Urs Vokinger

1501

Band 387

Propagation, Modification and Analysis of Partially Coherent Light Fields

UFO Atelier für Gestaltung & Verlag

Université de Neuchâtel
Institut de Microtechnique

Propagation, Modification and Analysis of Partially Coherent Light Fields

Thèse

Présentée à la Faculté des sciences
pour obtenir le grade docteur ès sciences

par

Urs Vokinger

UFO Dissertation Band 387

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Ein Titelsatz für diese Publikation ist bei Der Deutschen Bibliothek erhältlich.

Dissertation der Universität Neuchâtel
Datum der mündlichen Prüfung: 8. April 1999
Referenten: Prof. Dr. R. Dändliker
Dr. H.-P. Herzig
Dr. B. Gächter
Dr. B. Braunecker
Prof. Dr. F. Gori

UFO Atelier für Gestaltung & Verlag GbR
Allensbach
Internet http://www.ufo-verlag.de
Maus Druck & Medien GmbH, Konstanz
Erste Auflage 2000

ISBN 3-930803-87-9

IMPRIMATUR POUR LA THÈSE

Propagation, modification and analysis of partially coherent light fields

de M. Urs Vokinger

UNIVERSITÉ DE NEUCHÂTEL

FACULTÉ DES SCIENCES

La Faculté des sciences de l'Université de Neuchâtel sur le rapport des membres du jury,

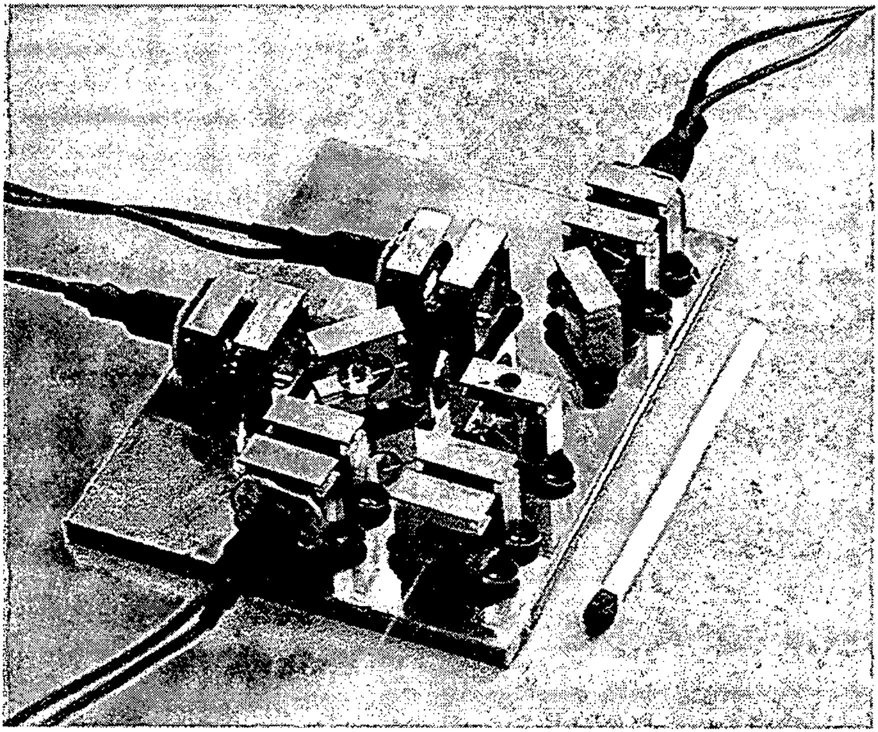
MM. R. Dändliker (directeur de thèse), H.-P. Herzig, B. Gächter (Heerbrugg), B. Braunecker (Heerbrugg) et F. Gori (Rome)

autorise l'impression de la présente thèse.

Neuchâtel, le 22 avril 1999

Le doyen:

F. Stoeckli



Beam combiner consisting of four pulsed laser diodes whose field is coupled into a fiber.

Summary

The goal of this work is to describe the design of an optical pointing system for rangefinders measuring the distance of non-cooperative target. The presented description yields efficiently possible design solutions if small size of the system and maximum irradiation of the target are required. The usability of novel assembling technology is tested by the realisation of an optical pointing system.

Commonly, when optical systems are designed, the paraxial geometrical optics is used to carry out the first rough design and then ray tracing to optimize the system. However, for pointing systems where light throughput, concepts of source combination and beam shaping are of major concerns, the classical geometrical approach of optics will not provide the necessary information to optimize a design. Therefore, we present in this work an alternative approach. It is based on the concept of wave propagation, phase space volume and entropy, and it yields more information about possible strategy in the optimization and realization of optical systems in general.

