

and find that

$$\sinh\mu = \frac{z}{z_0} . \quad (3.42)$$

Since  $\xi = \sinh\mu$  in the oblate spheroidal coordinate system, we now have a simple expression for the phase draw up  $\Phi$ . Equation (3.33) now becomes

$$\Phi = \tan^{-1}\xi . \quad (3.43)$$

Close study of equation (3.34) reveals that the spherical wavefronts in the classical Gaussian beam description possess a radius of curvature consistent with the vertex curvature of a family of oblate ellipsoids having an interfocal spacing of  $2z_0$ . Relating these ellipsoids to the oblate spheroidal coordinate system gives the radius of the focal ring as  $d = z_0$ . This value of  $d$  conflicts with that predicted by the amplitude hyperbola of Equation (3.35), where  $d = \sqrt{w_0^2 + z_0^2}$ . This ambiguity arises from the paraxial approximation where  $\theta$  is so small that  $\sin\theta$  and  $\tan\theta$  are indistinguishable. For the purposes of this chapter,  $d$  will be assumed to equal  $z_0$ .

In the limit of very small  $\theta$ , the exponential amplitude  $\exp[-(x^2+y^2)/w^2(z)]$  equals  $\exp[-kd(1-\cos\theta)]$  of Equation (3.28) where  $\eta = \cos\theta$ . Using the parametric equations in Equation (3.5), we see that

$$\sin^2\theta = \frac{x^2 + y^2}{d^2 \cosh^2\mu} . \quad (3.44)$$

Employing the trigonometric identity of Equation (3.10b) and substituting Equation (3.44) into the argument of the exponential amplitude in Equation (3.28) leads to

$$-kd(1 - \cos\theta) = -kd \left\{ 1 - \left[ 1 - \frac{x^2 + y^2}{d^2 \cosh^2 \mu} \right]^{1/2} \right\}. \quad (3.45)$$

Use of the binomial expansion and the paraxial approximation produces

$$\begin{aligned} -kd(1 - \cos\theta) &\cong -kd \left\{ 1 - \left[ 1 - \frac{x^2 + y^2}{2d^2 \cosh^2 \mu} \right] \right\} \\ &= -\frac{k(x^2 + y^2)}{2d \cosh^2 \mu}. \end{aligned} \quad (3.46)$$

Assuming a value of  $z_0$  for  $d$ , we can rewrite Equation (3.36) as

$$\frac{1}{w_0^2} = \frac{k}{2d}. \quad (3.47)$$

Using Equation (3.47) and substituting the expression for  $\cosh^2 \mu$  given by Equation (3.41) into Equation (3.46) yields

$$-kd(1 - \cos\theta) \cong \frac{x^2 + y^2}{w_0^2 [1 + z^2/z_0^2]} = -\frac{(x^2 + y^2)}{w^2(z)}. \quad (3.48)$$

The exponential amplitude for the wave function expressed by Equation (3.28) is seen to be exactly equivalent to that of a traditional Gaussian amplitude in the limit of small  $\theta$ .

Going further, we can compare the amplitude factor  $[\eta^2 + \xi^2]^{-1/2}$  in Equation (3.28) with  $w_0/w(z)$  in Equation (3.32). Again in the paraxial limit of  $\eta = 1$ , we note that

$$[\eta^2 + \xi^2]^{-1/2} \cong [1 + \xi^2]^{-1/2} = \cosh^{-1} \mu . \quad (3.49)$$

Since we established in Equation (3.40) that  $w(z)/w_0 = \cosh \mu$ , the amplitude factors are seen to be exactly equivalent. Therefore, the intensity decrease on axis [ $\eta = 1$ ] is the same for both models.

We can now rewrite the wave function  $\psi_0/S$  in the paraxial limit to obtain

$$\lim_{\eta \rightarrow 1} \psi_0/S = \frac{\exp \left\{ ikd\xi - i \tan^{-1} \xi - \frac{k}{2d} \left( \frac{x^2 + y^2}{1 + \xi^2} \right) \right\}}{kd (1 + \xi^2)^{1/2}} \quad (3.50)$$

The most notable difference between Equation (3.50) and the traditional Gaussian beam description of Equation (3.32) is that the wavefront is now explicitly a section of an oblate ellipsoid rather than a sphere.

### Summary

In this chapter, we have derived an exact solution to the scalar wave equation in oblate spheroidal coordinates. This solution was obtained using a transformation introduced by Flammer [1957] to convert a spherical wave in the prolate spheroidal coordinate system to a wave function with a Gaussian amplitude distribution in the oblate spheroidal coordinate system. On close analysis of this new wave function, we found that it resolved into the traditional description of a propagating Gaussian beam in the paraxial limit with one important exception; that is, the wavefront is a section of an oblate ellipsoid. This new expression possesses significant advantages over previous attempts to describe a Gaussian beam, including its simplicity and the elimination of the paraxial approximation. The wave function given in Equation (3.28) is merely the zero-order term of an entire family of exact solutions to the scalar Helmholtz equation which possess a fundamental Gaussian amplitude distribution along with predictably higher-order amplitude terms. This family will be derived in Chapter 5. Before going on, however, we shall demonstrate in the following chapter a further advantage of this new formulation for a propagating Gaussian beam. We shall develop a simple, elegant and powerful geometrical model for the beam which serves to interpret Equation (3.28) in the light of geometrical optics and allows for straight-line propagation of a Gaussian beam.

## CHAPTER 4

A GEOMETRICAL MODEL FOR THE FUNDAMENTAL MODE OF A  
PROPAGATING GAUSSIAN BEAM

Since Gaussian beams are used in a wide variety of optical systems and instruments, a geometrical model of the beam is an absolute necessity for system design and analysis. Ideally, the model should correspond to the mathematical description and interpret it in the light of geometrical optics. Many methods exist for predicting the first-order properties of a Gaussian beam as it traverses an optical system. Those that are wedded to the mathematical description of Kogelnik and Li suffer from the approximations inherent in that wave function, as well as the awkwardness of describing the beam in the Cartesian coordinate system. The more useful, and recently more popular, method is that introduced by Arnaud [1969, 1973, 1985] which unites the geometrical constructs of the oblate spheroidal coordinate system with the wavefronts and amplitude contours of the Kogelnik and Li model. In particular, Arnaud utilizes the geometrical concept of a ruled surface that is produced by the motion of a skew line. A ruled surface is a surface generated by the motion of a straight line, called a rectilinear generator, in three-dimensional space. A hyperboloid of revolution of one sheet,  $\eta = \text{constant}$  in the oblate

spheroidal coordinate system, is one example of a ruled surface, which in this case, results from the rotation of a straight line about an axis it does not cut. The straight line is therefore skewed to the axis of symmetry, the z axis, and is referred to as a skew line. The skew line lies on the surface of the hyperboloid and is everywhere tangent to it.

The skew line has enjoyed some popularity as the basis for a Gaussian beam model because of the simplicity and well-behaved nature of a straight line. Arnaud treats a complex representation of the skew line as a complex ray which obeys the laws of geometrical optics. Further work has since extended the representation of the fundamental mode of a Gaussian beam by complex rays [Keller and Streifer (1971), Einziger and Felsen (1982), and Herloski, Marshall and Antos (1983)], a concept Felsen [1976] vigorously disputes. In contrast, a real representation of the skew line leads to an elegant design tool for predicting the first-order properties of a Gaussian beam in an optical system. This method was first introduced by Shack [1983], and later developed more fully by Kessler and Shack [1984].

All of these constructs hinge on a geometrical interpretation of the traditional description of the fundamental mode of a Gaussian beam as given by Equation (3.32). In contrast, the geometrical model that will be developed in this chapter uses a real representation of the skew line to interpret the new mathematical description of such a beam as given by Equation (3.28). In the course of building this geometrical model, the properties of a skew line will be developed in

depth, incorporating characteristics noted by Arnaud [1985] and introducing others. In addition, the concept of the skew line as a real ray will be explored. Since this idea does have certain drawbacks, the skew line will ultimately be used to build a nonorthogonal coordinate system which provides an unambiguous framework for studying Gaussian beam propagation.

#### Skew-Line Generator of a Ruled Surface

In this section, we begin by discussing the geometrical characteristics of an individual skew line and its relationship to both the hyperboloid of revolution and the oblate ellipsoid in the oblate spheroidal coordinate system. Later, we shall expand the discussion to include the behavior of families of skew lines. This development constitutes the background for a later section in which the characteristics of a skew line are compared with the ray of geometrical optics.

In Figure 4.1, let NP represent a skew line revolving about the z axis, with ON the common perpendicular to the axis and NP in any position. Then ON has a constant length,  $W_0$ , and as the line NP rotates, the point N describes a circle of radius  $W_0$  in the x-y plane. The angle between the revolving line and the z axis is also a constant,  $\delta$ .

Let  $P(x,y,z)$  be any point on the skew line and let  $NP = \ell$ . Then the parametric equations of the locus of P, in terms of  $\ell$  and  $\phi$  are



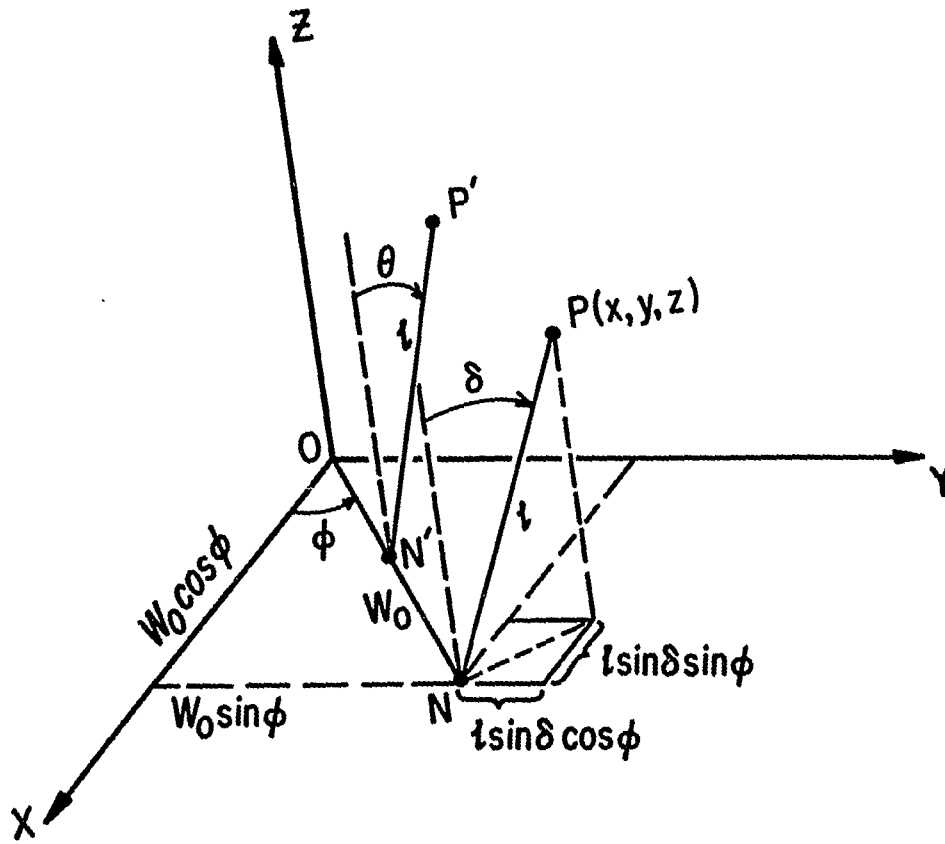


Figure 4.1 The orientation of a skew line  $l$  with respect to the  $z$  axis for two possible deviation angles,  $\delta$  and  $\theta$ .

$$\begin{aligned}
 x &= W_0 \cos \phi - l \sin \delta \sin \phi, \\
 y &= W_0 \sin \phi + l \sin \delta \cos \phi, \\
 z &= l \cos \delta.
 \end{aligned}
 \tag{4.1}$$

The point P lies on a surface expressible as an equation in x, y, and z. This equation will be independent of  $l$  and  $\phi$ , and we can write

$$\begin{aligned}
 x^2 + y^2 &= W_0^2 + l^2 \sin^2 \delta = W_0^2 + \frac{z^2 \sin^2 \delta}{\cos^2 \delta}, \\
 x^2 + y^2 - z^2 \tan^2 \delta &= W_0^2,
 \end{aligned}
 \tag{4.2}$$

$$\frac{x^2 + y^2}{W_0^2} - \frac{z^2}{W_0^2 / \tan^2 \delta} = 1.
 \tag{4.3}$$

This is the equation of a hyperboloid of revolution of one sheet, with a "waist" radius  $W_0$ , and  $\delta$  the angle of the asymptotic cone, or divergence. A cross section of this figure, for  $\phi = \text{constant}$ , is shown in Figure 4.2.

After the manner of Arnaud [1985], we can project the skew line NP onto a second plane parallel to the  $z = 0$  plane and located a distance  $\bar{z}$  away. The projected skew line forms the line segment PQ shown in Figure 4.3. The angle PO'Q has a constant value  $\alpha$  as the skew line rotates about the z axis. As  $\bar{z}$  increases,  $\alpha$  increases proportionately, while the deviation angle  $\delta$  remains constant. If  $\bar{z}$  remains fixed, various values of  $\alpha$  will generate different figures of revolution. Figure 4.4 demonstrates this progression. If  $\alpha = 0$ , the deviation angle is also zero and as the rectilinear generator rotates

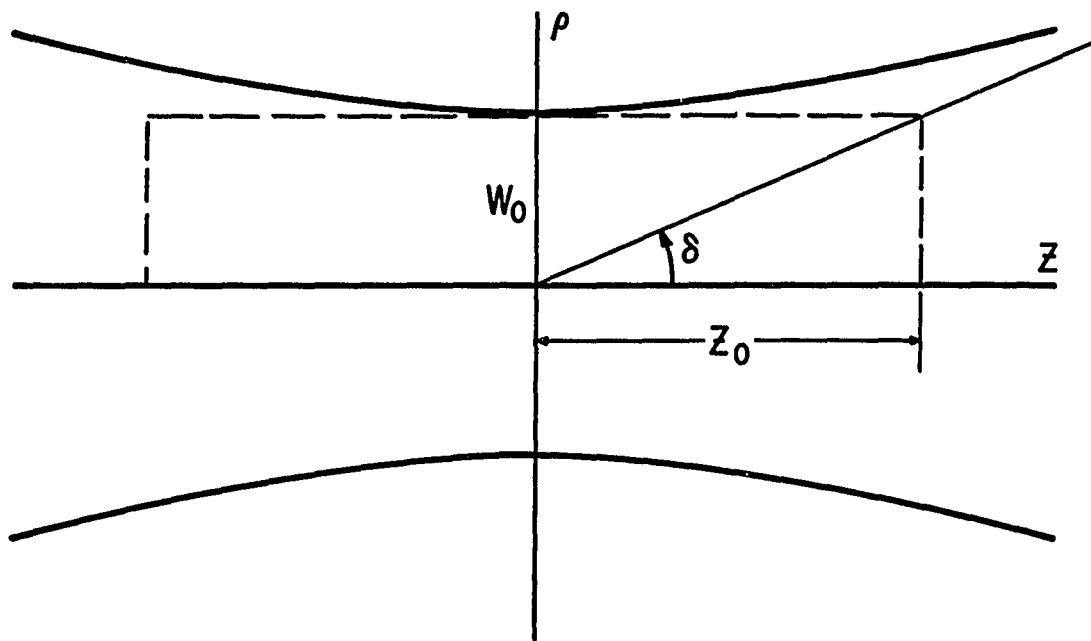


Figure 4.2 Cross section of the hyperboloid generated by the rotation of a skew line about the  $z$  axis.

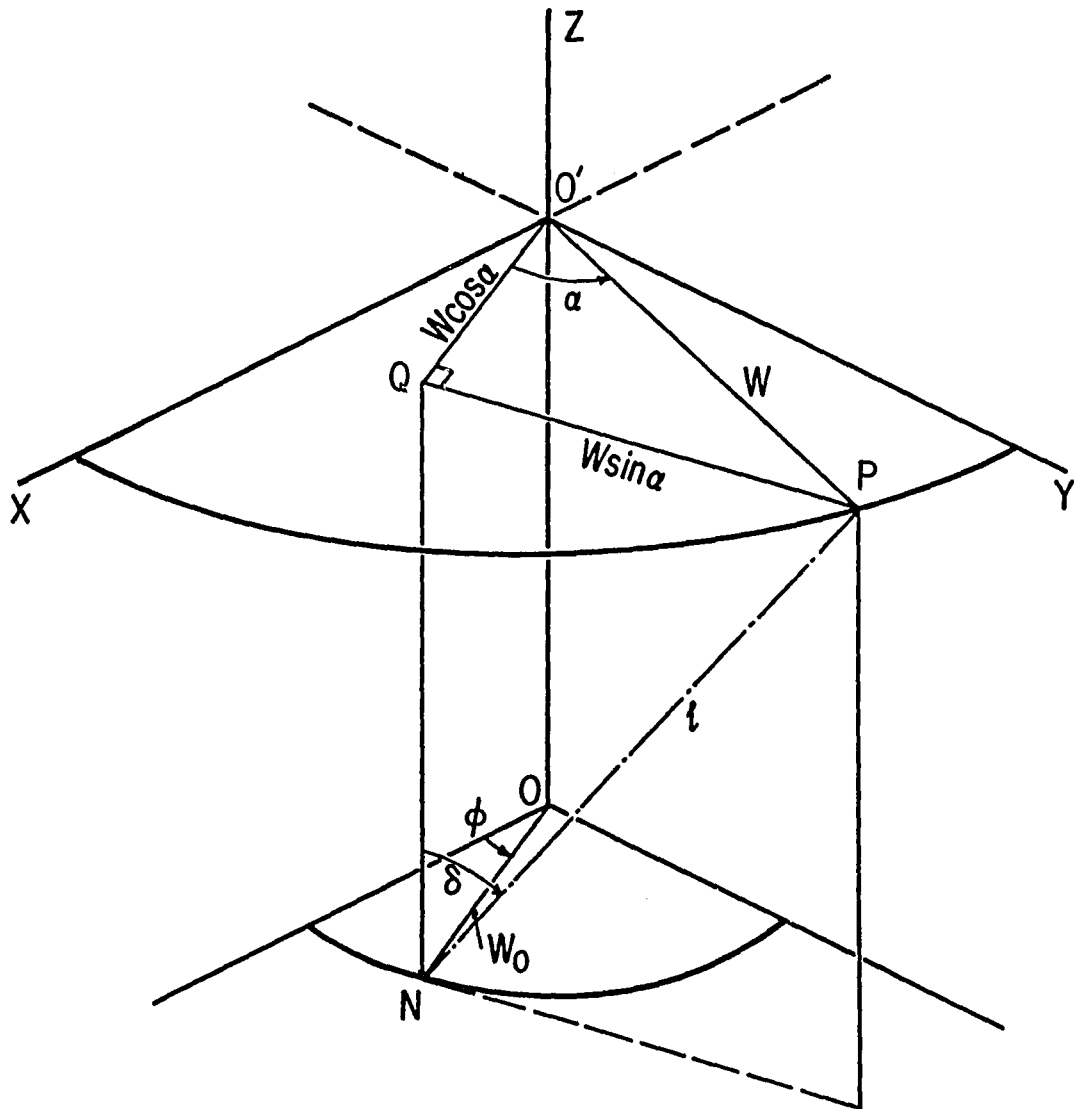


Figure 4.3 The skew line  $l$  projected onto an  $x$ - $y$  plane.

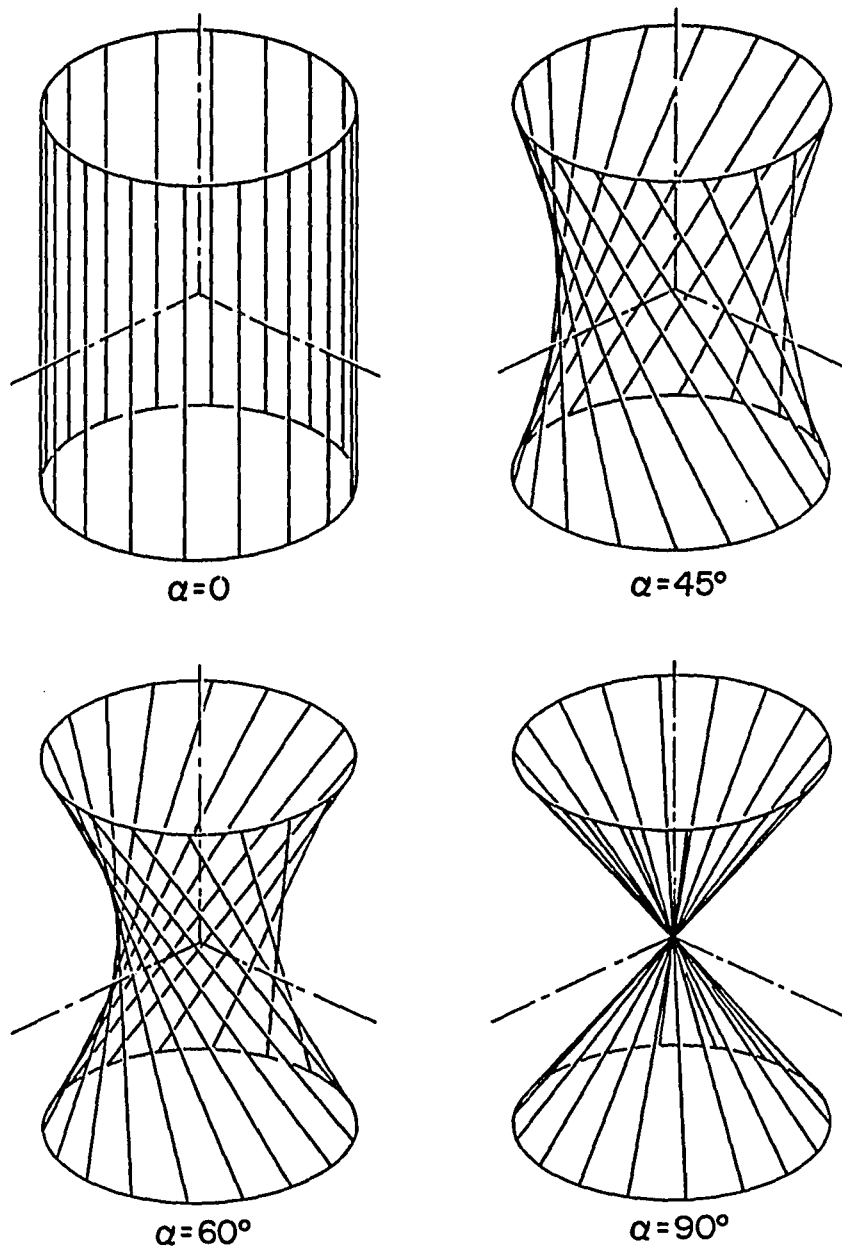


Figure 4.4 Different ruled surfaces for varying amounts of twist  $\alpha$  .

about the  $z$  axis, it sweeps out a cylinder. As  $\alpha$  increases, the figure becomes more "twisted," generating hyperboloids of revolution of increasing deviation angle and decreasing waist sizes. Finally, in the limit as  $\alpha \rightarrow \pi/2$ , the waist radius  $W_0 \rightarrow 0$  and the figure becomes a right circular cone.

The "twist" angle  $\alpha$ , the deviation angle  $\delta$ , and the distance from the plane of the waist  $\bar{z}$  are related to one another. Referring again to Figure 4.3, note that triangle  $PQO'$  is a right triangle with line segment  $PO'$ , given by  $W$ , the radius of the beam cross section in the  $z=\bar{z}$  plane. Line segments  $O'Q$  and  $PQ$  are then given by

$$O'Q = W \cos \alpha \quad , \quad (4.4)$$

$$PQ = W \sin \alpha \quad .$$

Since  $O'Q$  is the projection of the beam waist radius  $W_0$  onto this second plane, the value  $W \cos \alpha$  is always a constant.

Comparing Equation (4.3) with Equation (3.35), we see that

$$\frac{W_0^2}{\tan^2 \delta} = z_0^2 \quad . \quad (4.5)$$

Using Equation (4.4), we can rewrite Equation (4.5) as

$$\tan \delta = \frac{W \cos \alpha}{z_0} \quad . \quad (4.6)$$

The relationship between  $\tan\delta$  and  $\bar{z}$  can be seen in Figure 4.3 and is given by

$$\tan\delta = \frac{W\sin\alpha}{\bar{z}} \quad . \quad (4.7)$$

Now the dependence of the twist angle  $\alpha$  on the distance from the plane of the waist can be expressed explicitly by

$$\tan\alpha = \frac{\bar{z}}{z_0} \quad . \quad (4.8)$$

This is the same result as Equation (3.33) with  $\alpha$  the same angle as  $\Phi$ , the phase difference in the classical Gaussian beam description.