The realized system consists of four laser diodes whose emitted field is combined and coupled into a multi-mode fiber. The radiated beam at the output of the fiber can then be collimated by a lens. In the ideal case about 40% efficiency can be expected at the output of the fiber and by modifying the intensity distribution in the source plane (matching the phase space volume of the beam to the fiber) an efficiency of 100% is theoretically possible. In order to reduce production costs, the realized pointing system was assembled with low cost optical elements and, therefore, did not achieve the predicted performance. However, it shows the feasibility of a such a system and the correctness of the proposed approach.

For further development of the pointing system, it should be reconsidered to replace the four laser diodes by a single laser diode with a four times longer emitting area (i.e. a four times longer laser chip), which is then optically folded by micro-optical elements in order to match the phase space volume of the fiber. The resulting pointing system would be more compact, cheaper in production, and therefore more competitive on the market.

The work has also revealed that the automated assembling of optical (micro-) systems is still a problem and needs further creative engineering work.

Contents

Chapter 1

Introduction

The goal of this work is to describe the design of an optical pointing system for rangefinders measuring the distance of non-cooperative target. The presented description yields efficiently possible design solutions if small size of the system and maximum irradiation of the target are required. The usability of novel assembling technology (OSMD [1.1]) is tested by the realisation of an optical pointing system.

The best known application of optical pointing systems are laser pointers, pencil like handy light sources, commonly used in presentations and talks as a pointing aid. Pointing systems are also used for communication link between satellites and in rangefinders. Rangefinders are optical systems measuring the distance to an object or target. As for laser pointers, its optical pointing system emits a beam whose energy flow is confined in a small solid angle (collimated beam). However, there are significant differences: the emitted energy flow is much higher, because the object to which it shall be pointed is further away (kilometers). When one source is not able to achieve the necessary energy flow, the field of several sources is combined. Furthermore, the operating wavelength can be in the infrared (pointing to the object is then carried out by a cross hair in the binocular). This is not only because high power laser diodes are easier available in the infrared, but eye-safety standards [1.2] allow higher energy flow in the infrared beam, at 1550 nm about a factor 10^3 compared to visible light if the sources are pulsed.

Usually in industry, optical systems are designed by ray tracing, a method to evaluate the propagation of light in an optical system by geometrical optics. Ray tracing has been proven to be a versatile design tool, although it neglects the wave properties of light. The approach is meaningful as long as the source is incoherent, so that the effects of diffraction are smeared out by the superposition of many different diffraction patterns and as long as collimating or focusing are not critically limited by diffraction. However, if the source is coherent (laser sources) and collimation or focusing becomes critical, diffraction has to be included and ray tracing is not sufficient to describe the light propagation. For instance, ray tracing will neither reveal that light can never be focused to a geometrical point [1.3], nor reflect the asymmetric intensity distribution along the optical axis with respect to the focus of a uniform converging spherical wave diffracted at a circular aperture [1.4].

This thesis presents a design concept based on other ideas than ray tracing. It is based on the concept of wave propagation, phase space volume and entropy, and it yields more information about possible strategy in the optimization and realization of optical systems in general. The concept is valid at any degree of coherence of the source. It can be considered as a synthesis of known theories, but to the best of our knowledge has not yet been applied rigorously to the design of optical pointing systems or other optical systems.

Chapter 2 of this thesis reviews the properties and propagation laws of the deterministic scalar field. Chapter 3 presents the random field which is encountered in incoherent sources. It will be shown how such sources can be decomposed into deterministic fields, so that problems of propagation can be solved by the propagation laws of deterministic fields. In the following chapter the concept of entropy for electromagnetic waves is presented. In chapter 5 the phase space volume of a field is introduced. It will be shown how and under which conditions the phase space volume of a radiated field can be transformed or even reduced. Since combining the fields of different sources is important in order to increase the energy flow at the output of a pointing system, a chapter is dedicated to this problem. In this chapter the concept of entropy is used to find the possibilities of combining fields by using passive optical elements. Chapter 7 shows how to calculate numerically the propagation of deterministic scalar fields in the paraxial approximation by respecting the sampling theorem. In chapter 8 the design, realization and characterization of a prototype pointing system is carried out. The layout is made suitable to the requirements of the Optical Surface Mounted Device technique (OSMD). The conclusions of this thesis are presented in Chap. 9.