As Arnaud pointed out, there are two possible orientations for a skew line. That is, a hyperboloid of revolution can be generated by a skew line deviated to the right of vertical as shown in Figures 4.1 and 4.3, or by one deviated to the left. This would mean, in the case of the former, an angle  $\alpha$  "twisted" counterclockwise with respect to the positive  $z$  axis; and in the latter case, an angle  $\alpha$  twisted clockwise with respect to the  $z$  axis. Figure 4.5 demonstrates this principle. The counterclockwise twist will be considered a positive  $\alpha$  since it is the same direction as a positive  $\phi$ . Conversely, the clockwise twist will be taken to be a negative  $\alpha$ . Although the skew lines are equivalent in the sense that either will generate a hyperboloid of one sheet when rotated about the  $z$  axis, taken together they generate a set of nonorthogonal parametric lines, or coordinate curves, on the surface of the hyperboloid.

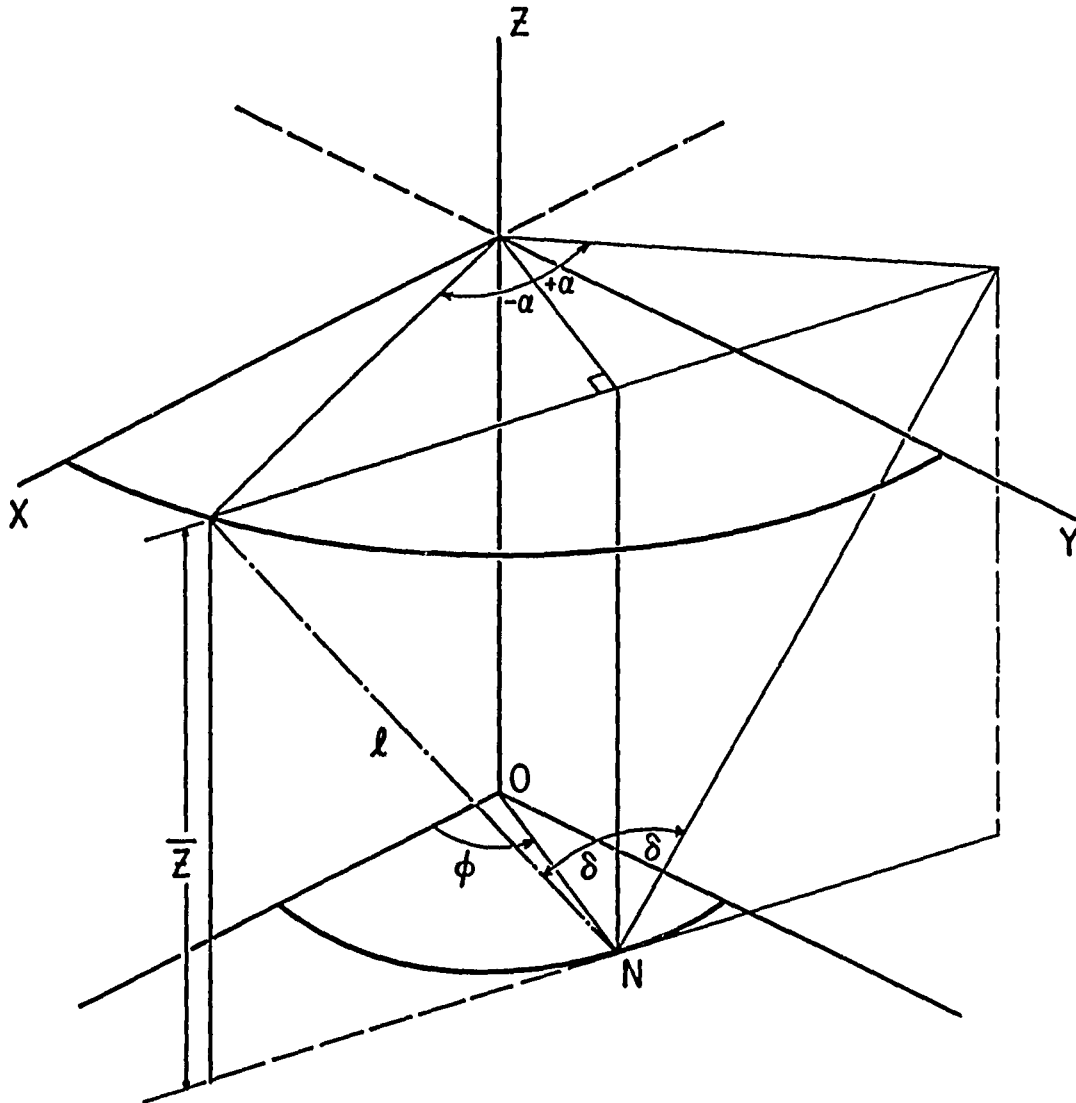


Figure 4.5 The two possible twist orientations for the skew line  $l$ .



The twist angle  $a$  can be used to describe the oblate spheroidal coordinate  $\mu$ , the radius of vertex curvature of the ellipse  $R_e$ , and the conic constant  $K$  of the ellipse. By doing so, we can relate abstract parameters, such as  $\mu$ , to a concept that lends itself to physical interpretations. Furthermore, we can develop a basis of comparison with well-known figures, such as circles and spheres, through the relationship of the twist angle to the vertex curvature and the conic constant. Ultimately, this will provide greater insight into the behavior of the Gaussian beam as it propagates. We shall now deal with the relationship between the twist angle and the oblate spheroidal coordinate  $\mu$ , and develop the remaining relationships later.

The oblate spheroidal coordinate  $\mu$  is a hyperbolic angle which can be equated to a circular angle, in this case  $a$ , through the gudermanian function. The gudermanian expresses the functional relationship between hyperbolic and circular angles without resorting to imaginary values. This function is written as

$$a = \text{gd}(\mu) \quad , \quad (4.9)$$

where

$$\text{gd}(\mu) = \int_0^{\mu} \frac{dt}{\cosh(t)} \quad , \quad (4.10)$$

and the inverse gudermanian is given by

$$\mu = \text{gd}^{-1}(a) = \int_0^a \frac{dt}{\cos(t)} \quad . \quad (4.11)$$

Performing the integration in both Equation (4.10) and Equation (4.11) leads to a pair of inverse relationships

$$a = 2 \tan^{-1}(e^{\mu}) - \pi/2 , \quad (4.12a)$$

and

$$\mu = \ln \tan \left[ \frac{\pi}{4} + \frac{a}{2} \right] . \quad (4.12b)$$

Expanding the tangent function in Equation (4.12b) leads to

$$e^{\mu} = \frac{1 + \sin a}{\cos a} , \quad (4.13a)$$

and

$$e^{-\mu} = \frac{1 - \sin a}{\cos a} . \quad (4.13b)$$

We can now derive a more explicit functional relationship between  $\mu$  and  $a$ . Using the definitions of the oblate spheroidal coordinate  $\xi$  given in Chapter 3 and the exponential definitions of  $\sinh\mu$  and  $\cosh\mu$ , we can write

$$\begin{aligned} \xi &= \sinh\mu = \tan a , \\ (\xi^2 + 1)^{1/2} &= \cosh\mu = 1/\cos a , \end{aligned} \quad (4.14)$$

$$\tanh\mu = \sin a .$$

Using the expressions in Equation (4.14) , we can rewrite the parametric equations for the oblate spheroidal system as

$$\begin{aligned} x &= \frac{d \sin \theta \cos \phi}{\cos \alpha} , \\ y &= \frac{d \sin \theta \sin \phi}{\cos \alpha} , \\ z &= d \tan \alpha \cos \theta . \end{aligned} \tag{4.15}$$

We can conclude from the foregoing discussion that the skew line and its corresponding twist angle can be used in describing the oblate spheroidal coordinate system in which we have described a new mathematical model for a propagating Gaussian beam. Therefore, they should be equally useful in describing the beam itself.

In Equation (4.3), varying  $W_0$  and  $\delta$  will generate a family of hyperboloids related by their common focus. As shown in Figure 4.1, each member of this family possesses its own rectilinear generator  $N'P'$  and deviation angle  $\theta$ , but all members have common perpendiculars to the  $z$  axis,  $ON' = W' \cos \alpha$ . Let the length of the skew line  $N'P' = \ell$  remain constant for all  $\theta$ . Then the parametric equations of the locus of  $P'(x', y', z')$  are similar to Equation (4.1) and we can write

$$\begin{aligned} x' &= W' \cos \phi = W' \cos \alpha \cos \phi - \ell \sin \theta \sin \phi , \\ y' &= W' \sin \phi = W' \cos \alpha \sin \phi + \ell \sin \theta \cos \phi , \\ z' &= \ell \cos \theta . \end{aligned} \tag{4.16}$$

We wish to find the surface containing  $P'$ , which is the endpoint of all skew lines having the same length  $l$  and twist  $a$  from the plane of the waist. The surface will be independent of  $\theta$  and  $\phi$ , and we can write

$$x'^2 + y'^2 = W'^2 \cos^2 a + l^2 \sin^2 \theta \quad ,$$

$$W'^2 = W^2 \cos^2 a + l^2 (1 - z'^2/l^2) \quad . \quad (4.17)$$

After some algebraic manipulation, Equation (4.17) becomes

$$\frac{W'^2}{l^2/\sin^2 a} + \frac{z'^2}{l^2} = 1 \quad . \quad (4.18)$$

This is the equation of an ellipse at a constant "distance"  $l$ , as measured along a skew line, from the plane of the waist of the hyperboloid of revolution. A family of skew lines with constant length  $l$  can be seen in Figure 4.6. The arc PA lies on the ellipsoid specified by the constant angle  $a$ . More will be said about this family later. Since the parameter  $l/\sin a > l$  for  $a < \pi/2$ , this particular ellipsoid has foci and semi-major axis located in the plane  $z = 0$  and is a figure of revolution about the  $z$  axis. As such, it is referred to as an oblate ellipsoid.

Equations (4.3) and (4.18) represent specific examples of more general families of hyperboloids of revolution of one sheet and oblate ellipsoids that are orthogonal. The general equation for a hyperboloid of revolution is given by

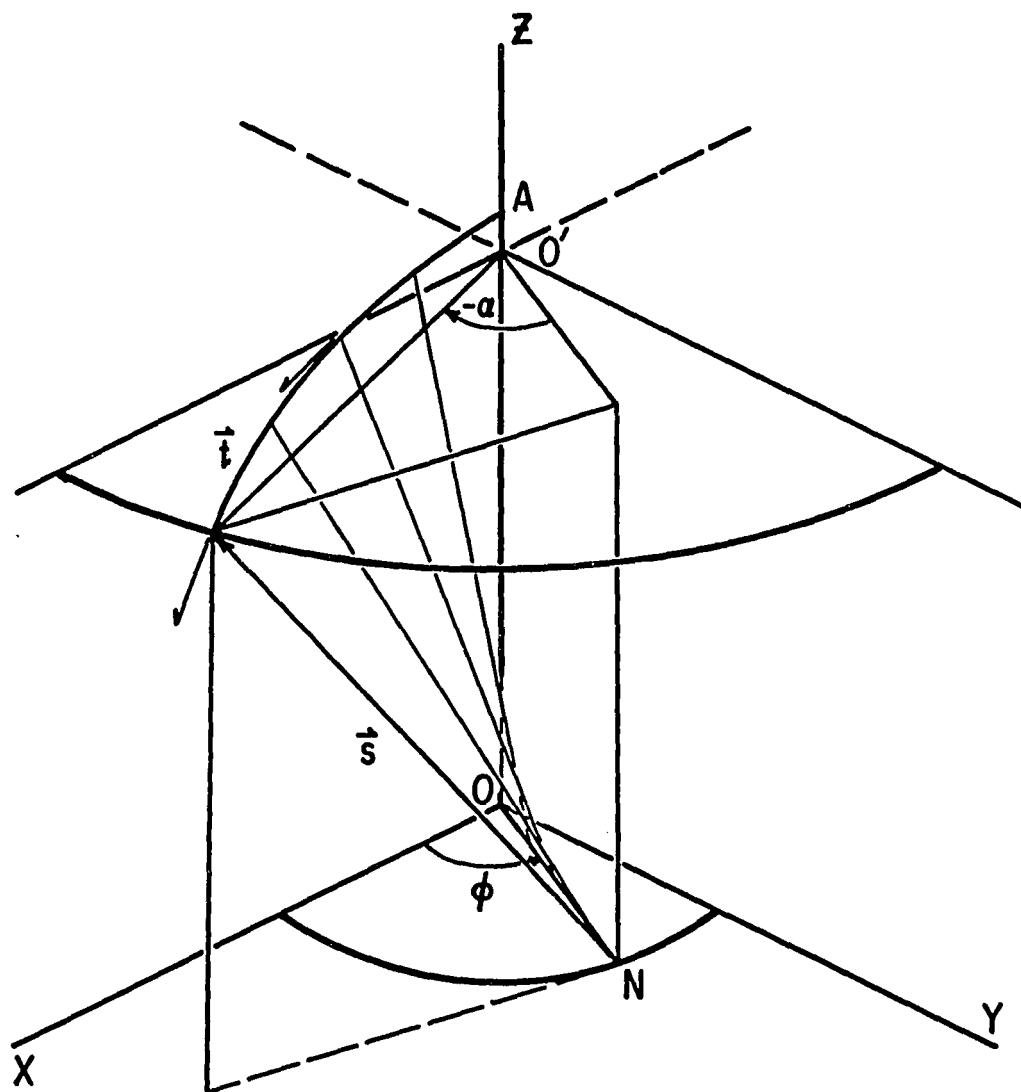


Figure 4.6 A clockwise fan of skew lines ending in elliptical arc PA.

$$\frac{\rho^2}{a_h^2} + \frac{z^2}{a_h^2 - d^2} = 1, \quad a_h < d \quad (4.19)$$

while for an ellipse it is

$$\frac{\rho^2}{a_e^2} + \frac{z^2}{a_e^2 - d^2} = 1, \quad a_e < d \quad (4.20)$$

where  $\rho^2 = x^2 + y^2$ . For a figure of revolution about  $z$ , the focus describes a ring of diameter  $2d$ . The parameters  $a_e$  and  $a_h$ , shown in Figure 4.7, refer to the semi-major axis of the ellipse and the waist radius of the hyperboloid, respectively.

Comparing Equation (4.18) with Equation (4.20), we note that

$$a_e^2 - d^2 = l^2 = \frac{l^2}{\sin^2 a} - d^2 \quad (4.21)$$

This expression can be reduced to

$$l = d \tan a \quad (4.22)$$

and Equation (4.20) becomes

$$\frac{\rho^2}{d^2 / \cos^2 a} + \frac{z^2}{d^2 \tan^2 a} = 1 \quad (4.23)$$

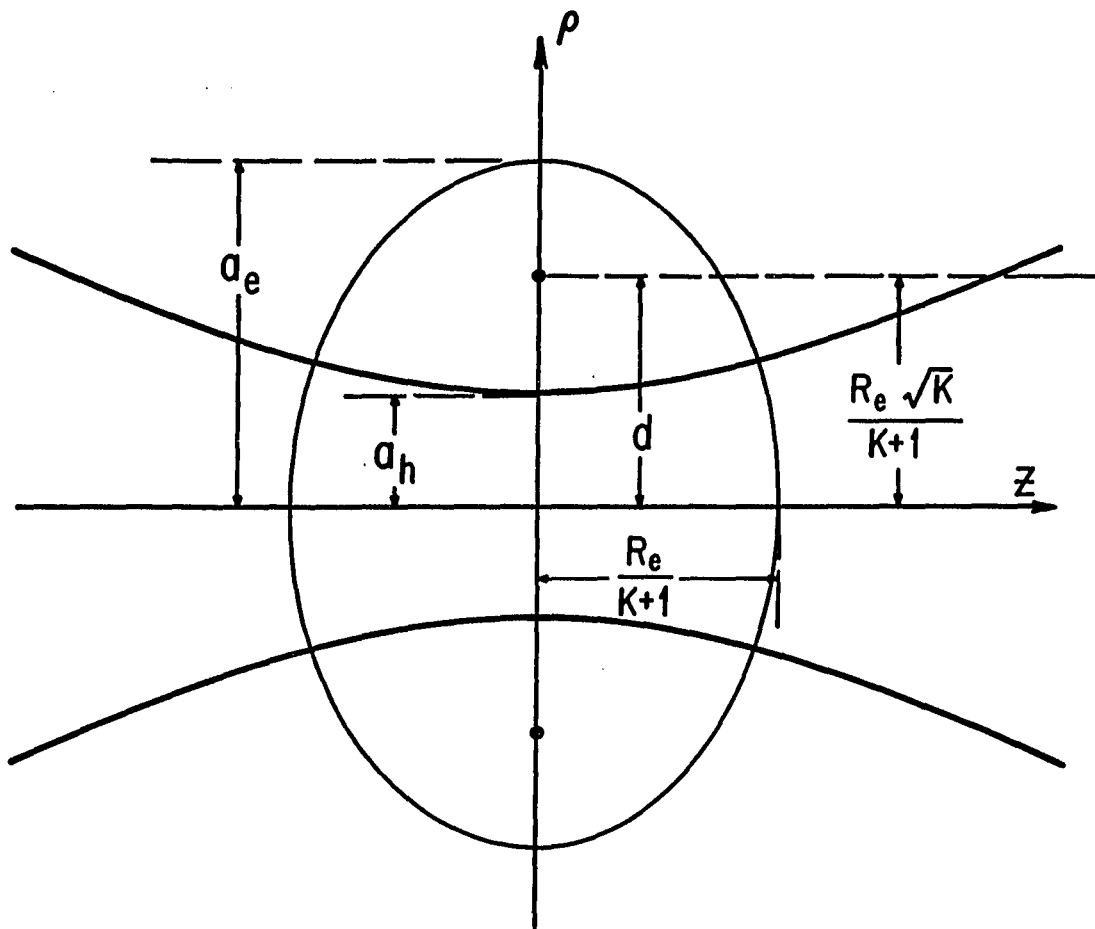


Figure 4.7 Cross section of an oblate ellipse and hyperbola with common foci at  $\pm d$ .

Note that the twist angle  $a$  determines not only the length of the skew line  $l$ , but also the specific ellipsoid in a family of ellipsoids with the same focus. Therefore, this angle can be used to ascertain both the conic constant and the radius of vertex curvature associated with each ellipsoidal surface. We will now establish the relationship between these two parameters and the twist angle.

Figure 4.7 demonstrates the connection between the focus distance  $d$ , the conic constant  $K$ , and the radius of vertex curvature  $R_e$ . The conic constant occurs in the expression for the sag of the ellipse and provides a basis of comparison with a sphere, for which  $K = 0$ . These relationships can be expressed by

$$d^2 = \frac{(R_e^2)K}{(K + 1)^2} \quad , \quad (4.24)$$

and

$$a_e^2 - d^2 = \frac{R_e^2}{(K + 1)^2} \quad . \quad (4.25)$$

Comparing Equations (4.20) and (4.23), we see that the lengths of the skew line and the semi-minor axis are equal. Making this substitution in Equation (4.25) yields

$$d^2 \tan^2 a = \frac{R_e^2}{(K + 1)^2} \quad . \quad (4.26)$$



Using the expression for  $d^2$  given by Equation (4.24), we can rewrite Equation (4.26) as,

$$\tan^2 a = \frac{R_e^2}{(K+1)^2} \cdot \frac{(K+1)^2}{R_e^2 K} = \frac{1}{K} . \quad (4.27)$$

Equation (4.27) leads to expressions for both the sine and the cosine of the twist angle in terms of the conic constant. These expressions are

$$\cos^2 a = \frac{K}{K+1} , \quad (4.28)$$

$$\sin^2 a = \frac{1}{K+1} .$$

A general expression for the hyperboloid of revolution can be obtained by comparing Equations (4.19) and (4.3). First, note that

$$a_h^2 - d^2 = - \frac{W_o^2}{\tan^2 \delta} = - \frac{a_h^2}{\tan^2 \delta} , \quad (4.29)$$

where

$$a_h = d \sin \delta . \quad (4.30)$$

Since  $\delta$  is a specific deviation angle and  $\theta$  represents all possible such angles between 0 and  $\pm\pi/2$ , replacing  $\delta$  by  $\theta$  produces a family of hyperboloids expressed by

$$\frac{\rho^2}{d^2 \sin^2 \theta} - \frac{z^2}{d^2 \cos^2 \theta} = 1 . \quad (4.31)$$

In this formation, the "Rayleigh range" for any specific hyperboloid is given by

$$z_0 = d \cos \theta \quad , \quad (4.32)$$

and the expression for the waist radius has the form

$$W_0 = d \sin \theta \quad . \quad (4.33)$$

So far, we have discussed the characteristics of a skew line as it pertains to both a hyperboloid of one sheet and an oblate ellipsoid. We began by showing how a straight line tilted at an angle  $\delta$  from the  $z$  axis and skewed to it forms a single hyperboloid upon rotation about the  $z$  axis. Next, we found that the endpoints of a family of skew lines all having the same length and twist formed a single oblate ellipsoid. We then expanded the equation for the hyperboloid to include all possible deviation angles  $\theta$  between 0 and  $\pi/2$  and derived a general equation for a family of hyperboloids with the same focus spacing. We performed the equivalent procedure for a family of ellipsoids. Along the way, we were able to relate the twist angle  $\alpha$  of the skew line to the skew-line length, the conic constant of the ellipse, the gudermanian of the hyperbolic angle  $\mu$  in the oblate spheroidal coordinate system, and the phase draw up  $\Phi$  in the traditional description of the fundamental mode of a Gaussian beam. We can also see from Equations (4.22), (4.8), and (3.33) that this phase draw up is directly proportional to the skew-line length, or

$$\tan\Phi = \tan\alpha = \frac{\ell}{d} . \quad (4.34)$$

We may expand on the relationship between an ellipsoid, specified by  $a$ , and the rectilinear generators of a family hyperboloids. Each member of this family of skew lines has the same length  $\ell$ , and a deviation angle  $\theta$  appropriate to its hyperboloid, and it intersects the plane of the waist along a line segment at an angle  $\phi$  from the  $x$  axis. We shall refer to this family, or aggregate, of skew lines as a fan, as depicted in Figure 4.6. The parametric equations for this family are given by Equation (4.15), with  $a$  and  $\phi$  constant. Each skew line, represented here as the vector  $\vec{s}$ , has the same length between the elliptical arc PA and the line segment ON. The intersection of each successive skew line with segment ON is a distance  $d\sin\theta$  from the origin where  $\theta$  is the deviation angle of the new skew line and represents another member in the family of hyperboloids. Furthermore, each skew line is perpendicular to the elliptical arc. This is intuitively true since each skew line lies on a hyperboloid, the orthogonal surface to an oblate ellipsoid. A more detailed proof of this orthogonal relationship is given in Appendix B.

#### A Geometrical Interpretation of $\exp[-i\tan^{-1}\xi/\eta]$

Both the classical mathematical model of a Gaussian beam given in Equation (3.32) and the new model of Equation (3.28) include arctangent factors in their exponential phase terms. The geometrical

interpretation of these terms, if any exists, is not intuitively obvious, but such an interpretation would provide much useful insight into the nature of a phase front as it propagates. In the case of the term  $\exp[-i \tan^{-1} \xi / \eta]$ , a simple geometrical explanation does exist.

Figure 4.8 displays the skew line PN once again at a deviation angle  $\theta$ , along with the attendant elliptical arc PA. Further, the figure designates two planes, one perpendicular to the z axis at the point O', and the other perpendicular to the z axis at point A. The O' plane intersects the hyperbola of deviation angle  $\theta$  in a circle with radius  $\rho = PO'$ . The A plane is the tangent plane to this oblate ellipsoid at the point (0,0,A). The distance along the z axis between these two planes is the sag of ellipse. As demonstrated earlier in this chapter, the distance AO must be the length of the skew line,  $d \tan a$ . Therefore, the equation for the sag is given by

$$\text{sag} = (1 - \cos \theta) d \tan a \quad . \quad (4.35)$$

The distance from plane O' to plane A along the skew line, designated  $\Delta l$  in the drawing, is simply  $\text{sag} / \cos \theta$  or

$$\Delta l = \frac{d \tan a}{\cos \theta} (1 - \cos \theta) \quad . \quad (4.36)$$

The increased length of the skew line translates directly into an increase in the twist angle  $a$ . We shall refer to this increased twist as  $\Delta a$ . The entire length of the skew line from the point N until it

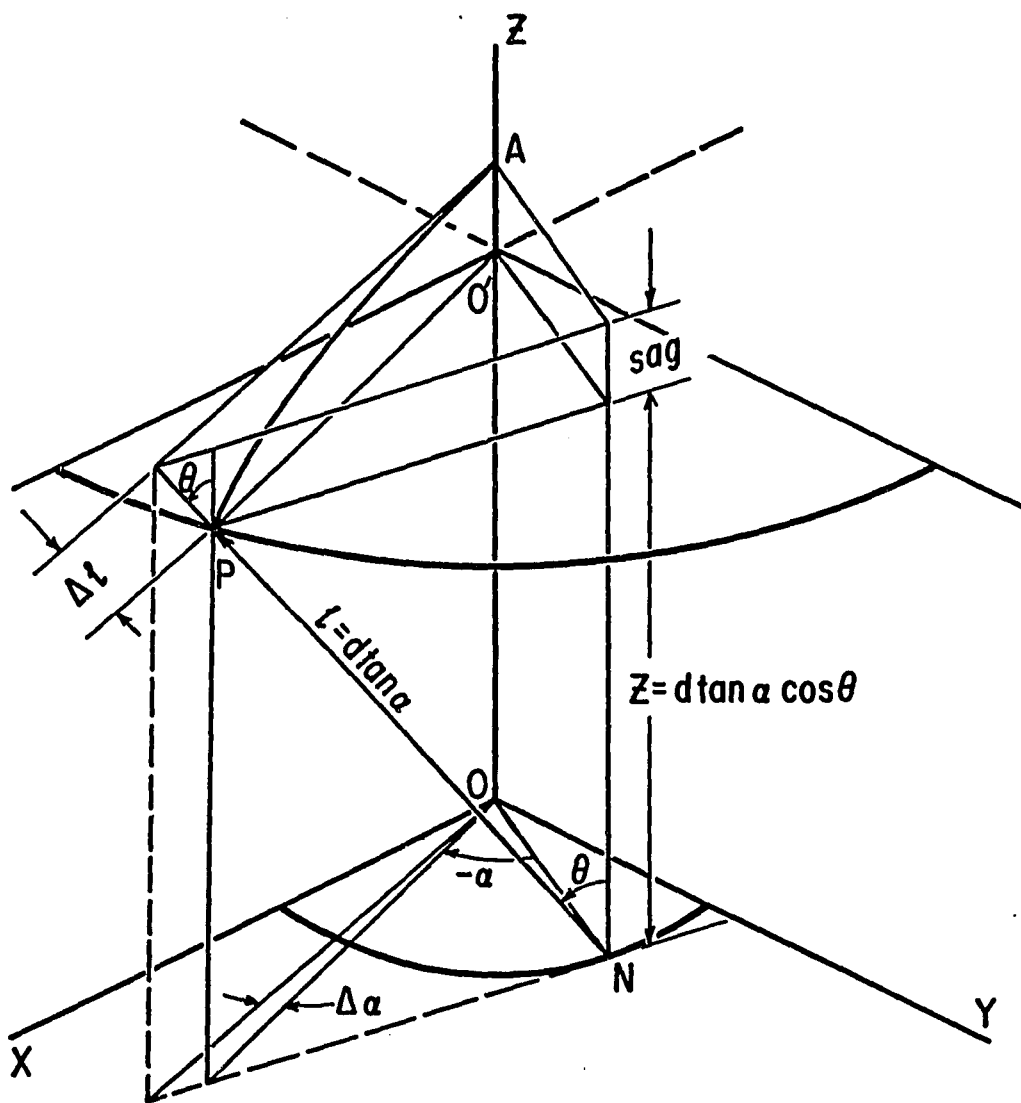


Figure 4.8 Differential change in twist angle,  $\Delta \alpha$  .

intersects the plane A is given by  $l + \Delta l$ . Given in terms of the twist angle, this equality can be written

$$l + \Delta l = d \tan(a + \Delta a) \quad . \quad (4.37)$$

Substituting the value of  $d \tan a$  for  $l$  and Equation (4.36) for yields

$$d \tan a + \frac{d \tan a}{\cos \theta} - d \tan a = d \tan(a + \Delta a) \quad . \quad (4.38)$$

Equation (4.38) reduces to

$$\frac{\tan a}{\cos \theta} = \tan(a + \Delta a). \quad (4.39)$$

From the definition of the oblate spheroidal coordinates, we can substitute  $\xi$  and  $\eta$  for the terms  $\tan a$  and  $\cos \theta$  respectively, and Equation (4.39) becomes

$$\frac{\xi}{\eta} = \tan(a + \Delta a) \quad . \quad (4.40)$$

Finally, by taking the arctangent of both sides, we obtain

$$\tan^{-1} \xi/\eta = a + \Delta a \quad . \quad (4.41)$$

Some algebraic manipulation is required to obtain an expression for  $\Delta\alpha$  in terms of  $\xi$  and  $\eta$ . This equation is

$$\Delta\alpha = \tan^{-1} \left[ \frac{\xi(1-\eta)}{\eta + \xi^2} \right]. \quad (4.42)$$

The entire exponential phase factor  $\exp[-i \tan^{-1} \xi / \eta]$  can then be rewritten as

$$\exp[-i \tan^{-1} \xi / \eta] = \exp \left[ -i \tan^{-1} \xi - i \tan^{-1} \left( \frac{\xi(1-\eta)}{\eta + \xi^2} \right) \right] \quad (4.43)$$

The difference between the exponential arctangent phase term of the classical Gaussian beam mathematical model and the one represented by Equation (3.28) is the phase term attributable to  $\Delta\alpha$ .

The exponential factor  $\exp[-i \tan^{-1} \xi / \eta]$  can be interpreted as representing the difference in phase that a wave disturbance would undergo in traveling along the skew line in the time required for the edge of the beam to advance to the  $z$  distance represented by plane A. In other words, as the beam propagates, its center will pass plane A at time  $t_1$ . At a later time  $t_2$ , the beam edge will pass plane A and in the time  $t_2 - t_1$ , the phase of the beam will have changed by  $\Delta\alpha$ . Another interpretation is that of an off-axis or wavefront error. That is, a wavefront is ideally a section of an oblate ellipsoid, but off-axis it is deviated from the ideal by an amount given by  $\Delta\alpha/k$ . For example,  $\alpha = 45^\circ$  at the Rayleigh range of any Gaussian beam. In the case of a highly divergent beam with a  $\theta$  of  $10^\circ$  at the  $1/e$  field

point, the wavefront will have a  $\lambda/8$  deviation from a perfectly oblate ellipsoid at the  $1/e$  radius at the Rayleigh range. We can expand on this calculation and compare the classical wavefront with the new model. In this case, there is a  $\lambda/6$  difference at the  $1/e$  radius of the Rayleigh range. Finally, a surface of constant phase in the new model will be curved toward the plane of the waist and inside the figure of the oblate ellipsoid. Although it is tempting to believe this makes the wavefront more spherical, the opposite is true. An oblate ellipsoid is always curved inside the figure of a sphere whose radius corresponds to the vertex curvature of the ellipse.

#### The Skew Line as a Ray

We have pointed out a number of interesting features of individual skew lines as well as fans of skew lines. In particular, we noted that the relationship between an elliptical arc and a fan of skew lines demonstrates three characteristics of wavefronts and their trajectories. First, the spatial separation as measured along any skew line of the fan between an elliptical arc and the plane of the waist remains constant. This statement can be generalized to include the separation between any two elliptical arcs of different twist angles and perpendicular to the same family of skew lines. Second, successive elliptical arcs are perpendicular to a skew line prompting comparisons between successive arcs on a wavefront and the wavefront's orthogonal trajectory or "ray". Finally, the skew line is a straight line, as is the current geometrical-optical model ray in a homogeneous medium.



Still further, we can prove that when a ray with a skew-line trajectory is reflected from an oblate-ellipsoidal mirror, the reflected ray also possesses a skew-line trajectory on the same hyperboloid as the incident ray. Proof of this requires a three-dimensional vector analysis and can be found in Appendix C. This particular characteristic of the skew line finds a useful application in modeling ray behavior in an optical resonator. The continuum of all such skew rays forms the hyperbolic amplitude contours. Such an envelope of rays has been discussed by Bykov and Vainshtein [1965], Kahn [1965], and Steier [1966].

All of the above demonstrates that the skew-line model of a propagating Gaussian beam as expressed in Equation (3.28) possesses strong parallels with the theorems of geometrical optics. However, two traits of a skew line prevent it from being defined as a ray. The first is that the skew line is not the gradient of the wavefront, an oblate ellipsoid with a Gaussian amplitude distribution. Second, if a skew line is to be seen as a ray, the possibility of its two different orientations must be considered equally likely. One way to include both positive and negative (counterclockwise and clockwise) twists in the geometrical description of a Gaussian beam is to treat the two skew lines as unit coordinate vectors,  $\hat{u}$  and  $\hat{v}$ , which together with the unit vector  $\hat{\theta}$ , forms a nonorthogonal coordinate system.

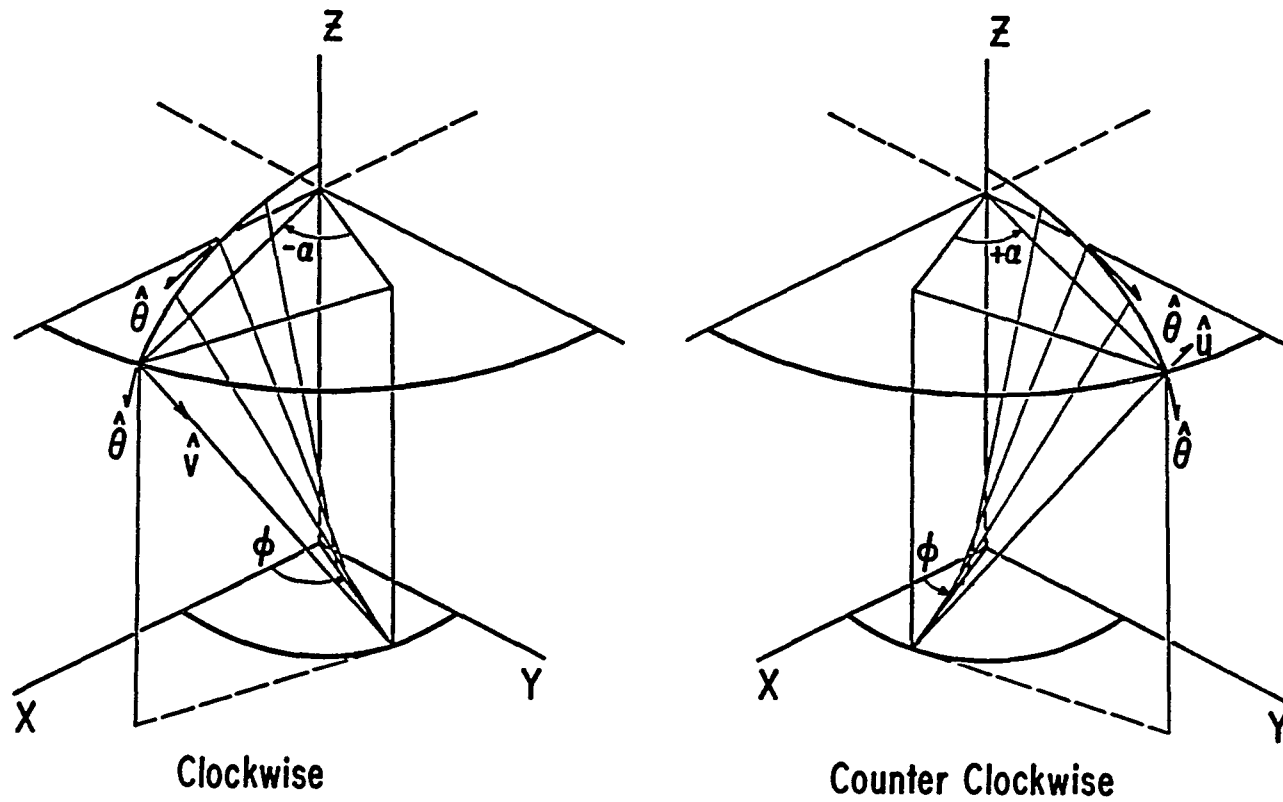


Figure 4.9 Clockwise and counterclockwise skew line fans.

The skew line fan of Figure 4.6 can be repeated for a counterclockwise twist and the resultant figure forms a mirror image of the clockwise fan. These mirror-image fans are demonstrated in Figure 4.9. The skew line  $\vec{s}$  of Figure 4.6 reappears as  $\hat{v}$  pointed in the opposite direction and the tangent vector to the ellipse,  $\vec{t}$  has been renamed  $\hat{\theta}$ . The skew line for a positive  $a$ , or counterclockwise twist is  $\hat{u}$ . Figure 4.9 sketches the skew line fans for a few discrete deviation angles. In actual practice, the fans are surfaces called right conoids, as shown in Figure 4.10. All of these conoids are ruled surfaces; they are formed by the motion of a straight line, the skew line, in three-dimensional space. Together with the oblate ellipsoid, these surfaces form three coordinate surfaces in a non-orthogonal coordinate system, which we shall now explore. An excellent discussion of general three-dimensional curvilinear coordinate systems can be found in Stratton's Electromagnetic Theory [1941].

#### Nonorthogonal Coordinate System

The coordinate system whose unit base vectors are shown in Figure 4.11 consists of two right conoids, one left-handed and the other right-handed, whose axis of symmetry is the  $z$  axis, and an oblate ellipsoid with focal ring of radius  $d$  in the  $x$ - $y$  plane. The right-handed conoid is the counterclockwise skew-line fan ( $+a$ ), so named because the skew lines seem to sweep in the direction of the

fingers of the right hand with the thumb pointed in the direction of the positive  $z$  axis as  $a$  varies from  $-\pi/2$  to  $+\pi/2$ . Similarly, the left-handed conoid is the clockwise skew-line fan ( $-a$ ). This time, as  $a$  varies from  $+\pi/2$  to  $-\pi/2$ , the skew lines sweep in the direction of the fingers of the left hand when the thumb is pointed in the direction of the positive  $z$  axis. The coordinate angle  $\theta$  is the deviation angle of the skew line with respect to the  $z$  axis. This coordinate system is not orthogonal since the coordinate surfaces do not intersect at right angles.

The parametric equations for this system are

$$\begin{aligned} x(u,v,\theta) &= \frac{d \sin \theta \cos \left( \frac{u+v}{2} \right)}{\cos \left( \frac{u-v}{2} \right)} , \\ y(u,v,\theta) &= \frac{d \sin \theta \sin \left( \frac{u+v}{2} \right)}{\cos \left( \frac{u-v}{2} \right)} , \\ z(u,v,\theta) &= d \cos \theta \tan \left( \frac{u-v}{2} \right) , \end{aligned} \tag{4.44}$$

where

$$u = \phi + a , \tag{4.45}$$

**Figure 4.10 Left-handed (clockwise) and right-handed (counterclockwise) right conoids.**

These are the continuous surfaces generated by the skew-line fans of Figure 4.9. The figure on the left is a left-handed right conoid corresponding to a clockwise twist, and the figure on the right is a right-handed right conoid corresponding to a counterclockwise twist. In both drawings, the lower circle represents the plane of the waist. The vertical dotted line represents the z axis in both.

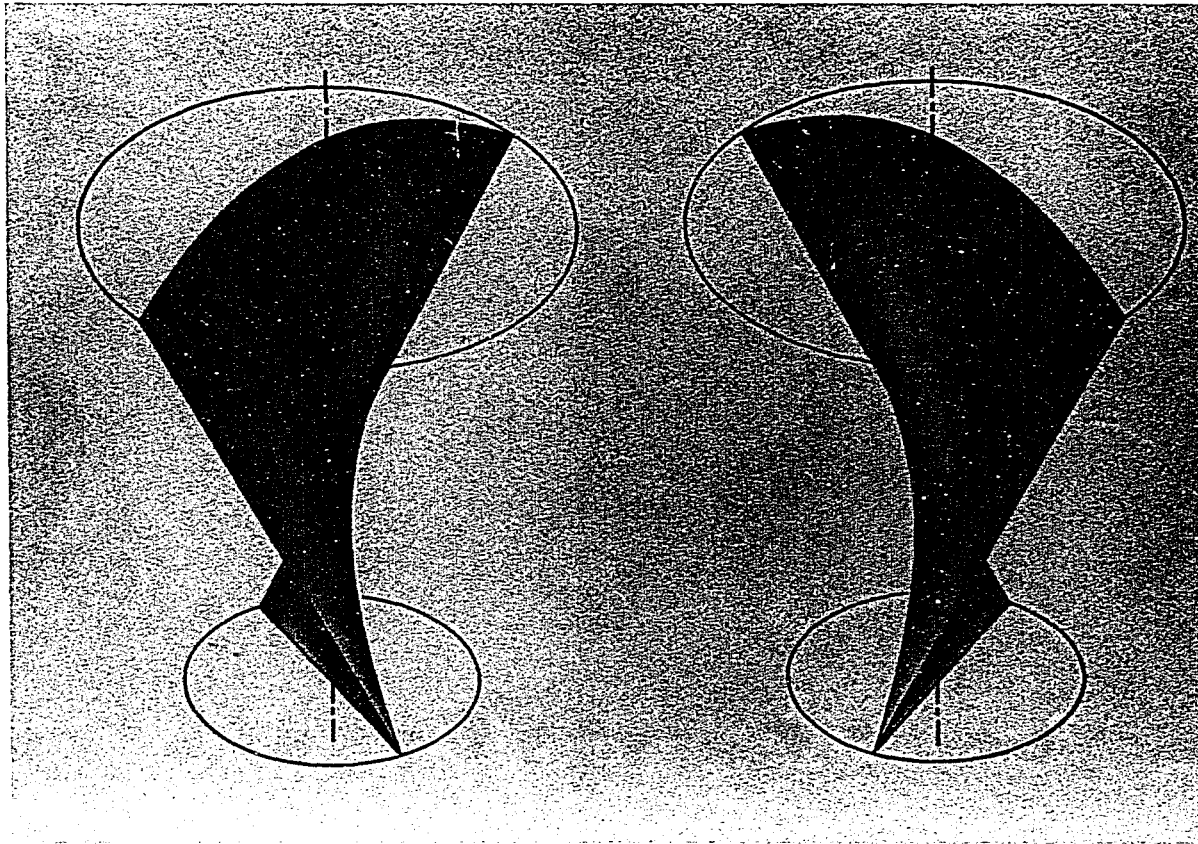


Figure 4.10 Left-handed (clockwise) and right-handed (counterclockwise) right conoids

Figure 4.11 The unit base vectors  $\hat{u}$ ,  $\hat{v}$ , and  $\hat{\theta}$  for a nonorthogonal coordinate system.

The two conoids join in an elliptical arc in the left-hand figure. Their junction in the plane of the waist is shown on the right.

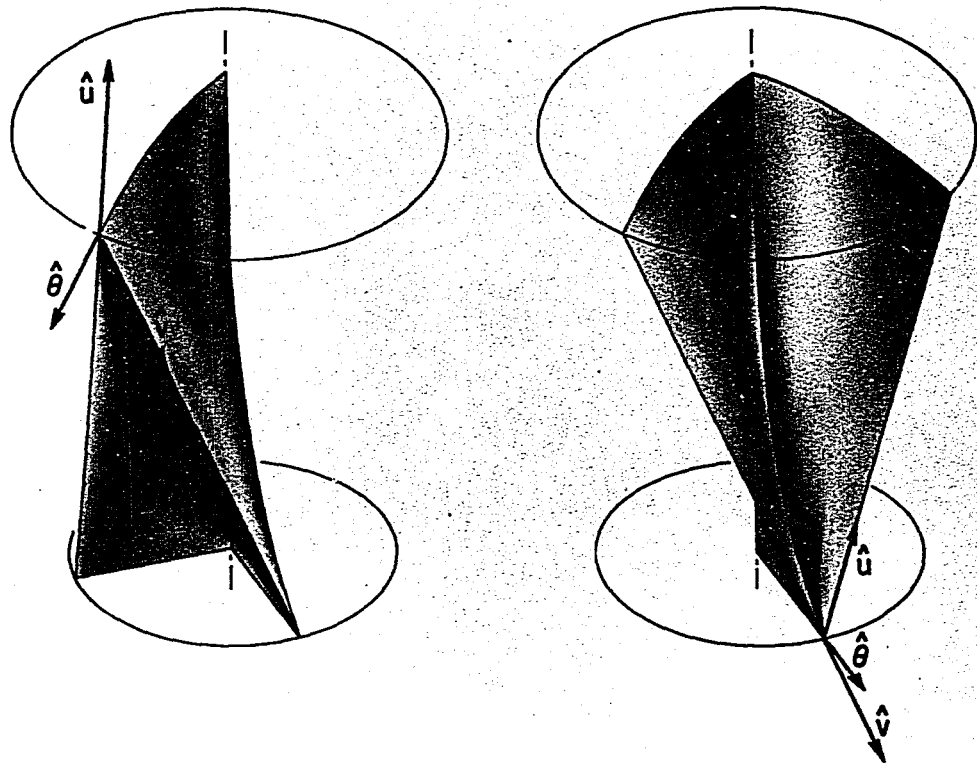


Figure 4.11 Unit base vectors  $\hat{u}$ ,  $\hat{v}$ , and  $\hat{\theta}$  for a nonorthogonal coordinate system



and

$$v = \phi - a . \quad (4.46)$$

Note that these equations strongly resemble the parametric equations of the oblate spheroidal coordinate system. In the latter system, we discussed surfaces on which one of the coordinate variables was constant and the unit base vectors,  $\hat{\xi}$ ,  $\hat{\eta}$ , and  $\hat{\phi}$ , were the normals to these surfaces. For the sake of simplicity, we shall not discuss the surfaces on which the coordinates  $u$ ,  $v$ , and  $\theta$  are constant. Rather, we shall outline the unit base vectors,  $\hat{u}$ ,  $\hat{v}$ , and  $\hat{\theta}$ , as the vectors tangent to the right-handed conoid, the left-handed conoid, and the oblate ellipsoid, respectively.

The position vector  $\vec{r}$  to any point in the three-dimensional space is given by

$$\vec{r} = x(u,v,\theta)\hat{x} + y(u,v,\theta)\hat{y} + z(u,v,\theta)\hat{z}, \quad (4.47)$$

where  $\hat{x}$ ,  $\hat{y}$ , and  $\hat{z}$  are the unit base vectors in the Cartesian coordinate system. A differential change in  $\vec{r}$  due to small displacements along the coordinate curves is expressed by

$$d\vec{r} = \frac{\partial \vec{r}}{\partial u} du + \frac{\partial \vec{r}}{\partial v} dv + \frac{\partial \vec{r}}{\partial \theta} d\theta . \quad (4.48)$$

If one moves a unit distance along any one of the coordinate curves, the change in  $\vec{r}$  is directed tangentially along that curve and has a magnitude of one. The vectors

$$\begin{aligned}\hat{u} &= \frac{\frac{\partial \vec{r}}{\partial u}}{\left| \frac{\partial \vec{r}}{\partial u} \right|} , \\ \hat{v} &= \frac{\frac{\partial \vec{r}}{\partial v}}{\left| \frac{\partial \vec{r}}{\partial v} \right|} , \\ \hat{\theta} &= \frac{\frac{\partial \vec{r}}{\partial \theta}}{\left| \frac{\partial \vec{r}}{\partial \theta} \right|} ,\end{aligned}\tag{4.49}$$

are the unit base vectors for the coordinate system. For the nonorthogonal system here, these vectors can be written in terms of the Cartesian base vectors as

$$\begin{aligned}\hat{u} &= -\sin\theta \sin v \hat{x} + \sin\theta \cos v \hat{y} + \cos\theta \hat{z} , \\ \hat{v} &= -\sin\theta \sin u \hat{x} + \sin\theta \cos u \hat{y} + \cos\theta \hat{z} , \\ \hat{\theta} &= \frac{\cos\theta \cos\phi}{(1-a^2)^{1/2}} \hat{x} + \frac{\cos\theta \sin\phi}{(1-a^2)^{1/2}} \hat{y} - \frac{\sin\theta \sin((u-v)/2)}{(1-a^2)^{1/2}} \hat{z} ,\end{aligned}\tag{4.50}$$

where  $a = \sin\theta \cos((u-v)/2)$ . Figure 4.11 shows these vectors at two different points in space.

In the left-hand drawing of Figure 4.11, the observation point resides on an oblate ellipsoid some distance from the plane of the waist. The vectors  $\hat{u}$ ,  $\hat{v}$ , and  $\hat{\theta}$  are shown as tangent to the conoids and the oblate ellipsoid respectively. The conoidal surfaces meet in an arc on this ellipsoid and are separated in the plane of the waist by the acute angle  $2a$ , where  $a$  is the twist angle associated with the ellipsoid. Likewise, the drawing on the right shows the observation point in the plane of the waist with the associated unit base vectors. The conoids meet along a line segment in the waist and are separated on an oblate ellipsoid by the acute angle  $2a$ .

The most useful aspect of these unit vectors is their relationship to the unit vectors of the oblate spheroidal system. Starting with Equation (4.50), some algebraic manipulation leads to the relations

$$\begin{aligned}\hat{\xi} &= \frac{\hat{u} - \hat{v}}{2(1 - a^2)^{1/2}}, \\ \hat{\phi} &= \frac{\hat{u} + \hat{v}}{2a}, \\ \hat{\eta} &= -\hat{\theta}.\end{aligned}\tag{4.51}$$

Clearly, the connection between the two systems is a simple one with the useful result that two straight-line vectors,  $\hat{u}$  and  $\hat{v}$ , can synthesize the tangential vectors,  $\hat{\xi}$  and  $\hat{\phi}$ , to two curves in space, a hyperbola and a circle.