References

[1.1] W. Andreasch, *Konzeption und Entwicklung einer Technologie zur automatisierten Oberflächenmontage optischer Elemente (Optical SMD)* (Swiss Federal Institute of Technology (EPFL), thesis No. 1591, 1996).

[1.2] International Standard, *Safety of laser products* (CEI/IEC 825-1:1993, Geneva, Switzerland).

[1.3] J. W. Goodman, *Introduction to Fourier optics* (McGraw-Hill Series in Electrical and Computer Engineering, New York, 1996, 2nd edition), pp. 96-125.

[1.4] J. H. Erkkila and M. E. Roger, "Diffracted fields in the focal volume of a converging wave, " J. Opt. Soc. Am. A **71**, 904-905 (1981).

CHAPTER 2

Propagation of deterministic scalar wave fields

2.1 Introduction

The aim of this chapter is to describe the propagation of deterministic electromagnetic wave fields as scalar fields. We shall start with the electromagnetic vector field and then introduce the scalar field. Next, the propagation of a monochromatic and deterministic scalar wave is formulated by different diffraction integrals. The Fourier transform is introduced, which enables to express the diffraction integrals by Fourier transforms and multiplication. Finally, we derive the equations for a scalar wave propagating in a lens like optical system. We will find that within the paraxial approximation, the intensity distribution between two conjugate planes of the lens is described by the imaging law of geometrical optics.

2.2 Wave equation and scalar field

An electromagnetic field is characterized by the electric field vector $\mathcal{E}$ and by the magnetic field vector $\mathcal{H}$. In a source free homogenous medium with relative dielectric constant ε_r and relative magnetic permeability μ_r, the vectors satisfy the homogenous wave equations [2.1]

$$\left(\nabla^2 - \frac{\varepsilon_r \mu_r}{c^2}\frac{\partial^2}{\partial t^2}\right)\cdot \mathcal{E}(\boldsymbol{r},t) = 0 \qquad \left(\nabla^2 - \frac{\varepsilon_r \mu_r}{c^2}\frac{\partial^2}{\partial t^2}\right)\cdot \mathcal{H}(\boldsymbol{r},t) = 0 . \tag{2.2-1}$$

In this equations ∇ is the Nabla operator, c the velocity of light in vacuum, $\boldsymbol{r}$ the vector of location with the components x, y, z and t is the time. In the entire chapter it is assumed that $\mathcal{E}(\boldsymbol{r},t)$ and $\mathcal{H}(\boldsymbol{r},t)$ depend on t deterministically. In this case Eq. (2.2-1) describes a propagating wave consisting of a deterministic electromagnetic vector field.

The flow of the electromagnetic power is governed by the Poynting vector

$$\mathcal{S}(\boldsymbol{r},t) = \mathcal{E}(\boldsymbol{r},t) \times \mathcal{H}(\boldsymbol{r},t), \tag{2.2-2}$$

which is the cross product of the electric and magnetic field vectors [2.1].

The intensity I, which is the energy flow across a unit area, is the time averaged Poynting vector projected onto the normal $\boldsymbol{n}$ of the unit area. Hence,

$$I(\boldsymbol{r}) = \boldsymbol{n}\cdot\langle \mathcal{S}(\boldsymbol{r},t)\rangle. \tag{2.2-3}$$

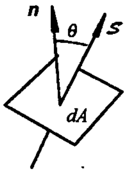


Figure 2.2-1: Energy flow through a unit area dA.

It is possible to summarize the behavior of the all components of $\mathcal{E}$ and $\mathcal{H}$ in Eq. (2.2-1) through a single scalar wave equation

$$\left(\nabla^2 - \frac{\varepsilon_r\mu_r}{c^2}\frac{\partial^2}{\partial t^2}\right)\cdot \mathcal{V}(\boldsymbol{r},t) = 0, \tag{2.2-4}$$

where $\mathcal{V}(r,t)$ represents any components of the vector field. The complex function $\mathcal{V}(r,t)$ in Eq. (2.2-4) describes a propagating scalar wave [2.1]. The intensity I introduced by Eq. (2.2-3) is then given by

$$I(\boldsymbol{r},\theta) = \langle \mathcal{V}^*(\boldsymbol{r},t)\cdot \mathcal{V}(\boldsymbol{r},t)\rangle\cdot\cos(\theta) = \langle |\mathcal{V}(\boldsymbol{r},t)|^2\rangle\cdot\cos(\theta), \tag{2.2-5}$$

where θ is the angle of incidence. The asterisk * stands for conjugate complex. The scalar complex wave $\mathcal{V}$ can be considered as a linear superposition of scalar waves oscillating at different frequencies ν which is [2.1]

$$\mathcal{V}(\boldsymbol{r},t) = \int_0^\infty U(\boldsymbol{r},\nu)\cdot\exp(-i2\pi\nu t)\cdot d\nu. \tag{2.2-6}$$