Although using a nonorthogonal coordinate system seems an unnecessary complication of the geometrical model, this particular coordinate system has some advantages over the oblate spheroidal coordinate system in describing a propagating Gaussian beam. Specifically, the normal to the oblate ellipsoid can be obtained from the difference between the two skew-line vectors,  $\hat{u}$  and  $\hat{v}$ . This gives us the advantage of using straight-line trajectories to predict the beam's behavior upon propagation, without neglecting the proper description of a wavefront gradient. Also, the nonorthogonal system is so closely allied with the oblate spheroidal coordinate system that straight-line propagation can be utilized without deviating from an exact mathematical description of the beam.

#### Summary

We have developed a simple but precise geometrical interpretation for the alternative mathematical description of the zero-order mode of a propagating Gaussian beam represented by Equation (3.28). This geometrical model is expressed in the oblate spheroidal coordinate system, and it is one of the coordinate surfaces, the hyperboloid of one sheet, of this system which forms the basis of the geometrical configuration. Specifically, the hyperboloid is an example

of a ruled surface which can be generated by a straight line, skew to the  $z$  axis, that rotates about the  $z$  axis.

The properties of an individual skew line as well as families of skew lines have been discussed in detail. We began by describing the hyperboloid generated by the skew line, and then related the skew line to the oblate ellipsoid with the twist angle concept. This relationship exists in the definition of the conic constant of the ellipse, the gudermanian of the hyperbolic angle  $\mu$  in the oblate spheroidal coordinate system, and in the length of the semi-minor axis of the ellipse. Furthermore, the length of the skew line from the plane of the waist to an ellipsoidal surface is directly proportional to the tangent of the twist angle.

One intriguing aspect of both the new and the traditional model of the Gaussian beam is the presence of a pure phase term which has an arctangent dependence. In the traditional model, the phase term is  $\exp[-ia]$ , where  $a$  is the twist angle described earlier. The interpretation of this term is that of a phase draw up, or the difference between the actual wavefront and a plane wave. In Equation (3.28) however, the term appears as  $\exp[-i \tan^{-1} \xi / \eta]$ , which represents the sag of the oblate ellipse as measured along the skew line.

The properties of an individual skew line can be expanded to include families of these lines, and it is in consideration of one type of family in particular, the fan, that strong parallels with the

ray of geometrical optics can be found. A fan of skew lines intersects any oblate ellipsoid in an arc and each member of the fan is perpendicular to this arc. Further, the spatial separation remains constant between two successive arcs, as measured along any member of the fan, much like the constant separation between any two successive wavefronts as measured along their orthogonal trajectories or rays. Proceeding with the idea of a skew line as a ray, we found in Appendix C that a ray with skew-line trajectory reflected from a mirror that is a section of an oblate ellipsoid yields its opposite member; that is, a ray with the skew-line trajectory of the same hyperboloid and equal, but opposite, twist angle.

Finally, we pointed out the failure of the skew line as a ray. Specifically, the skew line is not the gradient of an ellipsoidal wavefront with a Gaussian amplitude distribution and, since there are two possible orientations of each skew line, they must be considered equally likely. It was then suggested that the real power of the skew-line model might lie in using both orientations as two components in a nonorthogonal system, and the details of this system were discussed. Rather than being an unnecessary complication of the issue, using both of these skew-line vectors makes straight-line propagation possible, while providing the unambiguous framework necessary for locating points in space, identifying the wavefront gradient, and defining deformed surfaces such as aberrated wavefronts.

## CHAPTER 5

A COMPLETE FAMILY OF SOLUTIONS TO THE  
SCALAR WAVE EQUATION

In Chapter 3 we noted that Helmholtz's equation is a partial differential equation separable in eleven coordinate systems including the polar and oblate and prolate spheroidal coordinate systems. Separability means that a complete solution to the differential equation consists of a product of three functions, each of which is a solution to an ordinary differential equation. In polar coordinates, a solution to Helmholtz's equation can be written as

$$\psi(r, \theta, \phi) = R(r) \Theta(\theta) \Phi(\phi) \quad (5.1)$$

where  $R(r)$ ,  $\Theta(\theta)$ , and  $\Phi(\phi)$  are solutions to the separated radial, polar angular, and azimuthal equations respectively.

When the Helmholtz equation is separated in polar coordinates, the radial equation has the form

$$r^2 \frac{\partial^2 R(r)}{\partial r^2} + 2r \frac{\partial R(r)}{\partial r} + [k^2 r^2 - n(n+1)] R(r) = 0 \quad (5.2)$$

Solutions to this equation are the spherical Bessel functions,  $j_n(r)$ , spherical Neumann functions  $y_n(r)$ , and spherical Hankel functions of the first and second kind,  $h_n^{(1)}(r)$  and  $h_n^{(2)}(r)$ . The separated polar angular equation is the associated Legendre equation

$$\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} [\Theta(\theta) \sin\theta] + \left[ n(n+1) - \frac{m^2}{\sin^2\theta} \right] \Theta(\theta) = 0, \quad (5.3)$$

with solutions

$$\Theta(\theta) = P_n^m(\cos\theta) \quad . \quad (5.4)$$

Finally, the separated azimuthal equation is

$$\frac{1}{\Phi(\phi)} \frac{\partial^2 \Phi(\phi)}{\partial \phi^2} = -m^2, \quad (5.5)$$

with solutions

$$\Phi(\phi) = \exp[\pm im\phi] \quad . \quad (5.6)$$

The complete solutions to the Helmholtz equation is then the product of these separated solutions. Now, Equation (5.1) can be written as

$$\begin{aligned} \psi(r, \theta, \phi) &= R(r)\Theta(\theta)\Phi(\phi) \\ &= h_n^{(1)}(r) P_n^m(\cos\theta) \exp[\pm im\phi], \end{aligned} \quad (5.7)$$

using only one of the spherical Hankel functions for simplicity.

The same general method can be applied to the situation here, that is, to find the entire family of higher-order solutions that has as its zero-order term Equation (3.28). As we did in Chapter 3, we begin in the prolate spheroidal coordinate system where we described a spherical wave emanating from a point source at the focus  $z = -d$ . Equation (3.25) can be rewritten



$$\psi_{P/S} = e^{-ikd} h_n^{(1)} [kd(\xi + \eta)] . \quad (5.8)$$

where  $n = 0$ . In contrast, a spherical wave converging toward the same point source would be described as

$$\psi_{P/S} = e^{-ikd} h_n^{(2)} [kd(\xi + \eta)] . \quad (5.9)$$

using the spherical Hankel function of the second kind. The entire family of higher-order, diverging-wave solutions would then be

$$\psi_{mn}(\xi, \eta, \phi)_{P/S} = e^{-ikd} h_n^{(1)} [kd(\xi + \eta)] P_n^m[\cos\theta'] e^{\pm im\phi}, \quad (5.10)$$

where the angle  $\theta'$  is the angle  $r_+$  makes with the  $z$  axis. The value of  $\cos\theta'$  is given by

$$\cos\theta' = \frac{z + d}{r_+} = \frac{1 + \xi\eta}{\xi + \eta} . \quad (5.11)$$

The geometry of this situation is demonstrated in Figure 5.1.

Now we perform the transformation of Equation (3.22) on the Helmholtz operator  $\nabla^2 + k^2$  as well as to the individual arguments of both the spherical Hankel function and the associated Legendre function. The argument of the spherical Hankel function becomes

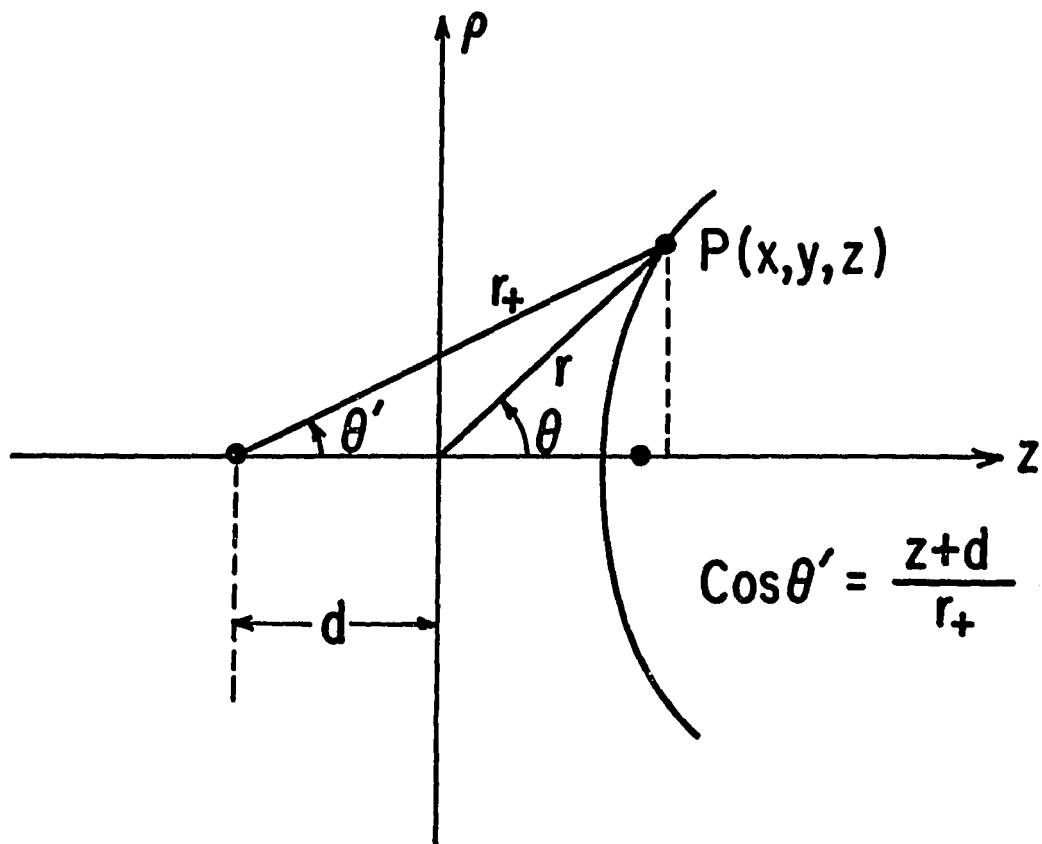


Figure 5.1 Geometry for a polar coordinate system whose origin has been shifted to one of the foci of the prolate spheroidal coordinate system.

$$r = kd_{P/S}(\xi_{P/S} + \eta) \rightarrow t = -ikd_{O/S}(\eta + i\xi_{O/S}) \quad (5.12)$$

$$t = kd_{O/S}[\xi_{O/S} - i\eta] .$$

The argument of the associated Legendre function is written

$$\cos\theta' = \frac{1 + \xi_{P/S}\eta}{\xi_{P/S} + \eta} \rightarrow s = \frac{1 + i\xi_{O/S}\eta}{\eta + i\xi_{O/S}}. \quad (5.13)$$

The complete family in oblate spheroidal coordinates is given by

$$\psi_{mn}(\xi, \eta, \varphi)_{O/S} = e^{-kd} h_n^{(1)}(t) P_n^m(s) e^{im\varphi}, \quad (5.14)$$

where

$$P_n^m(s) = \frac{1}{2^n n!} (1-s^2)^{m/2} \frac{d^{m+n}}{ds^{m+n}} (s^2-1)^n, \quad (5.15)$$

and

$$e^{-kd} h_n^{(1)}(t) = e^{-kd} (-i)^n \frac{e^{it}}{it} \sum_{j=0}^n \left[ -\frac{1}{it} \right]^j \frac{(n+j)!}{2^j j! (n-j)!}. \quad (5.16)$$

We can rewrite Equations (5.15) and (5.16) in terms of the oblate spheroidal coordinates. First, note that

$$s^2 - 1 = \frac{(1 - \eta^2)(1 + \xi^2)}{(\eta + i\xi)^2} \quad (5.17)$$

Now, Equation 5.15 becomes

$$P_n^m(s) = \frac{i^m}{2^n n!} \frac{(1-\eta^2)^{m/2} (\xi^2+1)^{m/2}}{(\eta^2 + \xi^2)^{m/2}} e^{-im \tan^{-1} \frac{\xi}{\eta}} \frac{d^{m+n}}{ds^{m+n}} \left[ \frac{(1-\eta^2)(\xi^2+1)}{(\eta+i\xi)^2} \right]^n, \quad (5.18)$$

and Equation (5.16) is

$$h_n^{(1)}(t) = \frac{(-j)^n e^{-kd(1-\eta)} e^{ikd\xi} e^{-i \tan^{-1} \xi/\eta}}{kd(\eta^2 + \xi^2)^{1/2}} \cdot \sum_{j=0}^n \frac{e^{-ij \tan^{-1} \xi/\eta}}{[-kd(\eta^2 + \xi^2)^{1/2}]^j} \frac{(n+j)!}{2^j j! (n-j)!} \quad (5.19)$$

Obviously, these functions are nonseparable in  $\eta$  and  $\xi$ , although they are separable in  $t$  and  $s$ . To prove that Equation (5.14) is an exact solution to the scalar Helmholtz equation, we must obtain the appropriate separated differential equations in  $t$  and  $s$ . This derivation will now be shown.

Solution of the Scalar Helmholtz  
Equation in Oblate Spheroidal Coordinates

Helmholtz's equation in oblate spheroidal coordinates is given by

$$(\nabla^2 + k^2)\psi_{0/s} = \frac{\partial}{\partial \xi} (\xi^2 + 1) \frac{\partial}{\partial \xi} \psi_{0/s} + \frac{\partial}{\partial \eta} (1 - \eta^2) \frac{\partial}{\partial \eta} \psi_{0/s} \quad (5.20)$$

$$+ \frac{\xi^2 + \eta^2}{(1 + \xi^2)(1 - \eta^2)} \frac{\partial^2 \psi_{0/s}}{\partial \phi^2} + k^2 d^2 (\xi^2 + \eta^2) \psi_{0/s} = 0.$$

We wish to convert this equation to a partial differential equation in  $t$  and  $s$ . To do this, we begin by finding the partial derivatives of  $t$  and  $s$  with respect to  $\xi$  and  $\eta$ . These partial derivatives are

$$\frac{\partial t}{\partial \xi} = kd \qquad \frac{\partial t}{\partial \eta} = -ikd \quad (5.21)$$

and

$$\frac{\partial s}{\partial \xi} = \frac{-i(1 - \eta^2)}{(\eta + i\xi)^2} \qquad \frac{\partial s}{\partial \eta} = -\frac{(\xi^2 + 1)}{(\eta + i\xi)^2} \quad (5.22)$$

Next, we convert the partial derivatives of  $\psi_{0/s}$  with respect to  $\xi$  and  $\eta$  to derivatives with respect to  $t$  and  $s$ . These derivatives can be written as

$$\begin{aligned} \frac{\partial \psi_{0/s}}{\partial \xi} = e^{im\phi} e^{-kd} \left\{ \frac{d}{dt} \left( h_n^{(1)}(t) \right) \left( p_n^m(s) \right) \frac{\partial t}{\partial \xi} + \right. \\ \left. + h_n^{(1)}(t) \frac{d}{ds} \left( p_n^m(s) \right) \frac{\partial s}{\partial \xi} \right\}, \end{aligned} \quad (5.23)$$

$$\begin{aligned} \frac{\partial \psi_{0/s}}{\partial \eta} = e^{im\phi} e^{-kd} \left\{ \frac{d}{dt} \left( h_n^{(1)}(t) \right) \left( p_n^m(s) \right) \frac{\partial t}{\partial \eta} + \right. \\ \left. + h_n^{(1)}(t) \frac{d}{ds} \left( p_n^m(s) \right) \frac{\partial s}{\partial \eta} \right\}. \end{aligned}$$

With further differentiation and some algebraic manipulation, the individual right-hand terms of Equation (5.20) become

$$\begin{aligned} \frac{\partial}{\partial \xi} (\xi^2 + 1) \frac{\partial \psi_{0/s}}{\partial \xi} = 2\xi \left\{ \frac{h'}{h} kd - \frac{P'}{P} \frac{i(1-\eta^2)}{(\eta+i\xi)^2} \right\} \psi_{0/s} \\ + (1+\xi^2) \left\{ \left( \frac{h'}{h} \right)^2 k^2 d^2 - \left( \frac{P'}{P} \right)^2 \frac{(1-\eta^2)^2}{(\eta+i\xi)^4} \right. \\ \left. - 2ikd \frac{h'}{h} \frac{P'}{P} \frac{(1-\eta^2)}{(\eta+i\xi)^2} \right. \\ \left. + k^2 d^2 \left[ \frac{h''}{h} - \left( \frac{h'}{h} \right)^2 \right] \right. \\ \left. - \frac{(1-\eta^2)^2}{(\eta-i\xi)^4} \left[ \frac{P''}{P} - \left( \frac{P'}{P} \right)^2 \right] \right\} \psi_{0/s}, \end{aligned} \quad (5.24)$$

$$\begin{aligned}
\frac{\partial}{\partial \eta} (1-\eta^2) \frac{\partial \psi_{0/S}}{\partial \eta} = & -2\eta \left\{ -ikd \frac{h'}{h} - \frac{P'}{P} \frac{\xi^2+1}{(\eta+i\xi)^2} \right\} \psi_{0/S} \quad (5.25) \\
& + (1-\eta^2) \left\{ -k^2 d^2 \left( \frac{h'}{h} \right)^2 + \left( \frac{P'}{P} \right)^2 \frac{(\xi^2+1)^2}{(\eta+i\xi)^4} \right. \\
& + 2ikd \frac{h'}{h} \cdot \frac{P'}{P} \cdot \frac{\xi^2+1}{(\eta+i\xi)^2} \\
& - k^2 d^2 \left[ \frac{h''}{h} - \left( \frac{h'}{h} \right)^2 \right] \\
& \left. + \frac{(\xi^2+1)^2}{(\eta+i\xi)^4} \left[ \frac{P''}{P} - \left( \frac{P'}{P} \right)^2 \right] \right\} \psi_{0/S} \quad ,
\end{aligned}$$

and

$$\frac{\xi^2 + \eta^2}{(\xi^2 + 1)(1 - \eta^2)} \frac{\partial^2 \psi_{0/S}}{\partial \vartheta^2} = - \frac{m^2 (\xi^2 + \eta^2)}{(1 + \xi^2)(1 - \eta^2)} \psi_{0/S} \quad (5.26)$$

In these equations,  $h_n^{(1)}(t)$  and  $P_n^m(s)$  have been abbreviated to  $h$  and  $P$  respectively. The single prime indicates differentiation with respect to the argument; a double prime indicates a second derivative with respect to the argument.

After some further algebraic manipulation and cancelling of common terms, Equation (5.20) reduces to

$$\begin{aligned}
& k^2 d^2 (\eta + i\xi)^2 \left[ 1 + \frac{h''}{h} \right] \psi_{0/S} + 2ikd (\eta + i\xi) \frac{h'}{h} \psi_{0/S} \\
& + \frac{P''}{P} \frac{(1 - \eta^2)(\xi^2 + 1)}{(\eta + i\xi)^2} \psi_{0/S} + 2 \frac{P'}{P} \frac{1 + i\xi\eta}{\eta + i\xi} \psi_{0/S} \quad (5.27) \\
& - \frac{m^2 (\eta + i\xi)^2}{(1 - \eta^2)(\xi^2 + 1)} \psi_{0/S} = 0 .
\end{aligned}$$

Making the appropriate substitutions for  $t$  and  $s$  leads to

$$\begin{aligned}
& \left\{ -t^2 \left[ 1 + \frac{h''}{h} \right] - 2t \frac{h'}{h} \right\} \psi_{0/S} + \quad (5.28) \\
& \left\{ -(1 - s^2) \frac{P''}{P} + 2s \frac{P'}{P} + \frac{m^2}{1 - s^2} \right\} \psi_{0/S} = 0 .
\end{aligned}$$

By setting both of these bracketed terms equal to an integer constant times  $\psi_{0/S}$ , we can separate the differential equations with respect to  $t$  and  $s$ . The choice of an integral separation constant and its specific value is predicated on obtaining a polynomial for  $P_n(s)$  as shown by Bell (1968). For an integer constant of  $-n(n+1)$ , the first term in Equation (5.28) becomes

$$t^2 h'' + 2th' + [t^2 - n(n+1)]h = 0 . \quad (5.29)$$

With an integer constant of  $n(n+1)$ , the second term in Equation (5.28) is



$$(1 - s^2) P'' - 2sP' + \left[ n(n+1) - \frac{m^2}{1 - s^2} \right] P = 0 . \quad (5.30)$$

Equation (5.29) is the spherical Bessel equation whose solutions are the spherical Bessel functions mentioned earlier; Equation (5.30) is the associated Legendre equation with solution  $P_n^m(s)$ . The wave function  $\psi_{mn}(\xi, \eta, \phi)_0/s$  as given by Equation (5.14) is indeed an exact solution to the scalar Helmholtz equation.

The amplitude distribution of each higher-order member of the family of solutions given by Equation (5.14) contains a fundamental Gaussian amplitude with higher-order amplitude terms. The Gaussian-like exponential amplitude term,  $\exp[-kd(1 - \eta)]$ , occurs for every member of the family, but as the order number increases, increasing orders of  $(\eta^2 + \xi^2)^{1/2}$  occur along with the appropriate coefficients. These terms will affect the exact amplitude distribution and the amplitude of each family member will vary with propagation, that is, with increasing  $\xi$ . To see this more clearly, we shall now explicitly expand the functions  $h_n^{(1)}(t)$  and  $P_n^m(s)$  for some individual orders.

Legendre Polynomials and  
Associated Legendre Functions

The associated Legendre functions are given by

$$P_n^m(s) = \frac{1}{2^n n!} (1 - s^2)^{m/2} \frac{d^{m+n}}{ds^{m+n}} (s^2 - 1)^n, \quad (5.31)$$

with  $P_n^m(s) = 0$  for  $m > n$ . With  $m = 0$ ,

$$P_n^0(s) = P_n(s) = \frac{1}{2^n n!} \frac{d^n}{ds^n} (s^2 - 1)^n, \quad (5.32)$$

which are the Legendre polynomials.

In Tables 5-1 and 5-2,

$$s = \frac{1 + i\xi\eta}{\eta + i\xi} = \frac{1 + (\xi\eta)^2}{\eta^2 + \xi^2} e^{i \tan^{-1} \xi/\eta} e^{-i \tan^{-1} \xi/\eta}, \quad (5.33)$$

and

$$\begin{aligned} s^2 - 1 &= \frac{(1 - \eta^2)(\xi^2 + 1)}{(\eta + i\xi)^2} \\ &= \frac{(1 - \eta^2)(\xi^2 + 1)}{\eta^2 + \xi^2} e^{-i2 \tan^{-1} \xi/\eta}. \end{aligned} \quad (5.34)$$

Table 5.1. Legendre polynomials

---


$$P_0(s) = 1$$

$$P_1(s) = \frac{1 + i\xi\eta}{\eta + i\xi} = \frac{\sqrt{1 + (\xi\eta)^2}}{\sqrt{\eta^2 + \xi^2}} e^{i \tan^{-1} \xi\eta} e^{-i \tan^{-1} \xi/\eta}$$

$$P_2(s) = \frac{1}{2}[3s^2 - 1]$$

$$= \frac{3}{2} \frac{(1 - \eta^2)(\xi^2 + 1)}{\eta^2 + \xi^2} e^{-i2 \tan^{-1} \xi/\eta} + 2$$

$$P_3(s) = \frac{1}{2}[5s^3 - 3s]$$

$$= \frac{(1 + i\xi\eta)}{\eta + i\xi} \left\{ 5 \frac{(1 - \eta^2)(\xi^2 + 1)}{(\eta + i\xi)^2} + 2 \right\}$$

$$P_4(s) = \frac{1}{8}[35s^4 - 30s^2 + 3]$$

$$= \frac{35}{8} \left[ \frac{(1 - \eta^2)^2(\xi^2 + 1)^2}{(\eta + i\xi)^4} \right] + 5 \frac{(1 - \eta^2)(\xi^2 + 1)}{(\eta + i\xi)^2} + 1$$


---

Table 5.2. Associated Legendre functions.

$$P_1^1(s) = \sqrt{1-s^2} = i \frac{\sqrt{1-\eta^2} \sqrt{\xi^2+1}}{\sqrt{\eta^2+\xi^2}} e^{-i \tan^{-1} \xi/\eta}$$

$$P_2^1(s) = 3s\sqrt{1-s^2} = 3i \frac{\sqrt{1-\eta^2} \sqrt{\xi^2+1} \sqrt{1+(\xi\eta)^2}}{\eta^2+\xi^2} e^{i \tan^{-1} \xi\eta - i 2 \tan^{-1} \xi/\eta}$$

$$P_2^2(s) = 3(1-s^2) = -3 \frac{(1-\eta^2)(\xi^2+1)}{\eta^2+\xi^2} e^{-i 2 \tan^{-1} \xi/\eta}$$

$$P_3^1(s) = \frac{3}{2}(5s^2-1)\sqrt{1-s^2}$$

$$= \frac{3}{2} i \frac{\sqrt{1-\eta^2} \sqrt{\xi^2+1}}{\sqrt{\eta^2+\xi^2}} e^{-i \tan^{-1} \xi/\eta} \left\{ 5 \frac{(1-\eta^2)(\xi^2+1)}{\eta^2+\xi^2} e^{-i 2 \tan^{-1} \xi/\eta} + 2 \right\}$$

$$P_3^2(s) = 15s(1-s^2)$$

$$= -15 \frac{\sqrt{1+(\xi\eta)^2} (1-\eta^2)(\xi^2+1)}{\sqrt{\eta^2+\xi^2}^3} e^{i \tan^{-1} \xi\eta - i 2 \tan^{-1} \xi/\eta}$$

$$P_3^3(s) = 15\sqrt{1-s^2}^3$$

$$= -15i \frac{\sqrt{1-\eta^2}^3 \sqrt{\xi^2+1}^3}{\sqrt{\eta^2+\xi^2}^3} e^{-i 3 \tan^{-1} \xi/\eta}$$

Spherical Hankel Functions

The spherical Hankel function of the first kind is given by

$$h_n^{(1)}(t) = (-i)^n \frac{e^{it}}{it} \sum_{j=0}^n \left( -\frac{1}{it} \right)^j \frac{(n+j)!}{2^j j! (n-j)!} . \quad (5.36)$$

The spherical Hankel function of the second kind is given by

$$h_n^{(2)}(t) = (-i)^n \frac{e^{-it}}{it} \sum_{j=0}^n \left( -\frac{1}{it} \right)^j \frac{(n+j)!}{2^j j! (n-j)!} . \quad (5.36)$$

In this particular case, the argument of the Hankel functions is given by

$$it = kd(\eta + i\xi) , \quad (5.38)$$

or

$$t = kd(\xi - i\eta) . \quad (5.39)$$

Table 5.3 lists the higher-order Hankel functions for the specific wave function  $\psi_{mn}(\xi, \eta, \phi)_{0/S}$  given by Equation (5.14).