For a propagating monochromatic wave

$$\mathcal{V}(\boldsymbol{r},t) = U(\boldsymbol{r})\cdot\exp(-i2\pi\nu t) \tag{2.2-7}$$

the complex amplitude $U(\boldsymbol{r})$ satisfies in a homogenous medium the ‘homogenous scalar Helmholtz equation’ [2.1]

$$\left(\nabla^2 + k^2\right)\cdot U(\boldsymbol{r}) = 0 . \tag{2.2-8}$$

The wavenumber

$$k = \frac{2\pi}{\lambda} = 2\pi\nu\frac{\sqrt{\varepsilon_r\mu_r}}{c} = 2\pi\nu\frac{n}{c}, \tag{2.2-9}$$

where λ is the wavelength in the medium, is given by the frequency ν, the vacuum speed of light c, and the refractive index n.

The simplest solution of the homogenous scalar Helmholtz equation is the plane wave given by [2.1]

$$U(\boldsymbol{r}) = A\cdot\exp(i\boldsymbol{kr}), \tag{2.2-10}$$

where A is the constant amplitude and the wavevector $\boldsymbol{k} = (k_x, k_y, k_z)$ points into the direction of propagation and has the length $|\boldsymbol{k}| = k$. Another simple solution is the diverging (outgoing) spherical wave [2.1]

$$U_l(\boldsymbol{r}_l) = \frac{A_0}{|\boldsymbol{r}_l - \boldsymbol{r}_0|}\exp\left(ik|\boldsymbol{r}_l - \boldsymbol{r}_0|\right), \tag{2.2-11}$$

where $\boldsymbol{r}_0$ is the origin of the source.

Figure 2.2-2: Two possible solutions of the homogenous scalar Helmholtz equation: left the plane wave, right the spherical wave.

2.3 Solutions of the scalar wave equation in form of diffraction integrals

2.3.1 The Kirchhoff diffraction integral

Kirchhoff expressed the solution of the homogenous Helmholtz equation (2.2-8) within a volume V in terms of the scalar field $U_{0I}(r_0)$ on the limiting closed surface S of the volume V (Fig. 2.3-1). The scalar field $U_I(r_I)_{HK}$ within this volume V can then be expressed by the integral

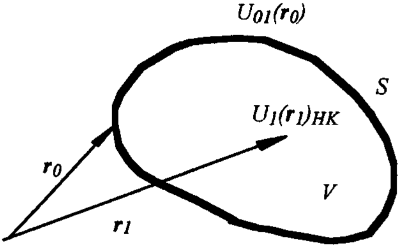


Figure 2.3-1: Volume V and surface S of the integral theorem of Helmholtz and Kirchhoff.

$$U_I(r_I)_{HK} = \frac{1}{4\pi}\iint_S \left\{ \frac{\partial U_{0I}(r_0)}{\partial n} G(r_I, r_0) - U_{0I}(r_0) \frac{\partial G(r_I, r_0)}{\partial n} \right\} ds, \tag{2.3-1}$$

where

$$G(r_I, r_0) = \frac{\exp(ik|r_I - r_0|)}{|r_I - r_0|} \tag{2.3-2}$$

is the auxiliary Green function (an outgoing spherical sphere) and $\partial/\partial n$ is the partial derivative in the outward normal direction at each point on the limiting surface. The Eq. (2.3-1) is known as the integral theorem of Helmholtz and Kirchhoff [2.2].

Let us consider an aperture of an infinite black screen perpendicular to the axis z, located at $z = 0$ (Fig. 2.3-2). The aperture is illuminated by a source emitting a monochromatic wave

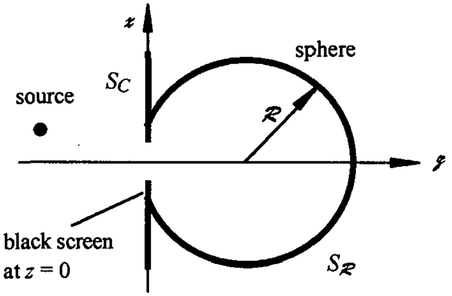


Figure 2.3-2: Black screen and sphere of the Kirchhoff diffraction integral.

placed somewhere in the half space $z < 0$. Then, the diffracted scalar field in the half space $z > 0$ can be evaluated by Eq. (2.3-1) with a surface of integration formed by the black screen S_C and a sphere $S_{\mathcal{R}}$ of radius $\mathcal{R}$. If the field on the spherical surface has the property

$$\lim_{\mathcal{R}\to\infty}\left\{\frac{\partial U_{01}(r_0)}{\partial n}G(r_1,r_0)-U_{01}(r_0)\frac{\partial G(r_1,r_0)}{\partial n}\right\}_{S_{\mathcal{R}}}=0, \qquad (2.3\text{-}3)$$

which is the Sommerfeld radiation condition, and if $\mathcal{R}$ is made to tend to infinity, the surface integral over the sphere $S_{\mathcal{R}}$ will vanish, so that the integration has only to be carried out over the aperture of the black screen. Hence,

$$U_1(r_1)_K=\frac{1}{4\pi}\iint_{\Sigma}\left\{\frac{\partial U_{01}(r_0)}{\partial n}G(r_1,r_0)-U_{01}(r_0)\frac{\partial G(r_1,r_0)}{\partial n}\right\}ds, \qquad (2.3\text{-}4)$$

where Σ is the area of the aperture. The integral of Eq. (2.3-4) is known as the Kirchhoff approximation or the Kirchhoff diffraction integral.

Evidently, the aperture on the black screen will not just limit the incident wave. Due to the interaction of wave and screen, the electromagnetic field will be perturbed at least in the vicinity of the rim of the aperture. Kirchhoff avoided this complication by introducing two boundary conditions [2.3]: first, the field distribution and its derivative on a black screen are exactly the same as they would be in absence of the screen and second, the field distribution and its derivative on the screen outside of the aperture (in the geometrical shadow) are both zero.

2.3.2 The Rayleigh-Sommerfeld diffraction integral

In the Rayleigh-Sommerfeld diffraction integral an alternative auxiliary Green function is used for the integral theorem of Helmholtz and Kirchhoff in Eq. (2.3-1). It consists of two spherical outgoing waves, placed symmetrically about the plane S_C (Fig. 2.3-3). This auxiliary Green function has the form

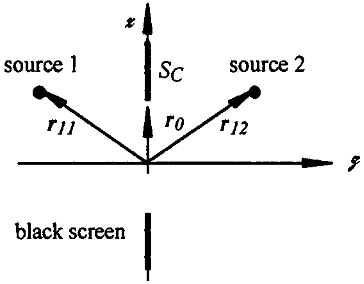


Figure 2.3-3: Notations for the alternative auxiliary Green function.

$$G(r_{11}, r_{12}, r_0)_{I,II} = \frac{\exp(ik|r_{11} - r_0|)}{|r_{11} - r_0|} \mp \frac{\exp(ik|r_{12} - r_0|)}{|r_{12} - r_0|} \qquad (2.3\text{-}5)$$

and yields two possible solutions depending on the phase shift between the two spherical waves, indicated by the minus and plus sign and referenced by the indices *I* and *II*, respectively. The solution referenced by *I* corresponds to the first Rayleigh-Sommerfeld diffraction integral and the solution referenced by *II*, to the second Rayleigh-Sommerfeld diffraction integral. Due to the imposed symmetry about the plane S_C it results that $|r_{11}-r_0| = |r_{12}-r_0|$ yielding for the function in Eq. (2.3-5) [2.2]

$$G(r_{1l}, r_{12}, r_0)_l \big|_{S_C} = 0 \qquad (2.3\text{-}6)$$

and

$$\left.\frac{\partial G(r_{1l}, r_{12}, r_0)_{II}}{\partial n}\right|_{S_C} = 0. \qquad (2.3\text{-}7)$$

Introducing Eq. (2.3-5) into Eq. (2.3-1), using the relations of Eq. (2.3-6) and (2.3-7) and applying the Sommerfeld radiation condition, the first Rayleigh-Sommerfeld diffraction integral takes the form [2.2]

$$U_l(r_l)_l = -\frac{1}{2\pi}\iint_{S_C} U_{0l}(r_0)\cdot\left(ik - \frac{1}{R}\right)\frac{\exp(ikR)}{R}\cdot\frac{z_l - z_0}{R}\cdot ds, \qquad (2.3\text{-}8)$$

where we have introduced the distance R as

$$R = \sqrt{(z_l - z_0)^2 + (x_l - x_0)^2 + (y_l - y_0)^2}. \qquad (2.3\text{-}9)$$