Table 5.3. Higher-order values of the function,

$$e^{-kd} h_n^{(1)}[kd(\xi - i\eta)]$$

---


$$h_0^{(1)}(t) = \frac{e^{-kd(1-\eta)} e^{ikd\xi} e^{-i\tan^{-1}\xi/\eta}}{kd\sqrt{\eta^2+\xi^2}}$$

$$h_1^{(1)}(t) = ih_0^{(1)}(t) \left\{ 1 - \frac{e^{-i\tan^{-1}\xi/\eta}}{kd\sqrt{\eta^2+\xi^2}} \right\}$$

$$h_2^{(1)}(t) = -h_0^{(1)}(t) \left\{ 1 - \frac{3e^{-i\tan^{-1}\xi/\eta}}{2kd\sqrt{\eta^2+\xi^2}} + \frac{3e^{-i2\tan^{-1}\xi/\eta}}{(kd)^2(\eta^2+\xi^2)} \right\}$$

$$h_3^{(1)}(t) = ih_0^{(1)}(t) \left\{ 1 - \frac{6e^{-i\tan^{-1}\xi/\eta}}{kd\sqrt{\eta^2+\xi^2}} + \frac{15e^{-i2\tan^{-1}\xi/\eta}}{(kd)^2(\eta^2+\xi^2)} - \frac{15e^{-i3\tan^{-1}\xi/\eta}}{(kd)^3\sqrt{(\eta^2+\xi^2)^3}} \right\}$$


---

### Summary

Two families of exact solutions to the scalar Helmholtz equation have been introduced--a new one in the oblate spheroidal coordinate system and its related family in the prolate spheroidal coordinate system. These solutions consist of products of spherical Hankel functions and associated Legendre functions which are not separable in their respective coordinate systems. Although nonseparability adds an additional level of complexity, it avoids the infinite-sum solutions derived upon first separating Helmholtz's equation in three dimensions. These solutions are not restricted by any approximations and are valid for any value of the product  $kd$  and at any point in space, with the obvious exception of the focal points.

We obtained the complete family of solutions by using the same Flammer transformation utilized in Chapter 3 to obtain the fundamental Gaussian solution in oblate spheroidal coordinates from a spherical wave in prolate spheroidal coordinates. We performed this conversion without resorting to the paraxial approximation or the concepts of complex point sources or complex space-time shifts. Just as the fundamental term was an exact solution to the scalar Helmholtz equation, so too are the higher-order solutions unencumbered by the paraxial approximation.

Each member of the family of solutions in the oblate spheroidal coordinate system possesses a fundamentally Gaussian amplitude distribution with higher-order amplitude terms, all of which vary upon propagation. In addition, the exponential phase is fundamentally a section of an oblate ellipsoid, also with higher-order terms. These additional phase terms are of the form  $\exp[-i \tan^{-1} \xi / \eta]$  which was shown in Chapter 4 to be the sum of the twist angle  $a$  and the sag of the oblate ellipsoid expressed as a differential twist angle  $\Delta a$ .

One way to view the family of solutions given in Equation (5.14) is as spherical wave functions with complex arguments. In the next chapter we shall compare this family of solutions to the body of literature dealing with complex-argument solutions to the scalar wave equation and include fields generated by complex point sources and complex space-time shifts. We shall also relate the geometrical model mapped out in Chapter 4 to previous work on complex rays.



## CHAPTER 6

AN HISTORICAL PERSPECTIVE ON GAUSSIAN  
BEAM PROPAGATION AND COMPLEX FUNCTIONS

When Kogelnik and Li developed their mathematical description for a propagating Gaussian beam, they defined a complex beam parameter  $q(z)$  given by Equation (2.20). This parameter is a complex phase term and a function of the wavefront radius and spot size. Since that time, complex quantities have appeared in different incarnations to describe a Gaussian beam and its propagation. In particular, the concept of a Gaussian beam arising from a complex point source has become quite popular since all the classical solutions for real point sources, such as scattering from spheres and diffraction from an edge can be converted into solutions for complex point sources by analytic continuation. While we have avoided the idea that a Gaussian beam originates from a complex point source, the body of research parallels the developments presented here. Other similarities can be noted in the study of complex-argument Hermite-Gaussian and Laguerre-Gaussian functions as well as in the idea of complex rays. We present in this chapter a comparison between this previous work and the model developed in Chapters 3 through 5.

Complex Arguments

Eigensolutions to the approximate wave equation given by Equation (2.18) are conventionally expressed as products of a Hermite polynomial of real argument and the complex Gaussian function  $\exp[-ik(x^2+y^2)/2q(z)]$ , where  $q(z)$  is the complex beam parameter of Equation (2.20). Siegman [1973] sought to express both the Hermite polynomial and exponential phase as a function of  $q(z)$  in order to make the eigenfunctions more symmetric and elegant. He did not vary the complex beam parameter so that the argument of the Hermite-Gaussian polynomials are functions of Cartesian coordinates. Siegman found that the complex-argument Hermite-Gaussian normal modes in two dimensions only are

$$\hat{\psi}_n(x, z) = \left( \frac{q_0}{q} \right)^{\frac{n+1}{2}} H_n(\sqrt{c} \cdot x) e^{-cx^2}, \quad (6.1)$$

where  $c = ik/2q$  and  $q_0$  is as given in Equation (2.24). Use of these eigensolutions led him to two interesting conclusions. The first was that these new functions are not orthonormal since their operator is not hermitian. Rather, they are biorthogonal with the eigenfunctions of the adjoint operator. That is, if  $\Phi_n(x)$  represents the adjoint solutions, then the biorthogonality relation can be expressed

$$\int_{-\infty}^{\infty} \hat{\Phi}_n^*(x) \hat{\psi}_m(x) dx = K_n \delta_{nm}, \quad (6.2)$$

where  $K_n$  is the normalization constant. The other conclusion Siegman reached was that the wavefronts of the higher-order modes of this new set were not spherical. The complex argument  $\sqrt{c} x$  in the hermite polynomials contributes an additional phase variation to the usual spherical phase of the conventional real-argument hermite polynomials. Siegman also pointed out that this procedure of employing complex arguments could be applied to the Laguerre-Gaussian polynomials.

Shin and Felsen [1976] showed that the complex-argument Hermite-Gaussian wave functions proposed by Siegman could be generated by assigning complex locations to the source points in a multipole expansion of the field. Specifically, they began with a spherical wave as the point source field and then defined a multipole field of order  $(m,n)$  as

$$G_{mn}(r,r') = \left(\frac{\partial}{\partial x}\right)^m \left(\frac{\partial}{\partial y}\right)^n \frac{e^{ik|r-r'|}}{4|r-r'|}, \quad (6.3)$$

where  $r$  is the distance from the origin to the observation point and  $r'$  is the distance between the origin and the source point. The source point is assigned a complex location such that

$$|r-r'| = [x^2+y^2+(z-ib)^2]^{1/2}. \quad (6.4)$$

The usual binomial expansion is used and the multipole field of Equation (6.3) can be written as higher-order derivatives of a complex

argument exponential. This is simply a Hermite polynomial of complex argument since

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} . \quad (6.5)$$

The result is complex-source-point multipole fields that coincide with the complex-argument beam modes of Siegman. Shin and Felsen thus established a link between higher-order complex argument beam fields and multipole expansions with complex source points.

Zauderer [1986] expanded on the use of complex argument Hermite-Gaussian and Laguerre-Gaussian beams by showing that they arise in correction terms of a perturbation expansion whose leading term is the zero-order, paraxial Gaussian beam mode. Expansions of this type have been used by other authors [Lax, Louisell and McKnight (1975), Agrawal and Pattanayak (1979), Couture and Belanger (1981), Agrawal and Lax (1983), and Takenaka, Yokota, and Fukumitsu (1985)] to develop an expression for the higher-order modes of a Gaussian beam that is an exact solution to the exact wave equation. Complex argument functions have played a pivotal role in the search for an exact Gaussian beam solution since they are, in a sense, analytically continued spherical waves which are exact solutions.

Previous work in complex-argument eigensolutions to the wave equation had been hindered by the reliance on the Cartesian coordinate

system. This difficulty was eliminated by Einziger and Raz [1987] who found a fundamental solution to the exact wave equation in oblate spheroidal coordinates, beginning in the now familiar regime of a complex-argument spherical wave. Their work comes closest to matching the conclusions of this research although their zero-order solution contains an error and they stopped short of developing an expression for the higher-order terms. Furthermore, they make use of the idea of complex space-time shifts in the generation of Gaussian beams. This bears some resemblance to complex-point-source theory which will be considered next.

#### Complex Point Sources

Kogelnik was the first to introduce an ingenious technique for obtaining a Gaussian beam from a spherical wave [Gordon 1968]. A spherical wave is an exact solution to the scalar Helmholtz equation and making the radius complex as in Equation (6.4) does not change the exact nature of the solution. After using the binomial expansion on the expression for the radius  $r$  and gathering together real and imaginary term in the exponent, the classical expression for a Gaussian beam is obtained. The expression for  $r$  is only an approximation and the resultant beam description is valid only in the paraxial region.

Deschamps [1971] echoed this idea and expanded on it. He stated that if the radius were made complex such that  $r = r' + ir''$  where the curves on which  $r'$  and  $r''$  were constant formed a system of

confocal ellipses and hyperbolas, the complex-source-point spherical wave would satisfy the exact wave equation rather than the parabolic one. This description, he felt, should be a better representation of the field. Deschamps went further to interpret the complex-source-point spherical wave as a bundle of complex rays emanating from the source. The tracing of such rays can be handled by geometrical optics methods such as matrix methods.

A number of authors have expanded on the idea of complex source points to obtain the higher-order modes for a Gaussian beam. Couture and Belanger [1981] expressed solutions to the exact Helmholtz equation as a power series and divided the Helmholtz operator into two differential operators. The first operator gave the paraxial or fundamental solution while the second represented the higher-order correction terms. They expressed these correction terms as higher-order complex-source-point spherical waves

$$A_{mn} = h_n^{(2)}(kR_c) P_n^m(\cos \theta_c) e^{\pm im\phi}, \quad (6.6)$$

where  $R_c = [\rho^2 + (z+iz_0)^2]^{1/2}$  and the complex polar angle  $\theta_c$  is given by

$$\cos \theta_c = \frac{z+iz_0}{R_c}. \quad (6.7)$$

This result is very similar to the developments of Chapter 5 although they are couched in the circular cylindrical coordinate system and Couture and Belanger provided no physical interpretations for the functions.

Both Hashimoto [1985] and Luk and Yu [1985] derived expressions for the higher-order beam modes using multipole sources. In the latter work, the authors chose an electric multipole that, when located at a complex point, generated the Hermite-Gaussian beam modes. Hashimoto simply used higher-order derivatives of the fundamental complex-source-point spherical wave to reach the same conclusions.

The complex-source-point method continues to be used in new applications. Norris [1986] employs them in the Gaussian beam summation method. The purpose here is to use Gaussian beams as an indirect means of computing the field of a point source in an inhomogeneous medium. This process avoids the singularities, such as caustics and foci, that occur in the field of a point source. Specifically, Norris describes a way of representing real point sources through complex point sources. He does this by noting that the field at an observation point due to a real point source is equivalent to a distribution of complex sources on the surface of a complex sphere centered on the point source. The Gaussian beam summation method can be used to great advantage in the study of beam refraction and other wave phenomena at an interface.

Felsen [1976] best summarized the advantages and disadvantages of the complex-source-point method. Specifically, it allows for the generation of rigorous beam solutions from real line-source or point-source solutions but the results obtained are not easily interpreted in physical terms. For this reason, he recommended a combination of the complex-point-source procedure and evanescent wave tracking to gain an understanding of the wave processes involved. He cited examples [Choudhary and Felsen (1973, 1974), Shin and Felsen (1974), Felsen and Shin (1975) and Wang and Deschamps (1974)] where this combined technique had been used to advantage in a variety of beam guiding and diffraction problems. Finally, he noted that the complex-point-source solutions provide rigorous results against which asymptotic methods can be tested.

The problem of gaining physical insight from the complex-source-point solutions is not a trivial one. In addition to evanescent wave tracking, another means of understanding the physical mechanisms involved is the use of complex rays. This subject will be considered next.

#### Complex Rays

Kogelnik (1965) showed that the laws for the transformation of Gaussian beams by an optical system are formally the same as the transformation laws obeyed by the spherical waves of geometrical optics. He did this by comparing the behavior of the complex beam



parameter  $q$  upon transfer and refraction with the radius of curvature of a spherical wave.

Consider two phase fronts of a spherical wave propagating in free space in the  $z$  direction. The radii of curvature are given by  $R_1$  followed by  $R_2$  and the perpendicular distance between them along the  $z$  axis is  $z$ . The radii are then related by the expression

$$R_2 = R_1 + z. \quad (6.8)$$

If the spherical wave passes through a thin lens of focal length  $f$ , the incident wave, of radius  $R_1$ , is converted into another spherical wave of radius  $R_2$ . This transformation is given by

$$\frac{1}{R_2} = \frac{1}{R_1} - \frac{1}{f}. \quad (6.9)$$

The rays associated with the spherical wave are perpendicular to the wavefront. The position  $x$  and slope  $x' = dx/dz$  of a paraxial ray are therefore related to the radius of the wavefront by

$$R = \frac{x}{x'}. \quad (6.10)$$

By tracing a paraxial ray through an optical system, one can determine the transformation of an incident wave  $R_1$  into a resultant wave  $R_2$ . This can be done using the ray matrix with elements  $A$ ,  $B$ ,  $C$ , and  $D$  for the system. The overall transformation is then given by

$$R_2 = \frac{AR_1 + B}{CR_1 + D} \quad (6.11)$$

This entire development can be echoed for a Gaussian beam by using the complex beam parameter  $q$ . We recall from Chapter 2 that  $q$  is given by

$$\frac{1}{q} = \frac{1}{R} - \frac{i\lambda}{\pi w^2} \quad (6.12)$$

where  $R$  is the radius of curvature of the beam and  $w$  is the beam spot size. We found in Chapter 2 that the value of  $q$  a distance  $z$  from the plane of the waist is given by

$$q_2 = q_0 + z \quad (6.13)$$

where

$$q_0 = \frac{i\pi w_0^2}{\lambda} \quad (6.14)$$

and  $w_0$  is the beam spot size at the waist. Because this transfer law is linear, we can define the relationship between an input parameter  $q_1$  and an output parameter  $q_2$  separated by a distance  $z$  as

$$q_2 = q_1 + z \quad . \quad (6.15)$$

After passing through a thin lens, the beam spot size remains unchanged although the radii of the input and output phase fronts are transformed in the same way as the radii of a spherical wave. Therefore, the transformation of the complex beam parameter through a thin lens is given by

$$\frac{1}{q_2} = \frac{1}{q_1} - \frac{1}{f} \quad . \quad (6.16)$$

Comparing Equations (6.15) and (6.16) with (6.8) and (6.9), we note that the laws for the transformation of a Gaussian beam upon refraction and transfer are the same as for a spherical wave with the complex beam parameter behaving in the same manner as a wavefront radius. We can therefore construct a Gaussian beam analog to Equation (6.10) which is given by

$$q = \frac{X(z)}{X'(z)} \quad , \quad (6.17)$$

where  $X'(z) = dX(z)/dz$ . Finally, we can construct a transformation law for a Gaussian beam through an optical system using the ray matrix which can be written as

$$q_2 = \frac{Aq_1 + B}{Cq_1 + D} \quad . \quad (6.18)$$

This is the Gaussian beam analog to Equation (6.11).

The parameter  $X(z)$  has been identified by Arnaud [1969] as a complex ray. He interpreted the complex ray as the skew-line generator of the Gaussian hyperboloid, that is, the surface of constant amplitude. He further identified its modulus and phase with the beam spot size and the phase of the on-axis field respectively. When projected onto a plane perpendicular to the  $z$  axis, this skew/complex ray resolves into two orthogonal components,  $A(z)$  and  $B(z)$ , in much the same way as shown in Figure 4.3. The expression for  $X(z)$  can now be written

$$X(z) = A(z) + iB(z) \quad . \quad (6.19)$$

The functions  $A(z)$  and  $B(z)$  are the  $x$  and  $y$  projections respectively of the skew line  $l$ . With the aid of Figure 4.3, they are determined to be

$$\begin{aligned} A(z) &= W \cos(\phi + \alpha) \\ B(z) &= W \sin(\phi + \alpha) \quad . \end{aligned} \quad (6.20)$$

At this point, we note two differences in notation between Arnaud [1985] and the work presented here. First, he measures the azimuthal angle  $\theta$  from the positive  $y$  axis instead of from  $x$ , such that  $\theta = 90 + \phi$ . The second and more subtle difference lies in the

phase of the complex ray. As a result of Arnaud's notation, the ray's phase at any point in the beam will depend upon the azimuthal angle  $X(z)$  as it traverses the plane of the waist. The mathematical expression for the fundamental mode of a Gaussian beam does not contain this extraneous angular dependence. Therefore, in order for the phase of the complex ray to correspond with the beam's mathematical description, one must calculate the difference between the ray's phase in one plane and its phase in the plane of the waist. To do this when performing a beam trace through an optical system requires an additional procedure using descriptive geometry. This involves projecting line ON in Figure 4.3 onto the x-y plane. Note that the notation used here does this automatically.

To find an explicit expression for  $X(z)$ , we shall consider a skew line with a zero azimuthal angle in the plane of the waist and assume that the origin of the  $z$  axis is coincident with the waist. Equation (6.20) then becomes

$$\begin{aligned} A(z) &= W \cos \Delta\Phi \quad , \\ B(z) &= -W \sin \Delta\Phi \quad , \end{aligned} \quad (6.21)$$

where we have noted that  $\alpha = -\Delta\Phi$ . From Figure 4.3 and Equations (4.7) and (4.5), we find that

$$\begin{aligned} A(z) &= w_0 \quad , \\ B(z) &= -iz \tan \delta = -iz \frac{w_0}{z_0} \quad . \end{aligned} \quad (6.22)$$

Noting from Equation (3.36) that  $z_0 = kw_0^2/2$ , we can express  $X(z)$  as

$$X(z) = w_0 - \frac{iz}{kw_0} \quad , \quad (6.23)$$

where  $k$  is the propagation constant for the medium and  $w_0$  is the beam spot size at the waist. The modulus of  $X(z)$  is written

$$|X(z)| = \left[ w_0^2 + \left( \frac{2z}{kw_0} \right)^2 \right]^{1/2} \quad (6.24)$$

This is precisely the same as Equation (3.35), and we see that the modulus of the complex ray equals the beam spot size,  $w$ . Likewise, the phase difference represented by  $X(z)$  is given by

$$\Delta\Phi = - \tan^{-1} \left( \frac{2z}{kw_0^2} \right) \quad , \quad (6.25)$$

which is simply  $-\alpha$  from Equation (4.8).

Arnaud goes on to demonstrate how the complex ray can be used to trace Gaussian beams by simply treating  $A(z)$  and  $B(z)$  as two real paraxial rays. He further remarks that the quantity

$$n \left[ \frac{\partial A(z)}{\partial z} B(z) - A(z) \frac{\partial B(z)}{\partial z} \right] = \frac{2}{k} \quad , \quad (6.26)$$

is invariant as the beam propagates in free space. The term in brackets is therefore the Lagrange invariant for the beam. Finally,

he points out that there are two symmetrical skew rays going through a given point on the beam surface, one of which can be considered rotated clockwise, the other counterclockwise. Due to symmetry, the tangent to the beam profile at that point is the bisectrix of the skew rays. He goes on to show that this bisectrix intersects the optical axis at the wavefront center of curvature. Therefore, by tracing  $A(z)$  and  $B(z)$  through an optical system one can obtain all the first-order properties of the beams, that is, beam sizes, waist locations, and radii of curvature.

The use of complex rays to trace a Gaussian beam through an optical system should be compared to the  $W-W_0$  diagram technique pioneered by Roland Shack and described in Appendix A. This technique projects the skew line together with the optical axis onto a transverse plane. Lens and mirrors are represented by a single point on the diagram and distances, beam and waist sizes, and radii of curvature are all calculated from simple geometric principles. This method derives its power and simplicity from the tracing of a single, real, skew ray as opposed to two real rays as Arnaud has done.

The aforementioned research in Gaussian beam propagation represents some of the highlights in describing a Gaussian beam and its propagation in a homogeneous medium as well as extending this description to other wave phenomena involving Gaussian beams such as scattering and diffraction. While similarities abound between this previous work and the research presented here there are four areas in which they differ clearly.

The first difference is the choice of coordinate system. With the exception of Einziger and Raz [1987], all of the previous work has concentrated on the Cartesian, polar, or circular cylindrical coordinate systems. Describing a hyperbola is cumbersome in these systems and makes the expressions more difficult to use in describing Gaussian beams. It also makes geometrical interpretations more unwieldy. Like Einziger and Raz, we have used the oblate spheroidal coordinate system in which the coordinate surfaces themselves describe the beam's amplitude and wavefronts. Unlike these authors, however, we have expanded the description to include the higher-order Gaussian beam modes and have developed a geometrical model as well.

Beyond the choice of coordinate systems, the most important differences between this work and its predecessors reside in the lack of approximations and in the closed-form solution as opposed to an infinite-series solution such as a perturbation expansion. Most authors have made use of the paraxial approximation or the parabolic wave equation, or both, to effect a solution to the scalar wave equation. Some have used the exact form of Helmholtz's equation but have developed only infinite series as a solution to the equation. They then compare the first term of the series to the zero-order mode of the traditional Hermite- or Laguerre-Gaussian polynomials. The solution here is closed form, that is, each term or mode in the family is an exact solution to the wave equation. This represents a significant increase in both simplicity and accuracy for the mathematical expression with which to study Gaussian beam propagation.



Finally, we have developed a geometrical interpretation for the mathematical expression. As we pointed out earlier, it is very difficult to grasp the physical significance of a point source at a complex location or complex rays. And, despite the usefulness of complex-argument functions to model complicated physical processes, they provide no insight into the wave phenomena. In contrast, we have generated a simple geometrical model by assembling many previous developments, expanding on others, and deriving new relationships. This model not only interprets our mathematical expression in the light of geometrical optics, but it affords the user greater physical insight into Gaussian beam propagation. Obviously, this research leaves many questions regarding the beam propagation and origin unanswered. These issues will be taken up in the next chapter.

## CHAPTER 7

## IDEAS FOR FUTURE RESEARCH

The model--both the mathematical expression and the geometrical construct--presented in this work represents a new way to express and conceive of propagating Gaussian beams. It provides some novel insights into the nature of these beams but raises even more questions about their origin as well as their behavior in an optical system. Specifically, there are two major issues that are left to be addressed. The first is the extension of these scalar-wave-equation solutions to the vector wave equation; the second is the application of these solutions to the resonator integral problem to ascertain whether Gaussian beams with elliptical wavefronts can indeed be generated by spherical resonators. There are, in addition, other topics of interest which will now be delineated.

Vector Wave Equation

As pointed out in Chapter 1, this research has been restricted to the scalar approximation of electromagnetic wave theory; that is, only the scalar amplitude of one transverse component of either the electric or magnetic field is considered in the study of the beam's propagation. In order to satisfy Maxwell's equations for a charge-free region, this scalar field must have a zero divergence. However, as Mukunda, Simon and Sudarshan [1985] have shown, this

condition requires that there be no spatial variation of the electric field in its direction of polarization. Gaussian beams by definition have a spatially modulated field, and therefore, a strictly scalar solution will not suffice to completely describe the beam's propagation; solutions to the vector wave equation must be investigated. Nonetheless, the scalar wave functions represent the point of origin of this expanded analysis since the only known general method of obtaining solutions to the vector wave equation is by applying certain vector differential operators to the scalar wave function.