For convenience, we place the plane S_C at $z_0 = 0$ perpendicular to the axis z. We introduce the vector ρ with the components x and y, which is always perpendicular to the axis z. Furthermore, we define the scalar field $U_0(\rho_0) = U_{0l}(\rho, z_0 = 0)$ and the scalar field $U_l(r_l) = U_l(\rho_l, z_l)$ with $|z_l| > 0$. We shall call the plane with $U_0(\rho_0)$ the (secondary) source plane and the plane with $U_l(\rho_l, z_l = constant)$ the observation plane. In many cases, it can be assumed that the field U_0 is limited by an aperture having the area Σ and that $R >> \lambda$. Under these conditions and using the relation $z_l/R = \cos(\theta)$, the first Rayleigh-Sommerfeld diffraction integral in Eq. (2.3-8) becomes [2.3]

$$U_l(r_l)_l = \frac{1}{i\lambda}\iint_{\Sigma} U_0(\rho_0)\cdot\frac{\exp(ikR)}{R}\cdot\cos(\theta)\cdot d\sigma, \qquad (2.3\text{-}10)$$

where $d\sigma$ is the differential surface in the source plane Σ. Using similar derivation and the same notations as for the first Rayleigh-Sommerfeld diffraction integral, it can be shown that the second Rayleigh-Sommerfeld diffraction integral takes the form [2.2]

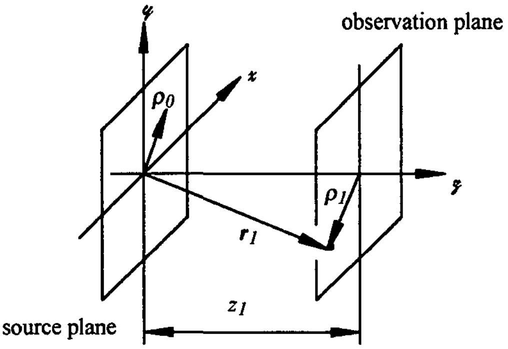


Figure 2.3-4: Definition of the geometry with ρ_0 and ρ_1.

$$U_1(r_1)_{II} = \frac{1}{2\pi}\iint_{\Sigma} \frac{\partial U_0(\rho_0)}{\partial n}\cdot\frac{\exp(ikR)}{R}\cdot d\sigma. \qquad (2.3\text{-}11)$$

It can be seen in Eq. (2.3-11) that in contrast to the first Rayleigh-Sommerfeld diffraction integral the partial derivative of the scalar field $U_0(\rho_0)$ must be known. However, usually just $U_0(\rho_0)$ is given so that the second Rayleigh-Sommerfeld diffraction integral is hardly used to evaluate the field of a propagating waves.

2.3.3 Fresnel diffraction integral

A second order approximation of the first Rayleigh-Sommerfeld diffraction integral is

$$U_1(x_1,y_1,z_1) = \frac{\exp(ikz_1)}{i\lambda z_1}\iint U_0(x_0,y_0)\cdot\exp\left(i\frac{k}{2z_1}\left[(x_1-x_0)^2+(y_1-y_0)^2\right]\right)\cdot dx_0dy_0, \qquad (2.3.\text{-}12)$$

which is the well known Fresnel (diffraction) integral or the Fresnel approximation. The approximation is based on the Taylor expansion [2.4] of the square root in Eq. (2.3-9). Using this expansion the distance R can be brought into the form

$$R = z_1 + \frac{(x_1-x_0)^2+(y_1-y_0)^2}{2\cdot z_1} - \frac{\left((x_1-x_0)^2+(y_1-y_0)^2\right)^2}{8\cdot z_1}\ldots \qquad (2.3\text{-}13)$$

The distance R of the term $cos(\theta)/R$ in Eq. (2.3-10) is approximated by the first term of the Taylor expansion. Furthermore, in the paraxial domain is $cos(\theta) \approx 1$ and therefore, $cos(\theta)/R \approx 1/z_I$. However, the distance R in the phase exponent of Eq. (2.3-10) has to be approximated with more care, because the wavenumber k is of the order of $10^7\ m^{-1}$ and therefore makes the phase very sensitive to the approximation of R. Thus, the first two terms of the Taylor expansion are retained in order to approximate the distance R in the phase exponent.

Expanding the brackets of the phase exponent, the Fresnel integral can be brought into the form

$$\begin{aligned} U_I(x_I,y_I,z_I) = &\frac{1}{i\lambda z_I}\cdot\exp\left(ik\left(z_I+\frac{x_I^2+y_I^2}{2z_I}\right)\right) \\ &\cdot\iint U_0(x_0,y_0)\cdot\exp\left(ik\frac{x_0^2+y_0^2}{2z_I}\right)\cdot\exp\left(-i2\pi\left(x_0\cdot\frac{x_I}{z_I\lambda}+y_0\cdot\frac{y_I}{z_I\lambda}\right)\right)dx_0\cdot dy_0 \end{aligned}, \qquad (2.3\text{-}14)$$

where the integral is recognized as the Fourier transform [2.5]. If we introduce the vector

$$q=\begin{pmatrix} q_x \\ q_y \end{pmatrix} \qquad (2.3\text{-}15)$$

with

$$q_x=\frac{x}{z\lambda} \qquad q_y=\frac{y}{z\lambda}, \qquad (2.3\text{-}16)$$

which are known as the spatial frequencies, we can introduce the Fourier transform by

$$FT\{U(\rho)\} \equiv \hat{U}(q)=\iint U(\rho)\cdot\exp(-i2\pi q\rho)\cdot d^2\rho, \qquad (2.3\text{-}17)$$

where the inverse transform is then given by [2.5]

$$FT^{-1}\{\hat{U}(q)\} \equiv U(\rho)=\iint \hat{U}(q)\cdot\exp(i2\pi q\rho)\cdot d^2q. \qquad (2.3\text{-}18)$$

The Fourier transform images the space domain into the (spatial) frequency domain, and vice versa for the inverse Fourier transform. $\hat{U}(q)$ is known as the complex angular spectrum and represents the wave $U(\rho)$ as a superposition of plane waves propagating in the directions

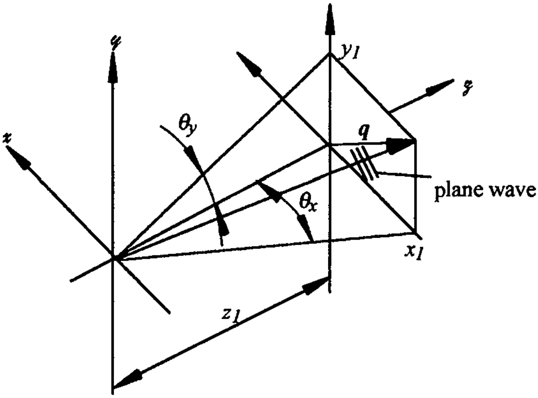


Figure 2.3-5: Spatial frequencies and angular spectrum of plane waves.

given by the angles $\theta_x \approx x/z$ and $\theta_y \approx y/z$ (paraxial approximation) in Eq. (2.3-16). Using the Fourier transform, Eq. (2.3-14) can be brought into the form

$$U_l(\rho_l, z_l) = \frac{\exp\left(ik\left(z_l + \frac{\rho_l^2}{2z_l}\right)\right)}{i\lambda z_l} \cdot FT\left\{U_0(\rho_0) \cdot \exp\left(ik\frac{\rho_0^2}{2z_l}\right)\right\}. \tag{2.3-19}$$

An extensive analysis of the error of the Fresnel approximation for a collimated beam, diffracted at the circular aperture having a radius ρ_a, was published by Forbes [2.6]. He compared the field U_F calculated by the Fresnel approximation (Eq. (2.3-19)) with the field U_{SF} calculated by the Rayleigh-Sommereld diffraction integral (Eq. (2.3-8)) and showed that for an error of $\varepsilon = |U_F - U_{SF}|/|U_{SF}|$ in the geometrically illuminated region, it is sufficient to have $z_l > \rho_a(\rho_a/(\lambda\varepsilon))^3$. In a succeeding publication [2.7], Forbes proposed an algebraic correction for the paraxial fields. It turns out that the Fresnel approximation can be improved by

$$\overline{U}_I(\rho_I,z_I) \approx \overline{U}_I(\rho_I,z_I)_F + \frac{jz_I}{2k}\cdot\left(\frac{\partial}{\partial z_I}\right)^2 \overline{U}_I(\rho_I,z_I)_F - \frac{z_I}{2k^2}\left[\left(\frac{\partial}{\partial z_I}\right)^3 \overline{U}_I(\rho_I,z_I)_F + \frac{z}{4}\left(\frac{\partial}{\partial z_I}\right)^4 \overline{U}_I(\rho_I,z_I)_F\right], \qquad (2.3\text{-}20)$$

where the index F indicates the field obtained by the Fresnel integral. The overbar is used to indicate that the associated field has been multiplied by $exp(-ikz_I)$ to suppress the most rapid variations in the paraxial field.