The vector wave equation is given by

$$\nabla(\nabla \cdot \vec{A}) - \nabla_x \nabla_x \vec{A} + k^2 \vec{A} = 0, \quad (7.1)$$

where  $\vec{A}$  represents the vector wave functions. Since the gradient operator commutes with the Laplacian, we can immediately obtain an irrotational solution to equation (7.1) from a solution  $\psi$  of the scalar wave equation by taking the gradient of  $\psi$ . Therefore,

$$\vec{A} = \nabla \psi, \quad (7.2)$$

is an irrotational solution to the vector wave equation.

Of greater significance, however, are the vector wave functions that satisfy the equation,

$$\nabla_{\mathbf{x}} \nabla_{\mathbf{x}} \vec{A} - k^2 \vec{A} = 0. \quad (7.3)$$

The solenoidal nature of the functions is ensured by writing the solutions in the form,

$$\vec{A} = \nabla_{\mathbf{x}}(\vec{a} \psi), \quad (7.4)$$

where  $\psi$  is again the scalar wave function derived in this research and  $\vec{a}$  is either a constant vector or the position vector  $\vec{r}$ .

From the foregoing we see that development of the vector wave functions from the scalar solutions derived in this research is a straightforward task. These results would allow a complete and exact description of a propagating Gaussian beam, provide the necessary framework for studying polarization effects in high-numerical-aperture beams, and provide much insight into the diffraction of Gaussian beams.

#### Resonator Analysis and Beam Origin

This research has shown that Gaussian beams with oblate ellipsoidal wavefronts can exist and propagate as a scalar phenomenon. At the same time, it brings up the intriguing question of whether such beams can be generated by optical resonators consisting of spherical or ellipsoidal mirrors. The latter type of resonator

would seem intuitively obvious since the mirrors would coincide with surfaces of constant phase in the beam. However, ideal resonators are conceived of having perfect spheres as end mirrors, regardless of the steady-state beam spot size on the mirror. Therefore, the diffraction of an oblate ellipsoidal wavefront by a perfectly conducting sphere, with an oblate ellipsoidal mirror as a special case, is the problem most properly addressed.

It is interesting to note as shown in Chapter 2 that Slepian and Pollack [1961] found prolate spheroidal wave functions to be solutions to the resonator integral given by Equation (2.8). In light of the strong relationship between Gaussian amplitude wave functions and prolate spheroidal wave functions indicated here, there may be a definite correlation between these oblate-spheroidal solutions to the wave equation and the fields diffracted by a spherical mirror. It is not clear whether this is a scalar or vector diffraction phenomenon; nor is it clear that the usual Fresnel approximation can be made.

Finally, it would be instructive to compare both modulus and phase functions of the traditional resonator modes--both Hermite-Gaussian and Laguerre-Gaussian--with this new family of Gaussian amplitude functions. Each of the circular Laguerre modes represents a linear combination of a certain number of the rectangular Hermites and vice versa. It is quite probable that each of these new Gaussian modes will prove to be a linear combination of either of the traditional modes. Furthermore, the presence of Brewster angle windows, as well as any mirror tilts or misalignments,

usually cause any real laser to oscillate in the rectangularly symmetric Hermite form rather than in circularly symmetric modes. It would therefore be most useful to compare the new Gaussian modes with the traditional Hermite forms.

### Diffraction

In Chapter 2, we discussed the approximations inherent in the Kogelnik and Li description of a Gaussian beam. Specifically, this traditional expression springs from the resonator analysis of Boyd and Gordon which relied upon the Rayleigh-Sommerfeld diffraction formulation. This formula approximates the point spread function, which is the normal derivative of a spherical wave, as a spherical wave. Since a detailed analysis of the resonance phenomena is beyond the scope of this work, this issue has not been addressed. However, the Flammer transformation employed here appears promising as a basis for reconstructing an exact diffraction formula.

The point spread function given in Equation (2.36) can also be written,

$$\hat{n} \cdot \nabla \left( -\frac{e^{ikr}}{kr} \right) = 2\pi \frac{z}{r} i h_1^{(1)}(kr) \quad , \quad (7.5)$$

where we have chosen the  $z$  axis as the normal direction for the sake of simplicity. The equation for a field  $u(x',y',z')$  in an  $x'-y'$  plane a distance  $z'$  from the initial field can be expressed as

$$u(x',y',z') = \iint_{-\infty}^{\infty} 2\pi \frac{(z'-z)}{r} h_1^{(1)}(kr) u(x,y,z) dx dy, \quad (7.6)$$

where  $u(x,y,z)$  is the initial field and  $r = [(x'-x)^2 + (y'-y)^2 + (z'-z)^2]^{1/2}$

The integrand in Equation (7.6) contains a Hankel function. This is a complex sum of a spherical Bessel function of the first kind,  $j_1(kr)$ , and a Neumann function,  $y_1(kr)$ . Assuming that the initial field  $u(x,y,z)$  is well-known and well-behaved, solving this integral equation hinges upon the behavior of  $h_1^{(1)}(kr)$ . Both Watson [1980] and Oberhettinger [1984] discuss ways of solving integrals of this type. In particular, Oberhettinger substitutes a modified Bessel function of the third kind,  $K_{n+1/2}(kr)$ , for  $h_n^{(1)}(kr)$  since the former converges as  $r \rightarrow \infty$ . This substitution is given as

$$h_1^{(1)}(kr) = -\frac{i}{2\pi} e^{-i\pi 3/4} \sqrt{\frac{\pi}{2z}} K_{3/2}(-iz). \quad (7.7)$$

The next step in the solution, as illustrated by Watson, would be to make an integral substitution for the  $K$  Bessel function, but it is here that the insolubility of Equation (7.6) becomes apparent. After substituting Equation (7.7) into (7.6), we see that

$$u(x', y', z') = \iint_{-\infty}^{\infty} ik(z'-z) u(x, y, z) \cdot \sqrt{\frac{\pi}{2}} \frac{K_{3/2}(-ikr)}{(-ikr)^{3/2}} dx dy . \quad (7.8)$$

The integral substitution for  $K$  is given by Watson as

$$\frac{K_{\nu}(z)}{z^{\nu}} = \frac{1}{2^{\nu+1}} \int_0^{\infty} \exp \left\{ -t - \frac{z^2}{4t} \right\} \frac{dt}{t^{\nu+1}} , \quad (7.9)$$

provided that  $R(z^2) > 0$ . In the case of Equation (7.8), the argument of the  $K$  function is purely imaginary since the radius  $r$  is real. Since the integral in Equation (7.9) does not converge, the diffraction integral of Equation (7.8) cannot be solved as it is expressed. This difficulty would be avoided completely if the argument of  $K_{\nu}(z)$  were complex. In other words, if the spherical waves of Equation (7.5) were transformed into complex-point-source waves, the diffraction integral could conceivably be solved without approximations. This is the same transformation used in Chapter 3. After utilizing the transformation of Equation (3.22), the radius  $r$  would become complex and the point spread function would be transformed from the normal derivative of a spherical wave into the normal derivative of a Gaussian beam. The suggestion of this new Green's function was broached by Einziger and Raz [1987].



Gaussian Beams with Elliptical Cross-Section

The geometrical model of a Gaussian beam developed in Chapter 4 utilized the concept of a skew-line generator of a hyperboloid of revolution to simulate circularly symmetric beams. This construct is not unique to figures of revolution. A general hyperboloid of one sheet, that is, one having an elliptical cross section in a plane perpendicular to the  $z$  axis, is also a ruled surface and can be generated by the motion of a skew line about the  $z$  axis. This general model can then be used to represent the elliptical Gaussian beams produced by semiconductor lasers.

In the case of a general hyperboloid, the skew line joins pairs of points with a constant difference of azimuthal angle on two equal ellipses in parallel planes that are symmetrically displaced with respect to the origin. The planes containing the ellipses are, of course, perpendicular to the  $z$  axis. The coordinates of the points on the skew line are given by

$$P: [a \cos(\phi - \alpha), b \sin(\phi - \alpha), c] ,$$

and

(7.10)

$$P': [a \cos(\phi + \alpha), b \sin(\phi + \alpha), -c] .$$

The equation of the elliptic hyperboloid can be written as

$$\frac{x^2}{a^2 \cos^2 a} + \frac{y^2}{b^2 \cos^2 a} - \frac{z^2 \tan^2 a}{c^2} = 1, \quad (7.11)$$

where  $a$  and  $b$  are the semi-major axes of the ellipse in the  $x$  and  $y$  directions, respectively. Further, the sign of  $a$  can be changed without altering the equation. This means that the general hyperboloid can be produced by a positive or negative twist, just as in the case of the hyperboloid of revolution.

#### Optical Design and Aberration Theory

This present work can be expanded not only into areas of physical optics such as diffraction and general electromagnetic wave theory, but it also has potential for use in geometrical optics and optical design. Both the plane wave and the spherical wave have long been used as ideal beams in optical system design. They both have a constant amplitude on a surface transverse to the direction of propagation and are, in general, well-behaved exact solutions to the scalar wave equation. Up to now, there have been no similarly well-behaved functions with a non-uniform amplitude to serve as an ideal model for beams whose behavior and amplitude distribution fit neither the plane wave nor the spherical wave models, such as beams with Gaussian apodized profiles.

Use of this new Gaussian amplitude function is greatly enhanced by the straight-line geometrical model given in Chapter 4. The power and elegance of using the skew line to predict the first order properties of a Gaussian beam in an optical system is amply demonstrated by Shack's  $W-W_0$  diagram explained in Appendix A.

Beyond the first-order design advantages, this new model of a Gaussian beam serves as a starting point for studying the aberrations of propagating Gaussian beams, but the unusual nature of the ideal wavefront might require a different architecture for the wave aberration function. Wave aberration theory deals with the difference between an aberrated wavefront and a reference sphere centered on the ideal focal point. For the case of a beam with an elliptical wavefront, the wave aberration function could consist of the difference between the aberrated wavefront and an ideal oblate ellipse whose "focus" is the plane of the waist and therefore at a fixed location. Another possibility would be to maintain the current wave aberration function definition but treat the difference between the reference sphere and the oblate ellipsoid as a first-order aberration, somewhat in the same manner as Petzval curvature. In this situation, the reference sphere radius would be the vertex radius of the ellipse. This method has the distinct disadvantage of allowing the location of the center of the reference sphere to vary in a nonlinear

fashion along the axis as the wavefront propagates. A third possibility for the formulation of a wave aberration function would be to center the reference sphere on axis in the plane of the beam waist after the manner of Hopkins [1952]. The radius of the reference sphere would then be the length of the skew line to the ideal oblate ellipsoid.

In addition to redefining the wave aberration function, a beam with an elliptical wavefront and Gaussian amplitude distribution poses other questions in the study of aberrations. These issues include, but are not confined to, the effect of refraction by a plane, sphere, parabola, cylinder or ellipsoid on the beam; the effect of reflection from a plane, sphere, cylinder, parabola or ellipsoid; the effect on the beam wavefront and amplitude upon propagation through a gradient index medium, such as the atmosphere; and the effect produced by diffractive elements such as gratings and holographic elements.

#### Summary

These new Gaussian functions coupled with the skew-line geometrical model provide a starting point for a number of exciting new areas of research in both physical optics and geometrical optics. The physical optics areas include exact Gaussian-amplitude solutions to the vector wave equation and a reevaluation of the resonator integral problem in order to investigate the beam's origin. On the other hand, future research in geometrical optics includes the study

of how aberrations affect the propagation of Gaussian beams in optical systems as well as in gradient index materials. In addition, the model can be expanded to include elliptical Gaussian beams in order to examine the propagation of Gaussian beams generated by semiconductor lasers.

## CHAPTER 8

## CONCLUSIONS

This research has generated an entirely new family of exact, closed-form solutions to the scalar wave equation in the oblate spheroidal coordinate system. Although we have used them here to model the Gaussian-amplitude fields that are, in general, produced by a laser resonator, their usefulness reaches beyond the confines of resonator physics. The (0,0) order of this family has a Gaussian amplitude distribution and a phase term which is basically a section of an oblate ellipsoid. This wave function has a singularity at the focal ring of the oblate spheroidal coordinate system in the same fashion that a spherical wave has a singularity at its origin. Indeed, the utility and fundamental importance of this new wave function can be compared with both a plane wave and a spherical wave. It represents a way to describe exactly the propagation of beams with a non-uniform amplitude, something which neither the spherical wave nor plane wave functions permit. Such beams can be generated by other means than a laser resonator, for example, incoherent, monochromatic beams on which a Gaussian apodization function has been imposed.

Next, we expanded the derivation of this new class of functions to include all of the higher-order terms of these Gaussian amplitude expressions. To do so, we used an elegant transformation

introduced by Carson Flammer [1957] that, in effect, turns a spherical wave into a Gaussian beam. This is not a coordinate transformation but allows for a transmutation from one real coordinate system (the prolate spheroidal coordinate system) to another (the oblate spheroidal coordinate system) by changing the nature of two coordinate parameters. This is a very powerful technique and, judging from what it allows us to do in this problem, should be investigated further, especially in the area of diffraction theory.

In addition to the derivation of this new family of solutions to the scalar Helmholtz equation, we have developed a geometrical model for the (0,0) order. The purpose of this model was to provide physical insight into the various mathematical terms that appear in the expression. Specifically, we used the skew line generator of a hyperboloid of revolution to mimic a geometrical-optical ray. A major limitation of the concept of a skew line as a ray is that the skew line is not the gradient of the suggested wavefront--an oblate ellipse with a Gaussian amplitude distribution. Another difficulty arises from the dual nature of the skew line, that is, that two skew lines of opposite twist can generate the same hyperbola. The concept is useful nonetheless and allows us to trace equal distances from one oblate ellipse to another just as the distance between two adjacent wavefronts is constant along their corresponding rays. The skew line

also permits an elegant interpretation to the phase term  $\exp [-i \tan^{-1} \xi / \eta]$  as the sag of the oblate ellipse as measured along the skew line. This suggests that a fruitful setting for the study of Gaussian beams is not in a meridional plane despite the appearance of rotational symmetry. Rather, the use of a nonorthogonal coordinate system would more effectively permit the implementation of this model.

Finally, a nonorthogonal coordinate system based on the skew line is presented as a more profitable environment in which to study this new family of wave functions. This coordinate system makes use of the double skew line generators of a single-sheet hyperboloid as coordinate curves on the hyperbola. The third coordinate curve is the oblate ellipse. Nonorthogonality need not be viewed as a disadvantage here since it is possible to revert to the orthogonal oblate spheroidal coordinate system by taking the sum and difference of the two skew line vectors. In the meantime, we have provided a rigorous, unambiguous framework for studying the propagation of Gaussian beams based on the simplicity of straight lines.

This research provides a new solution to a very old problem in wave propagation. Furthermore, it has been interpreted against a background that supplies a considerable amount of physical insight into the wave phenomena. It is hoped that this combined approach will extend the usefulness of this research into new areas of wave propagation, diffraction, and scattering.



**APPENDIX A:**

**FIRST ORDER GAUSSIAN BEAM PROPAGATION**

**USING SKEW LINE MODEL**

A number of techniques have been suggested and outlined for the first-order design and analysis of laser systems. That is, as a Gaussian beam propagates through a system of refractive or reflective elements, systematic methods exist for predicting beam and waist sizes, waist locations and wavefront curvatures (in the classical description). Unfortunately, most of these methods are unwieldy and difficult to use, thereby providing little insight into the behavior of the beam as it propagates through the system.

Kogelnik (1965) developed a set of equations to describe the propagation and refraction of Gaussian beams based on the classical model. In particular, he derived the transformation law for the complex beam parameter  $q$ , discussed in Chapter 2. This law, called the ABCD law, allows one to trace Gaussian beams through any optical system defined by the elements A, B, C, and D of its ray matrix.

The transformation law for  $q$  leads to analogies between transmission line and laser beam problems. As a result, Gaussian beam behavior upon refraction and transfer can be represented graphically on a circle diagram similar to impedance charts, such as the Smith chart. Such a diagram was first proposed by Collins (1964) and Li (1964) and is shown in Figure A.1. Kogelnik and Li (1966) described the relationship between this circle diagram and  $q$  and cited related work on the same concept.

Arnaud (1969, 1973, 1980) showed that a fundamental Gaussian beam propagating through an optical system can be generated by the rotation about its axis of a skew ray that obeys the laws of

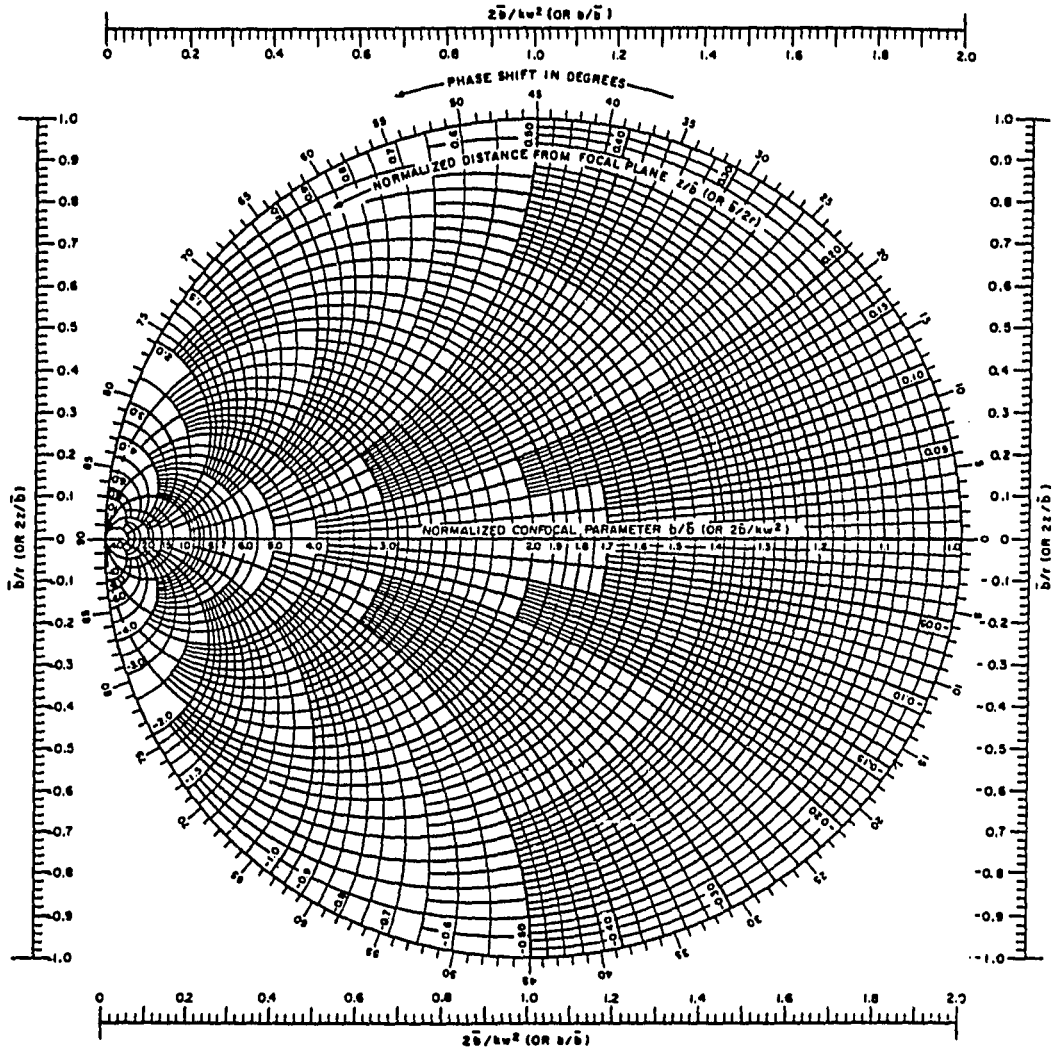


Figure A.1 The Gaussian beam chart

geometrical optics. The projections of this skew ray onto two perpendicular meridional planes form the real and imaginary parts of a complex ray,  $X(z)$ . The modulus of  $X(z)$  equals the beam size at any point and the phase gives the on-axis phase of the beam in the classical description. The ability to represent a Gaussian beam by two rays provided a method of analyzing laser systems using ordinary ray tracing.

A far more straightforward, insightful, and powerful method for investigating the behavior of Gaussian beams in an optical system was introduced by Shack (1982). This method, referred to hereafter as the  $WW_0$  diagram, utilizes many of the same concepts introduced by Delano (1963) for use in the  $Y\bar{Y}$  diagram, a technique for synthesizing optical systems. Both procedures trace a skew line or ray through an optical layout. In the case of the  $Y\bar{Y}$  diagram, this skew ray provides both marginal and chief ray information. It is similar to Arnaud's complex ray,  $X(z)$ , in that the marginal and chief rays are the projections of the skew ray in two perpendicular meridional planes. The  $WW_0$  diagram, on the other hand, traces the rectilinear generator of the Gaussian beam's hyperbolic envelope in terms of the beam parameters in the spherical Gaussian beam expression.

The traditional scalar description of the fundamental mode of a Gaussian beam can be written:

$$u(\rho, z) = \frac{w_0}{W} \exp \left\{ -i \left( kz - \Phi + \frac{k\rho^2}{2R(z)} \right) - \frac{\rho^2}{w^2(z)} \right\} \quad (\text{A.1})$$

where  $u(\rho, z)$  represents the light amplitude,  $R(z)$  is the radius of curvature of the spherical wavefront,  $w$  is the radius of the beam where the amplitude is  $1/e$  its on-axis value and  $z$  is the distance along the axis from the waist. Further, the phase term  $\Phi$  is given by

$$\Phi = \tan^{-1} \frac{z}{z_0} \quad . \quad (\text{A.2})$$

We also have the relationships,

$$R(z) = - \frac{(z^2 + z_0^2)}{z} \quad , \quad (\text{A.3})$$

and,

$$w^2(z) = w_0^2 \left[ 1 + \frac{z^2}{z_0^2} \right] \quad . \quad (\text{A.4})$$

The beam parameters  $z_0$  and  $w_0$  are known as the Rayleigh range and beam waist respectively and are related by

$$z_0 = \frac{kw_0^2}{2} \quad , \quad (\text{A.5})$$

where  $k$  is the propagation constant of the medium. Equations (A.3) and (A.4) can also be written,

$$w_0^2 = - \frac{w(z)^2}{1 + \left( \frac{kw^2(z)}{2R(z)} \right)^2} \quad (\text{A.6})$$

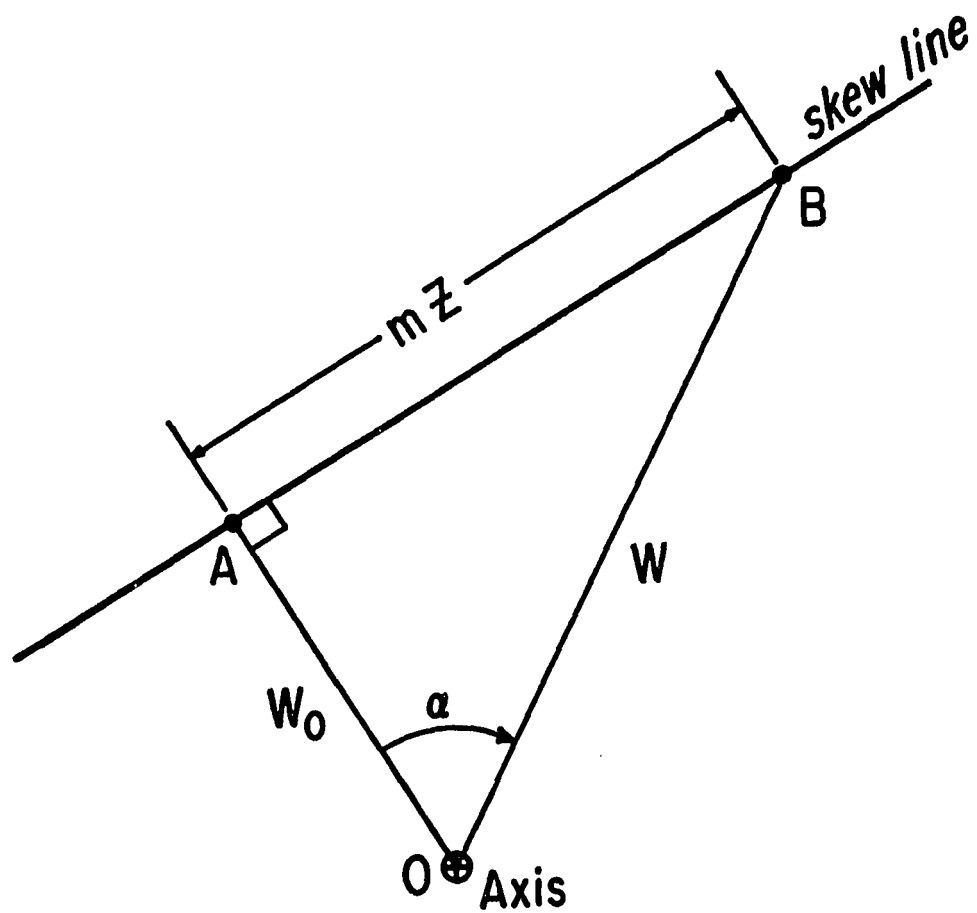


Figure A.2 The skew line and optic axis projected onto a plane

$$z = \frac{-R(z)}{1 + \left(\frac{2R(z)}{kw^2(z)}\right)^2} \quad (A.7)$$

The  $WW_0$  diagram reduces the complexity of these equations by a simple and powerful graphical technique.

The projection of the skew line and the axis of rotation onto a plane forms the basis of the  $WW_0$  diagram. The result of this projection is shown in Figure A.2, where the axis appears as a point and becomes the origin of the diagram and the skew line occurs as a straight line. The shortest distance from the point to the line, that is, from the axis to the rectilinear generator, is a perpendicular and corresponds to the beam waist,  $w_0$ . Some other point, B, on the skew line represents a different cross-sectional plane whose beam width  $w$  is the distance from that point to the origin (axis). Further, the length of the skew line from this arbitrary point to the beam waist is proportional to the actual axial distance between cross-sectional planes. Finally, the angle, AOB, formed by the construction lines,  $w_0$  and  $w$ , is the phase difference  $\Delta\Phi$  between the spherical wavefronts of the beam at those points. It is also the twist angle  $\alpha$  discussed in Chapter 3.

To find the value of the proportionality constant  $m$ , let  $\alpha = 45^\circ$  in Figure A.2. At this point,  $z = z_0$ ,  $w = \sqrt{2} w_0$  and,

$$\frac{mz_0}{w} = \sin\alpha \quad (A.8)$$

Substituting equation (A.5) into (A.8) leads to,

$$m = \frac{w_0}{z_0} = \frac{2}{kw_0} \quad (\text{A.9})$$

Note that the axial distance is proportional to the area of triangle AOB. That is,

$$K_{\Delta AOB} = \frac{1}{2} \left( \frac{2z}{kw_0} \right) w_0 = \frac{z}{k} \quad (\text{A.10})$$

The sign convention for the area is positive when the area is swept by the vector  $\vec{OB}$  rotating clockwise about the origin. Area is therefore considered positive for negative  $a$ .

A very useful but little known characteristic of Gaussian beams is the relationship between two surfaces whose wavefronts have a phase difference of  $90^\circ$ . These surfaces have radii of curvature that are equal and opposite such that the center of curvature of one lies on the surface of the other. Shack calls these "sister surfaces", and they are uniquely determined by the  $WW_0$  diagram. Since they have a  $90^\circ$  phase difference, any two perpendicular beam width lines possess a "sister" relationship.

Two representative sister surfaces are shown in Figure A.3 with its corresponding  $WW_0$  diagram given in Figure A.4. The angle between  $W_2$  and  $W_1$  is  $90^\circ$  so the sum of twist angles  $a_1$  and  $a_2$  must also be  $90^\circ$ . This leads to the requirement that the cotangent of  $a_1 - a_2$  equal zero. The difference between the two twist angles is required to preserve the sign convention for negative distance. This condition can be expressed as



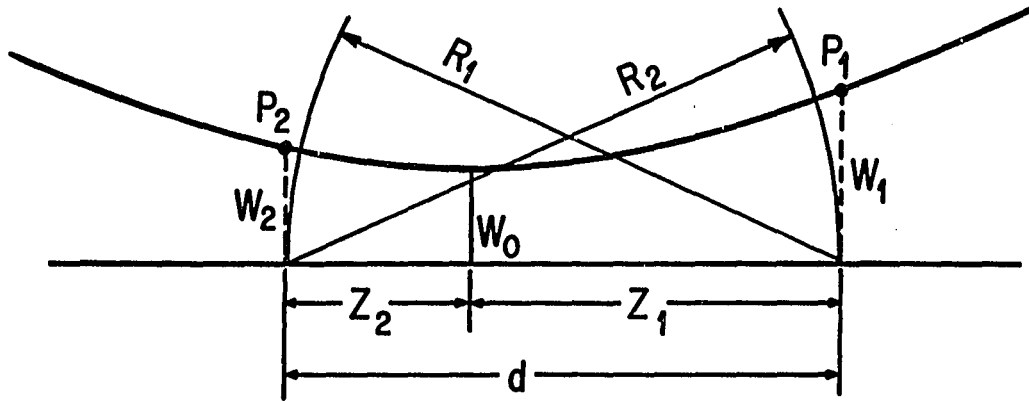


Figure A.3 Sister surfaces in a Gaussian Beam

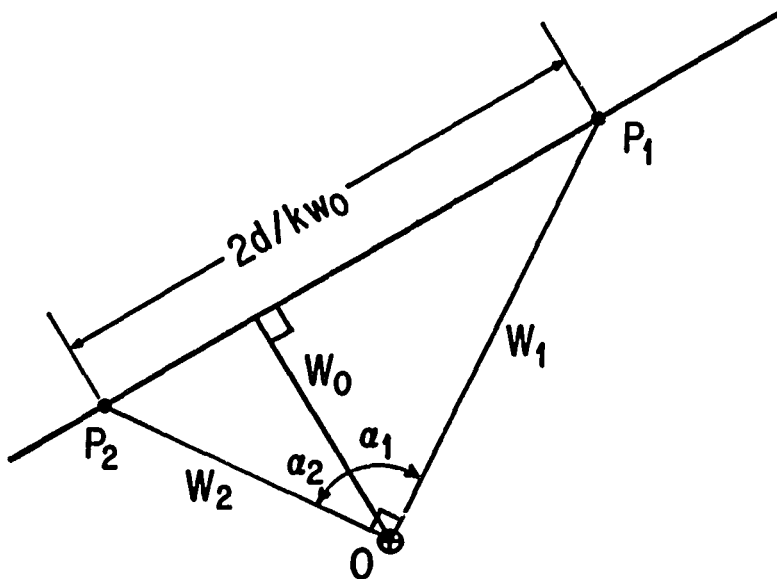


Figure A.4 Sister surface representation on the  $W$ - $W_0$  diagram

$$\cot(\alpha_1 - \alpha_2) = \frac{\cot \alpha_1 \cot \alpha_2 + 1}{\cot \alpha_1 - \cot \alpha_2} = 0 \quad (\text{A.11})$$

The numerator in (A.11) must equal zero or,

$$-1 - \frac{1}{\tan \alpha_1} = \frac{1}{\tan \alpha_2} \quad (\text{A.12})$$

Substituting Equation (A.2) into (A.12) results in an equation for  $z_1$  and  $z_2$  in terms of the Rayleigh range where

$$-z_1 z_2 = z_0^2 \quad (\text{A.13})$$

We can now find expressions for  $R_1$  and  $R_2$  by substituting Equation (A.13) into (A.3). This manipulation results in

$$R_1 = - \frac{(z_1^2 + z_0^2)}{z_1} = -(z_1 - z_2) \quad (\text{A.14})$$

We note by inspection of figure A.3 that the value of  $R_2$  is defined to be:

$$R_2 = z_1 - z_2 \quad (\text{A.15})$$

Therefore, from (A.14) and (A.15) we see that

$$R_1 = -R_2 \quad (\text{A.16})$$

Writing Equation (A.4) explicitly for planes 1 and 3 gives

$$\begin{aligned}
 W_1 &= \left[ \frac{2z_o}{k} \left( 1 + \frac{z_1^2}{z_o^2} \right) \right]^{1/2}, \\
 W_2 &= \left[ \frac{2z_o}{k} \left( 1 + \frac{z_2^2}{z_o^2} \right) \right]^{1/2}.
 \end{aligned}
 \tag{A.17}$$

Multiplying  $W_1$  and  $W_2$  and utilizing the expression for  $z_o^2$  given in Equation (A.13), we see that

$$W_1 W_2 = \frac{2z_o}{k} \left[ \left( \frac{z_2 - z_1}{z_2} \right) \cdot \left( \frac{z_1 - z_2}{z_1} \right) \right]^{1/2}.
 \tag{A.18}$$

Substituting (A.14) and (A.13) into (A.18) yields

$$W_1 W_2 = \frac{2R_2}{k} = \frac{-2R_1}{k}.
 \tag{A.19}$$

Finally, similar manipulations of Equation (A.17) shows that

$$\frac{1}{W_1^2} + \frac{1}{W_2^2} = \frac{1}{W_0^2} \quad . \quad (\text{A.20})$$

From the above equations, particularly (A.19), it becomes apparent that the radius of curvature at any point in the beam can be obtained using sister surface relationships.

After applying this development to the  $WW_0$  diagram in Figure A.4, note that the area of triangle  $P_2OP_1$  is given by,

$$K_{\Delta P_2OP_1} = \frac{1}{2} W_2 W_1 = \frac{d}{k} \quad , \quad (\text{A.21})$$

where  $d$  is the distance between the sister surfaces and  $d = |R_2| = |R_1|$ . Equation (A.21) is a restatement of (A.19) such that,

$$d = \frac{kW_1W_2}{2} = R \quad . \quad (\text{A.22})$$

The radius of curvature at any point in the beam can be obtained graphically in a simple and straightforward manner.

The  $WW_0$  diagram can now be expanded for use in predicting the first-order properties of a Gaussian beam as it propagates through an imaging system. At a thin lens, the beam widths of both input and output beams must match and the wavefront curvature of one is converted to the wavefront curvature of the other. The lens can be depicted on the diagram as a point which "bends" the input skew line in accordance with the power  $\phi$  of the lens. The thin lens equation is

$$\frac{n'}{l'} = \frac{n}{l} + \phi, \quad (\text{A.23})$$

where  $l$  is the distance from the lens to the "object" plane,  $l'$  is the distance from the lens to the conjugate "image" plane and  $n$  and  $n'$  are the refractive indices in the object and image space respectively. As shown in Figure A.5, if  $l$  and  $l'$  are made equal to  $R$  and  $R'$  respectively, the conjugate (object and image) planes are sister surfaces to the input and output surfaces at the lens. Since these surfaces have an overall phase difference of  $180^\circ$ , one can conclude that a straight line through the origin identifies a pair of conjugate planes.

Translating this layout to the  $WW_0$  diagram in Figure A.6, we note that the lens is represented by a point and has a "clear aperture" equal to twice the beam width  $W_1 = W_1'$ . The lens "bends" the input skew line and in doing so generates an output beam with different characteristics (divergence angle, waist size and Rayleigh range) from the input beam. Since a straight line through the origin corresponds to conjugate planes, input and output beam waists can never be conjugate in a focal system. That is, with any kind of imaging system an input beam waist can never be imaged into the output beam waist. Because of this, if a beam waist resides in the front focal plane of a lens, the output beam waist will appear in the rear focal plane.

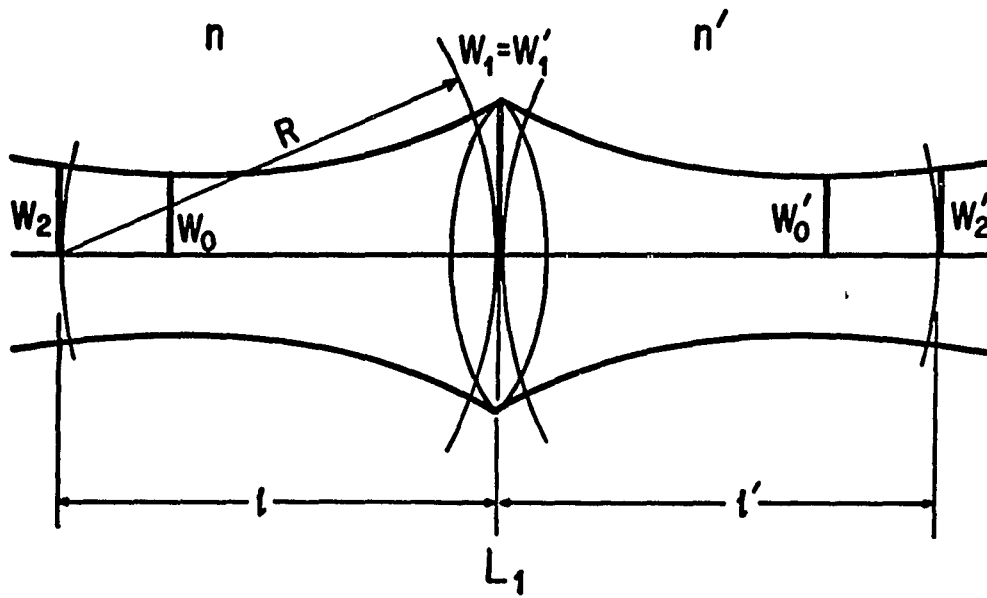


Figure A.5 Transformation of a Gaussian beam by a lens

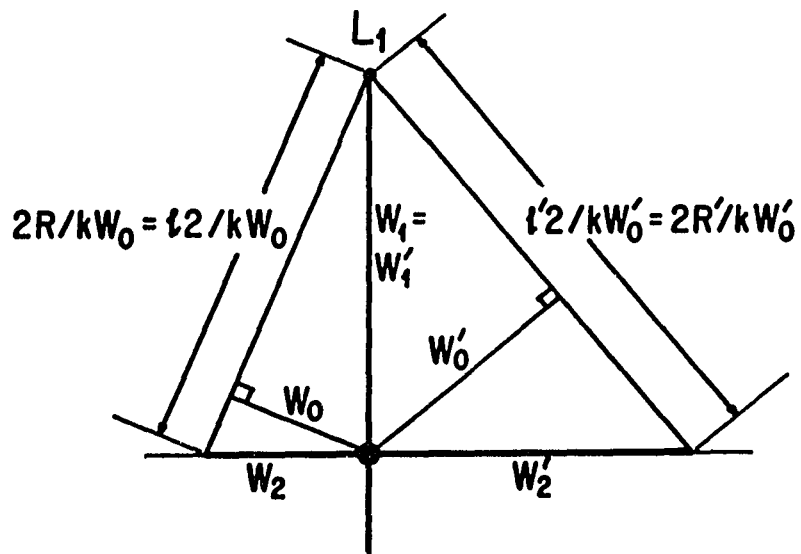


Figure A.6 Lens transformation on the  $WW_0$  diagram

The focal plane construction in the  $WW_0$  diagram is very similar to that of the  $\bar{Y}\bar{Y}$  diagram. Figure A.7 demonstrates that the line connecting the front focal plane  $F$  of lens  $L_1$  with the origin is parallel to the output beam. Likewise, the line connecting the rear focal plane  $F'$  to the origin is parallel to the input beam line. If the input and output beams have a  $90^\circ$  phase difference, their construction lines will be perpendicular. As a result, beam waists for both the input and output beams reside in the front and rear focal planes respectively.

Expanding on this insight, we can find the locus of all possible beam waists that lie in the front and rear focal planes of a lens. As shown in Figure A.8, we construct a circle of diameter equal to the beam width at the lens and containing the lens element. The left-hand semicircle is a trace of all possible beam waists in the front focal plane of the lens; the right-hand semicircle traces the output beam waists in the rear focal plane. The locus of all possible beam waists is a circle because the inscribed figure, a rectangle, always has the same diagonal that is equal to the circle diameter.

The circle construction also leads to a particularly simple method of predicting the size of a beam on a target. The axial distance from a lens element to a target is proportional to the triangular area swept out on the  $WW_0$  diagram. In Figure A.9, the lens element is  $L_1$  with a beam width  $W_1$  and the target lies somewhere on the line  $T$ . The distance from the lens to the target is given by

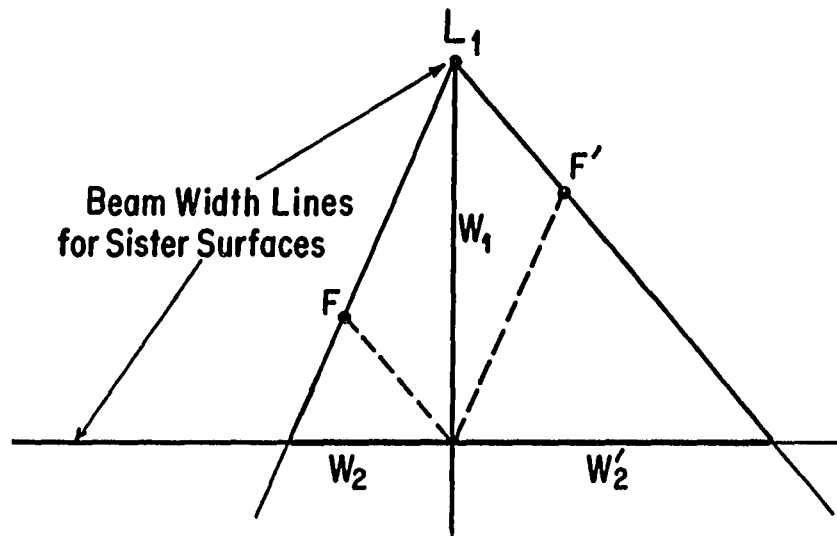


Figure A.7 Front and rear focal planes for a lens  $L_1$

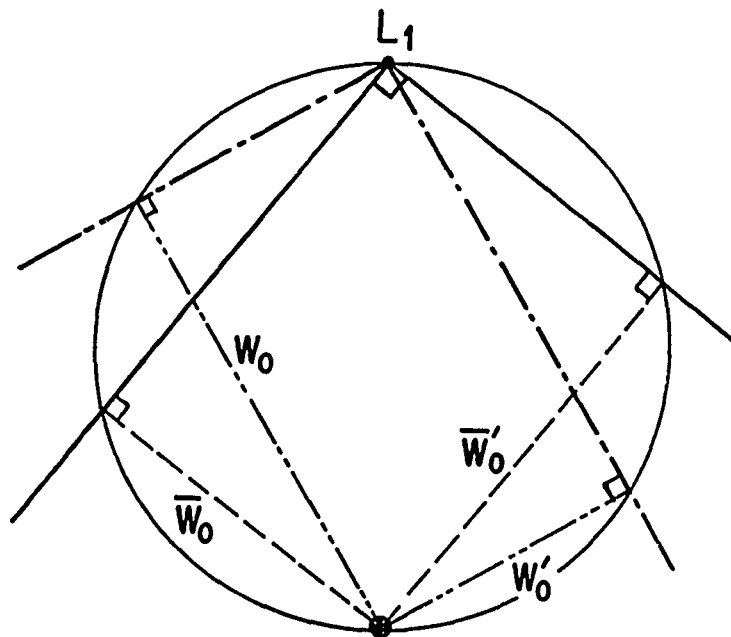


Figure A.8 Beam waist locus for lens  $L_1$



$$d = \frac{k\ell W_1}{2} . \quad (\text{A.24})$$

Since  $\ell$  and  $W_1$  are constant,  $d$  is constant, and we conclude that the locus of all points equidistant from the lens is a line parallel to the line connecting the lens to the origin and a distance  $\ell$  away. The target is designated as point  $T_1$ , and the beam size on the target is  $W_T$ . The line connecting  $L_1$  to  $T_1$  represents the output beam formed by the lens and hitting the target. This beam has a waist that occurs between the lens and the target. The distance locus is outside the beam waist locus and it becomes apparent that the beam waist will always be smaller than the smallest beam size on the target.

Figure A.10 demonstrates the situation for locating a beam waist on the target. The distance locus now falls inside the beam waist locus. Since the line  $T$  intersects the circle in two places, two beam waists, of different widths, can be located on the target. The smallest possible beam width on the target will be of width  $\ell$  when the target occurs at  $T_1$  in Figure A.10. The dashed line from  $L_1$  to  $T_1$  represents the output beam in this situation and  $W_0'$ , the beam waist. Note that even here the smallest beam size on the target will be larger than the corresponding beam waist. Finally, the maximum range at which a beam waist can occur on the target happens when the distance locus  $T$  is tangent to the circle. In this situation,  $\ell = W_1/2$  and,

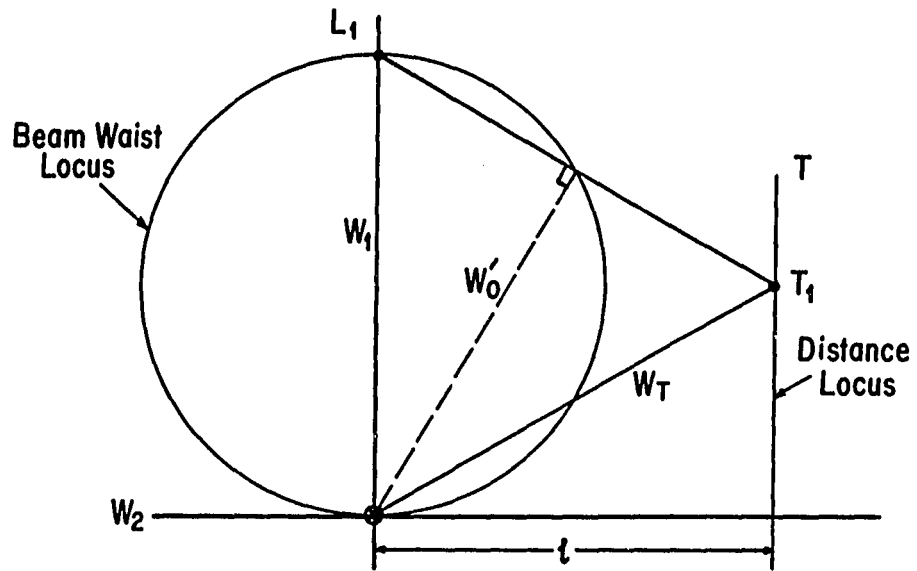


Figure A.9 Distance locus for target beyond beam waist

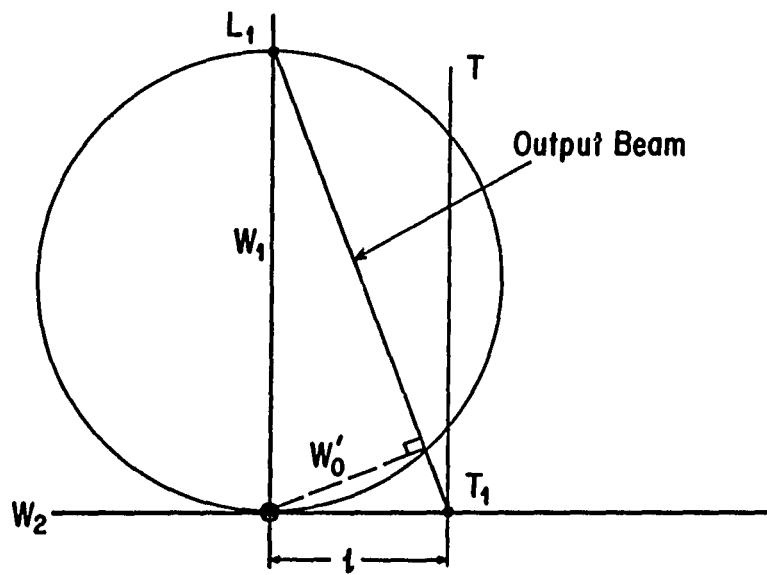


Figure A.10 Configuration for locating beam waist on target

$$d = \frac{k}{2} \cdot \frac{W_1}{2} \cdot W_1 = k \left( \frac{W_1}{2} \right)^2 \quad (\text{A.25})$$

Figure A.11 illustrates the procedure for producing a spot of a certain size in the output beam of a lens. This will appear as a circle around the origin with a radius equal to the desired beam width. A beam waist will occur where the spot size locus and beam waist locus intersect. On the other hand, an output beam size equal to the beam size at the lens shows up as the dashed circle around the origin in Figure A.11. This circle has a radius equal to the diameter of the beam waist locus. The maximum distance from the lens at which this beam size occurs is designated as R on the diagram and is given by

$$d = \frac{k}{2} (W_1)^2 \quad (\text{A.26})$$

This is the maximum range for which the beam can be kept as restricted as possible for as long as possible. For a target located anywhere on this dashed circle, the waist for the beam between the lens and the target will always occur halfway between the two.

The  $WW_0$  diagram has a major drawback in terms of scaling. It can be a very effective tool only when the beam or element size is on the order of the beam waist. This results from the value of the Lagrange invariant which determines the diameter to length scaling of an imaging system. In the case of Gaussian beams, the Lagrange invariant is

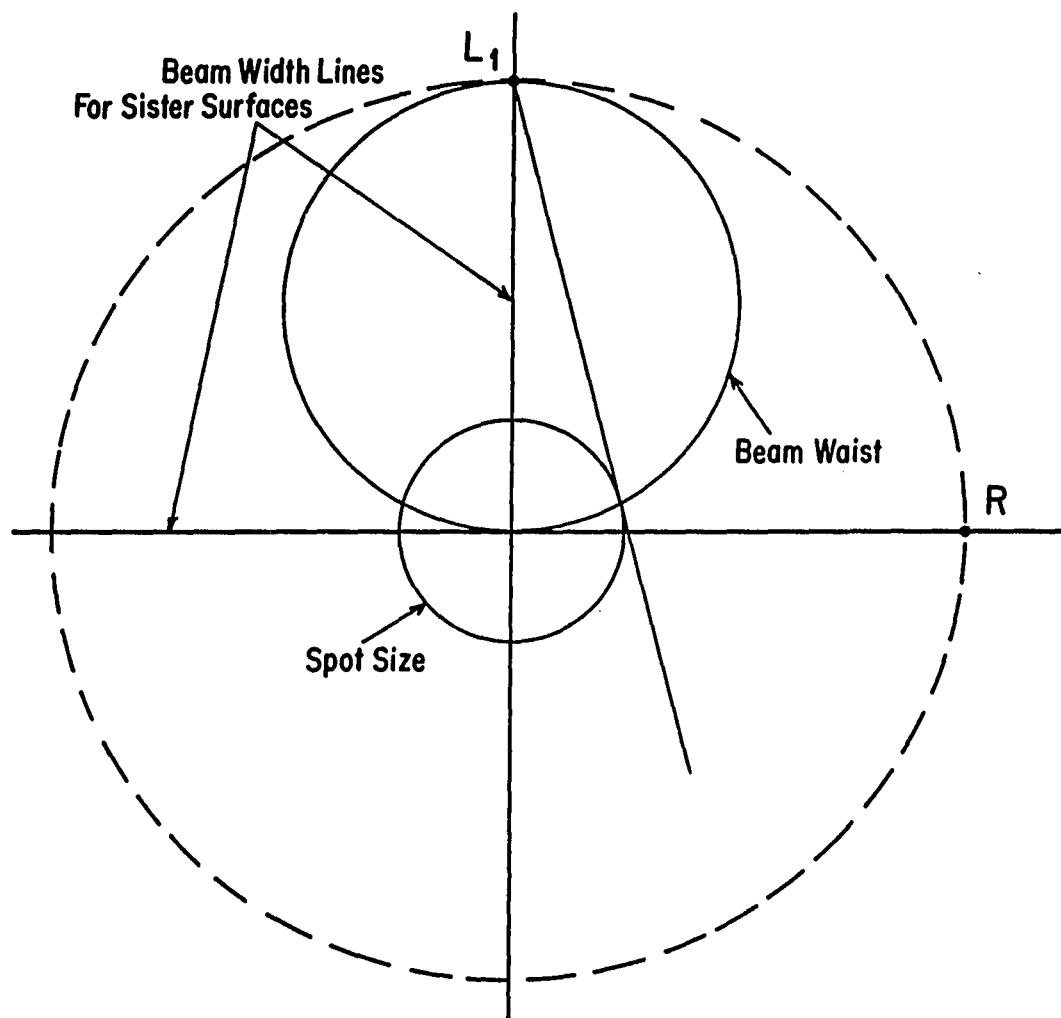


Figure A.11 Configuration for designated spot size in output beam of lens

$$\mathcal{H} = \frac{2}{k} = \frac{\lambda}{\pi} . \quad (\text{A.27})$$

Since we are not free to change the beam wavelength, neither can we change the scaling. Furthermore, the axes of the  $WW_0$  diagram are not separable so we cannot rescale at will. However, in some cases the beam waist is indistinguishable from the sister surface, permitting the use of orthogonal, and therefore rescaleable, axes. Even with this limitation, however, the  $WW_0$  diagram exhibits significant advantages in simplicity, clarity, and insight over other methods of Gaussian beam propagation.

APPENDIX B:  
PROOF THAT AN ARC OF AN OBLATE ELLIPSE IS  
PERPENDICULAR TO ITS ATTENDANT SKEW-LINE FAN

The parametric equations of equation (4.15) describe the arc PA in Figure 4.7 when  $a$  and  $\phi$  are held constant. These equations are therefore functions of  $\theta$  only and have the form

$$\begin{aligned}x(\theta) &= \frac{d \sin\theta}{\cos a} \cos(\phi + a) . \\y(\theta) &= d \frac{\sin\theta}{\cos a} \sin(\phi + a) , \\z(\theta) &= d \tan a \cos\theta .\end{aligned}\tag{B.1}$$

where we have made the substitutions for  $\sinh\mu$  and  $\cosh\mu$  given in Equation (4.14).

The angle  $\phi$  is measured from the x axis to the line segment ON in the plane of the waist only. The azimuthal angle in any other plane perpendicular to the z axis is given by  $\phi \pm a$ . The angle  $a$  determines the ellipse of interest and the length of the skew line. The tangent to the elliptical arc is given by

$$\vec{t} = \frac{\partial \vec{r}}{\partial \theta} = \frac{\partial x(\theta)}{\partial \theta} \hat{x} + \frac{\partial y(\theta)}{\partial \theta} \hat{y} + \frac{\partial z(\theta)}{\partial \theta} \hat{z} ,\tag{B.2}$$

where  $\hat{x}$ ,  $\hat{y}$ , and  $\hat{z}$  are Cartesian unit vectors. Differentiating with respect to  $\theta$  yields

$$\vec{t} = \frac{d \cos\theta \cos(\phi \pm a)}{\cos a} \hat{x} + \frac{d \cos\theta \sin(\phi \pm a)}{\cos a} \hat{y} - d \tan a \sin\theta \hat{z} .\tag{B.3}$$

The + or - sign for  $a$  determines whether the arc in question is for a counterclockwise (+ $a$ ) or clockwise (- $a$ ) twist.

The vector representation for a skew line  $\vec{s}$ , is given by

$$\vec{s} = (x' - x) \hat{x} + (y' - y) \hat{y} + (z' - z) \hat{z} . \quad (\text{B.4})$$

In the plane of the waist,  $a = 0$  and the parametric Equation of (B.1) become

$$\begin{aligned} x &= d \sin \theta \cos \phi , \\ y &= d \sin \theta \sin \phi , \\ z &= 0 . \end{aligned} \quad (\text{B.5})$$

On the elliptical arc, the coordinates of a point are  $x'$ ,  $y'$ , and  $z'$  and is described by the parametric equations

$$\begin{aligned} x'(\theta) &= \frac{d \sin \theta \cos(\phi \pm a)}{\cos a} \\ y'(\theta) &= \frac{d \sin \theta \sin(\phi \pm a)}{\cos a} , \\ z'(\theta) &= d \tan a \cos \theta . \end{aligned} \quad (\text{B.6})$$

Substituting Equations (B.5) and (B.6) into (B.4) gives

$$\vec{s} = \mp d \tan a \sin \theta \sin \phi \hat{x} \pm d \tan a \sin \theta \cos \phi \hat{y} + d \tan a \cos \theta \hat{z} . \quad (\text{B.7})$$



The magnitude of  $\vec{s}$  is  $d \tan \alpha$ , the length of the skew line. Note that regardless of the choice of  $\alpha$ , the magnitude of the  $\hat{x}$  and  $\hat{y}$  components are opposite in sign.

Taking the dot product of  $\vec{t}$  and  $\vec{s}$  leads to

$$\begin{aligned} \vec{t} \cdot \vec{s} &= \mp d^2 \tan \alpha \frac{\sin \theta \cos \theta}{\cos \alpha} \cos(\theta \pm \alpha) \sin \theta \\ &\quad \pm d^2 \tan \alpha \frac{\sin \theta \cos \theta}{\cos \alpha} \sin(\theta \pm \alpha) \cos \theta \\ &\quad - d^2 \tan^2 \alpha \sin \theta \cos \theta . \end{aligned} \quad (\text{B.8})$$

Simplifying terms results in

$$\vec{t} \cdot \vec{s} = d^2 \frac{\tan \alpha}{\cos \alpha} \sin \theta \cos \theta [\mp \sin \theta \cos(\theta \pm \alpha) \pm \cos \theta \sin(\theta \pm \alpha) - \sin \alpha] . \quad (\text{B.9})$$

Next, we expand the term in brackets to find that

$$\begin{aligned} \vec{t} \cdot \vec{s} &= d^2 \frac{\tan \alpha}{\cos \alpha} \sin \theta \cos \theta [\mp \sin \theta (\cos \theta \cos \alpha \mp \sin \theta \sin \alpha) \\ &\quad \pm \cos \theta (\sin \theta \cos \alpha \pm \cos \theta \sin \alpha) - \sin \alpha] \quad (\text{B.10}) \\ &= \frac{d^2 \tan \alpha}{\cos \alpha} \sin \theta \cos \theta [\sin \alpha (\mp \sin^2 \theta \mp \cos^2 \theta) - \sin \alpha] , \end{aligned}$$

and finally

$$\vec{t} \cdot \vec{s} = 0 . \quad (\text{B.11})$$

Therefore, the skew line is perpendicular to the elliptical arc PA.

**APPENDIX C:**

**REFLECTION OF A SKEW-LINE RAY FROM AN ELLIPTICAL MIRROR**

In geometrical optics, the laws for reflection are:

1. The reflected ray lies in the plane formed by the incident ray and the surface normal.
2. The reflected ray forms the equal but opposite angle to the normal as the incident ray.

Since the reflective surface is an oblate ellipse, its normal is simply the unit normal for the oblate ellipsoid  $\hat{\xi}$  in the oblate spheroidal coordinate system. The three unit normals for this system are given by

$$\begin{aligned}\hat{\xi} &= \frac{\xi\sqrt{1-\eta^2}}{\sqrt{\xi^2+\eta^2}} \cos(\phi \pm \alpha) \hat{x} + \frac{\xi\sqrt{1-\eta^2}}{\sqrt{\xi^2+\eta^2}} \sin(\phi \pm \alpha) \hat{y} + \eta \frac{\sqrt{\xi^2+1}}{\sqrt{\xi^2+\eta^2}} \hat{z} , \\ \hat{\eta} &= -\eta \frac{\sqrt{\xi^2+1}}{\sqrt{\xi^2+\eta^2}} \cos(\phi \pm \alpha) \hat{x} - \eta \frac{\sqrt{\xi^2+1}}{\sqrt{\xi^2+\eta^2}} \sin(\phi \pm \alpha) \hat{y} + \xi \frac{\sqrt{1-\eta^2}}{\sqrt{\xi^2+\eta^2}} \hat{z} , \quad (\text{C.1}) \\ \hat{\phi} &= -\sin(\phi \pm \alpha) \hat{x} + \cos(\phi \pm \alpha) \hat{y} ,\end{aligned}$$

where  $\hat{x}$ ,  $\hat{y}$ , and  $\hat{z}$  are the unit vectors in the Cartesian coordinate system. We have made use here of the definition of the azimuthal angle  $\phi \pm \alpha$  in a plane other than the plane of the waist. Furthermore, we shall employ the definitions for  $\tan\alpha$  and  $\cos\alpha$  given in Eq. (4.14).

Since the incident ray has the same trajectory as a skew line, its vector representation can be found from Eq. (B.7). For a skew-line trajectory with a clockwise twist,  $(-\alpha)$ , the incident ray is the unit vector

$$\hat{u} = \sin\theta \sin\phi \hat{x} - \sin\theta \cos\phi \hat{y} + \cos\theta \hat{z} . \quad (\text{C.2})$$

In order to find the plane containing  $\hat{\xi}$  and  $\hat{u}$ , it suffices to find the normal to both  $\hat{\xi}$  and  $\hat{u}$ , since this will also be the normal to the plane containing the two vectors. The unit normal can be found by taking the cross-product of  $\hat{\xi}$  for a

negative  $\alpha$  and  $\hat{u}$  and is given by

$$\frac{\hat{u} \times \hat{\xi}}{|\hat{u} \times \hat{\xi}|} = - \frac{\eta \cos(\phi - \alpha) \sqrt{\xi^2 + 1}}{\sqrt{\xi^2 + \eta^2}} \hat{x} - \frac{\eta \sin(\phi - \alpha) \sqrt{\xi^2 + 1}}{\sqrt{\xi^2 + \eta^2}} \hat{y} + \frac{\xi \sqrt{1 - \eta^2}}{\sqrt{\xi^2 + \eta^2}} \hat{z}. \quad (C.3)$$

This is simply  $\hat{\eta}$ .

Next, we calculate the angle between the incident ray and the surface normal which is

$$\cos^{-1}[\hat{u} \cdot \hat{\xi}] = \cos^{-1} \left[ \frac{\sin^2 \alpha \sin^2 \theta + \cos^2 \theta}{\cos \alpha \sqrt{\tan^2 \alpha + \cos^2 \theta}} \right]. \quad (C.4)$$

Reduction of the term on the right leads to

$$\cos^{-1}[\hat{u} \cdot \hat{\xi}] = \cos^{-1} \left[ \sqrt{\sin^2 \alpha + \cos^2 \theta} \cos^2 \alpha \right]. \quad (C.5)$$

In determining the reflected ray,  $\hat{v}$ , we know that it must lie in the plane formed by  $\hat{u}$  and  $\hat{\xi}$  and must therefore be perpendicular to  $\hat{\eta}$ . Also, the angle  $\hat{v}$  makes with  $\hat{\xi}$  must be equal and opposite to  $\cos^{-1}(\hat{u} \cdot \hat{\xi})$ . Expressed in terms of a dot product, this latter condition is

$$\hat{v} \cdot \hat{\xi} = - \hat{u} \cdot \hat{\xi}. \quad (C.6)$$

Finally, the angle formed by  $\hat{u}$  and  $\hat{v}$  must be twice that of  $\hat{u} \cdot \hat{\xi}$ . This condition can also be expressed in terms of a dot product as

$$\hat{u} \cdot \hat{v} = 1 - 2[\hat{u} \cdot \hat{\xi}]^2. \quad (C.7)$$

Taken together, these conditions will produce the three direction cosines of the vector  $\hat{v}$ .

If the vector  $\hat{v}$  is given by

$$\hat{v} = a\hat{x} + b\hat{y} + c\hat{z}, \quad (C.8)$$

then the three simultaneous equations to find  $a$ ,  $b$ , and  $c$  are

$$-\cos \theta \cos(\phi - \alpha)a - \cos \theta \sin(\phi - \alpha)b + \sin \alpha \sin \theta c = 0,$$

$$\sin\alpha \sin\theta \cos(\phi - \alpha)a + \sin\alpha \sin\theta \sin(\phi - \alpha)b + \cos\theta c = -\sin^2\alpha - \cos^2\theta \cos^2\alpha, \quad (\text{C.9})$$

$$\sin\theta \sin\phi a - \sin\theta \cos\phi b + \cos\theta c = 1 - 2(\sin^2\alpha + \cos^2\theta \cos^2\alpha).$$

The resultant reflected ray is given by

$$\hat{v} = \sin\theta \sin(\phi - 2\alpha)\hat{x} - \sin\theta \cos(\phi - 2\alpha)\hat{y} - \cos\theta\hat{z}, \quad (\text{C.10})$$

which is a skew-line vector on the same hyperbolic envelope as  $\hat{u}$  pointed in the direction of the waist with endpoint at the point of reflection.

## REFERENCES

- Abramowitz, Milton and Stegun, Irene, Editors, Handbook of Mathematical Functions, National Bureau of Standards, Washington, D.C. (1964).
- Agrawal, G. P., and M. Lax, "Free-Space Wave Propagation beyond the Paraxial Approximation", *Phys. Rev A* 27, 1639 (1983).
- Agrawal, G. P., and D. N. Pattanayak, "Gaussian Beam Propagation beyond the Paraxial Approximation", *JOSA* 69, 575 (1979).
- Arfken, George, Mathematical Methods for Physicists, Academic Press, Orlando (1985).
- Arnaud, Jacques A., "Representation of Gaussian Beams by Complex Rays", *Applied Optics* 24, 538 (1985).
- Arnaud, Jacques A., "Hamiltonian Theory of Beam Mode Propagation", Progress in Optics, 11, E. Wolf editor, North Holland (1973).
- Arnaud, Jacques A., "Degenerate Optical Cavities II: Effect of Misalignments", *Applied Optics* 8, 1909 (1969).
- Arnaud, Jacques A., "Degenerate Optical Cavities III: Effect of Aberrations", *Applied Optics* 9, 1192 (1970).
- Arnaud, Jacques A., and Herwig Kogelnik, "Gaussian Light Beams with General Astigmatism", *Applied Optics* 8, 1987 (1969)
- Bateman, Harry, The Mathematical Analysis of Electrical and Optical Wave-Motion, Dove Publications (1955), first edition, Cambridge University Press, p. 92 (1914).
- Bell, Robert J. T., Elementary Treatise on Coordinate Geometry of Three Dimensions, Macmillan and Company, London pp. 312-321 (1912).
- Bell, W. W., Special Functions for Scientists and Engineers, van Nostrand Company, Ltd., London (1968).
- Boyd, C. D. and J. P. Gordon, "Confocal Multimode Resonator for Millimeter Through Optical Wavelength Masers", *Bell System Technical Journal* 40, 489 (1961).

- Boyd, G. D. and Herwig Kogelnik, "Generalized Confocal Resonator Theory", Bell System Technical Journal 41, 1347 (1962).
- Bracewell, Ron, The Fourier Transform and its Applications, McGraw-Hill, pp. 227-233 (1962).
- Bykov, V. P., and L. A. Vainshtein, "Geometric Optics of Open Resonators", Sov. Phys. JETP 20, 338 (1965).
- Choudhary, Satinath and Leopold B. Felsen, "Analysis of Gaussian Beam Propagation and Diffraction by Inhomogeneous Wave Tracking", Proc. IEEE 62, 1530 (1974).
- Choudhary, Satinath and Leopold B. Felsen, "Asymptotic Theory for Inhomogeneous Waves", IEEE Trans. on Antennas and Propagation, AP-2, 730 (1973).
- Collins, Stuart A., "Analysis of Optical Resonators Involving Focusing Elements", Applied Optics 3, 1263 (1964).
- Couture, Marc and Pierre A. Belanger, "From Gaussian Beam to Complex-Source-Point Spherical Wave", Physical Review A 24, 325 (1981).
- Delano, E., "First-Order Design and YY Diagram", Applied Optics 2, 1251 (1963).
- Deschamps, G. A., "Gaussian Beam as a Bundle of Complex Rays", Electronics Letters 7, 684 (1971).
- Einzig, P. D. and L. B. Felsen. "Evanescent Waves and Complex Rays", IEEE Trans. on Antennas and Propagation AP-30, 594 (1982).
- Einzig, P. D. and S. Raz, "Wave Solutions under Complex Space-time Shifts", JOSA A 4, 3 (1987).
- Felsen, Leopold B., "Geometrical Theory of Diffraction, Evanescent Waves, Complex Rays and Gaussian Beams", Geophys. J. R. Astr. Soc. 79, 77 (1984).
- Felsen, Leopold B., "Evanescent Waves", J. Opt. Soc. Am. 66, 751 (1976).
- Felsen, Leopold B. and Sang-Yung Shin, "Rays, Beams, and Modes Pertaining to the Excitation of Dielectric Waveguides", IEEE Trans. on Microwave Theory and Techniques, MTT-23, 150 (1975).
- Flammer, Carson, Spheroidal Wave Functions, Stanford University Press, Palo Alto, CA (1957).



- Fox, A. G. and Tingye Li, "Resonant Modes in a Maser Interferometer", Bell System Technical Journal 46, 453 (1961).
- Gaskill, Jack D., Linear Systems, Fourier Transforms and Optics, John Wiley and Sons, New York (1978).
- Goodman, Joseph W., Introduction to Fourier Optics, McGraw-Hill, San Francisco (1968).
- Gordon, J., "Elements of Laser Theory" in Laser Technology and Applications, S. Marshall, Editor; McGraw-Hill, NY (1968).
- Harris, Franklin S., Jr., Michael S. Tavenner, and Richard L. Mitchell, "Single Slit Fresnel Diffraction Patterns", J. Opt. Soc. Am. 59 (3) p. 293 (1969).
- Hashimoto, M., "Beam Waves with Sources at Complex Location", Electronics Letters 21, 1096 (1985).
- Herloski, R., S. Marshall, and R. Antos, "Gaussian Beam Ray-Equivalent Modeling and Optical Design", Appl. Optics 22, 1168 (1983).
- Hopkins, H. H., "The Wave Aberration Associated with Skew Rays", Proceedings of the Physical Society B 65, 934 (1952).
- Ito, Michiaki, "Theory of Ellipsoidal Waves and Seidel Aberrations of Gaussian Beams", Japanese Journal of Applied Physics 12, 856 (1973).
- Ito, Michiaki, "Common Aspects of Gaussian Beams, Ellipsoidal Waves and Multipole Radiation", Japanese Journal of Applied Physics 12, 1914 (1973).
- Kahan, W. K., "Geometric Optical Derivation of Formula in a Spherical Mirror Resonator", Appl. Opt. 4, 758 (1965).
- Kogelnik, Herwig and Tingye Li, "Laser Beams and Resonators", Proceedings of the IEEE 54, 1312 (1966).
- Kogelnik, Herwig, "On the Propagation of Gaussian Beams of Light Through Lenslike Media Including Those with a Loss or Gain Variation", Appl. Opt. 4, 1962 (1965).
- Kogelnik, Herwig W., "Imaging of Optical Modes in Resonators with Internal Lenses", Bell System Technical Journal 44, 455 (1965).
- Keller, J. B. and W. Streifer, "Complex Rays with an Application to Gaussian Beams", J. Opt. Soc. Am. 61, 40 (1971).

- Kessler, David, and R. V. Shack, "First-Order Design of Laser Systems with the  $YY$  Diagram", *J. Opt. Soc. Am. A* 1, 1219 (1984).
- Kuper, Thomas G., "A New Perspective on Electromagnetic Diffraction Theory", Ph.D. Dissertation, The University of Arizona, Optical Sciences Center (1983).
- Li, Tingye, "Dual Forms of the Gaussian Beam Chart", *Applied Optics* 3, 1315 (1964).
- Luk, Kwai-Man and Yu, Ping-Kong, "Generation of Hermite-Gaussian Beam Modes by Multipoles with Complex Source Points", *J. Opt. Soc. Am* 2, 1818 (1985).
- Morse, Philip and Feshbach, Herman, Methods of Theoretical Physics, McGraw-Hill (1953).
- Mukunda, N., Simon R., and Sudarshan, E.C.G., "Fourier Optics for the Maxwell Field: Formalism and Applications", *J. Opt. Soc. Am* 2, 416 (1985).
- Norris, A. N., "Complex-Point-Source Representation of Real Point Sources and the Gaussian Beam Summation Method", *J. Opt. Soc. Am. A* 3, 2005 (1986).
- Oberhettinger, Fritz, Professor Emeritus, Oregon State University, Department of Mathematics, Private Communication (1984).
- Oberhettinger, Fritz, Tables of Bessel Transforms, Springer-Verlag, Berlin (1972).
- Otis, Gabriel, "Application of the Boundary-Diffraction-Wave Theory to Gaussian Beams", *J. Opt. Soc. Am* 64, 1545 (1974).
- Papoulis, Athanasios, The Fourier Integral and Its Applications, McGraw-Hill, pp. 227-233 (1962).
- Shack, Roland V., Professor of Optical Sciences, The University of Arizona, Class Notes (1983).
- Shack, Roland V., The Optical Sciences Center, The University of Arizona, seminar, October 1982.
- Shin, Sang-Yung and Felsen, L. B., "Gaussian Beam Modes by Multipoles with Complex Source Points", *J. Opt. Soc. Am.* 67, 699 (1977).
- Shin, Sang-Yung and Felsen, L. B., "Gaussian Beams in Anisotropic Media", *Applied Phys* 5, 239 (1974).

- Siegman, Anthony E., "Hermite-Gaussian Functions of Complex Argument as Optical-Beam Eigenfunctions", *J. Opt. Soc. Am* 63, 1093 (1973).
- Siegman, Anthony E., An Introduction to Lasers and Masers, McGraw-Hill, New York (1971).
- Slepian, D. and Pollack, H. O., "Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty - I", *Bell System Technical Journal*, 40, 43 (1961).
- Smith, P. F., Gale, A. S. and Neelley, J. H., New Analytic Geometry, Ginn and Company, Boston, p. 304-305 (1928).
- Sommerville, D. M. Y., Analytic Geometry of Three Dimensions, University Press, Cambridge, pp. 114-117 (1951).
- Stavroudis, Orestes N., The Optics of Rays, Wavefronts and Caustics, Academic Press, New York (1972).
- Steier, W. J., "The Ray Packet Equivalent of a Gaussian Light Beam", *Appl. Opt.* 5, 1229 (1966).
- Stratton, Julius A., Electromagnetic Theory, McGraw Hill, New York, (1941)
- Struik, Dirk J., Lectures on Classical Differential Geometry, Second Edition, Addison-Wesley Publishing Co., Reading, MA (1961).
- Takenaka, Takashi, Yokota, Mitsuhiro, and Fukumitsu, Otozo, "Propagation of Light Beams Beyond the Paraxial Approximation", *J. Opt. Soc. Am A* 2, 826 (1985).
- Vainshtein, L. A. "Open Resonators with Spherical Mirrors", *Soviet Physics JETP* 18, 471 (1964).
- Wang, Wei-Yi D., and Deschamps, Georges, "Application of Complex Ray Tracing to Scattering Problems", *Proc. IEEE* 62, 1541 (1974).
- Watson, G. N., A Treatise on the Theory of Bessel Functions, Cambridge University Press (1922) reprinted (1980).
- Zauderer, Erich, "Complex Argument Hermite-Gaussian and Laquerre-Gaussian Beams", *J. Opt. Soc. Am. A* 3, 465 (1986).