2.3.4 Fraunhofer approximation

The Fraunhofer approximation is an approximation of the Fresnel integral. At sufficiently large distance z_I exp $(ik^2/2z_I)$ in Fourier transform of Eq. (2.3-19) becomes approximately unity within the aperture of $U(\rho_0)$. Therefore, Eq. (2.3-19) may be approximated by

$$U_I(\rho_I,z_I) = \frac{\exp\left(ik\left(z_I + \frac{\rho_I^2}{2z_I}\right)\right)}{i\lambda z_I}\cdot FT\{U_0(\rho_0)\}, \qquad (2.3\text{-}21)$$

which is known as the Fraunhofer approximation [2.3].

2.3.5 Other forms of Rayleigh-Sommerfeld and Fresnel integral

The equations (2.3-8) and (2.3-12), which are the Rayleigh-Sommerfeld and Fresnel diffraction integrals, respectively, are in fact convolutions. Using the relation

$$\int\int_{-\infty}^{+\infty} U(x,y)\cdot h(X-x,Y-y)\cdot dx\cdot dy = FT^{-1}\{FT\{U(x,y)\}\cdot FT\{h(x,y)\}\}, \qquad (2.3\text{-}22)$$

which is the convolution theorem of the Fourier transform [2.5], a convolution can be calculated by Fourier transforms and a multiplication. This alternative is of great importance if the integrals in equations (2.3-8) and (2.3-12) have to be evaluated numerically (Chap. 7). Applying the theorem to the Rayleigh-Sommerfeld integral gives

$$U_1(\rho_1,z_1)_1 = FT^{-1}\left\{FT\{U_0(\rho_0)\}\cdot\exp\left(i2\pi z_1\cdot\sqrt{\left(\frac{1}{\lambda}\right)^2-q_x^2-q_y^2}\right)\right\}, \qquad (2.3\text{-}23)$$

where the Fourier transform pair

$$FT\left\{-\frac{1}{2\pi}\left(ik-\frac{1}{R}\right)\cdot\frac{z_1}{R}\cdot\frac{\exp(ikR)}{R}\right\} = \exp\left(i2\pi z_1\sqrt{\left(\frac{1}{\lambda}\right)^2-q_x^2-q_y^2}\right) \qquad (2.3\text{-}24)$$

has been used. This Fourier transform pair can be obtained by modification of the Weyl expansion formula [2.8]. For the Fresnel integral we get

$$U_1(\rho_1,z_1) = \exp(ikz_1)\cdot FT^{-1}\left\{FT\{U_0(\rho_0)\}\cdot\exp\left(-i\pi\lambda z_1\cdot\left(q_x^2+q_y^2\right)\right)\right\}, \qquad (2.3\text{-}25)$$

where the Fourier transform pair [2.3]

$$FT\left\{\exp\left(\frac{ik}{2z_1}\rho_0^2\right)\right\} = i\lambda z_1\cdot\exp\left(-i\pi\lambda z_1\cdot\left(q_x^2+q_y^2\right)\right) \qquad (2.3\text{-}26)$$

has been used.

Comparing the argument of the exponents of the right hand side in Eqs. (2.3-23) and (2.3-25), it can been shown straightforwardly, that the arguments in the exponents in Eq. (2.3-25) are the first two terms of the Taylor expansion of the square root function in exponent of Eq. (2.3-23). This Taylor expansion in the frequency domain results in an error of the Fresnel diffraction integral in Eq. (2.3-25) with respect to the Rayleigh-Sommerfeld integral in Eq. (2.3-23) which increases linearly with increasing z. This is in contrast to the Fresnel integral of Eq. (2.3-19), where the Taylor expansion of the root is in the space domain and the error decreases with increasing distance z.

Moreover, it can be seen by the Eqs. (2.3-10), (2.3-19), (2.3-23), and (2.3-25) that in the space domain the scalar field $U_1(\rho_1,z_1)$ is a linear superposition of spherical waves (for the Fresnel integral the spherical waves are parabolas, i.e. a paraxial approximation of spheres) and that in the frequency domain the scalar field is a linear superposition of plane waves. In Sec. (2.2) we have introduced both waves as solutions of the homogenous scalar Helmholtz equation. Thus, the Rayleigh-Sommerfeld and Fresnel diffraction integrals are solutions of this equation.

2.4 Imaging of deterministic fields in paraxial approximation

We investigate the relation of scalar fields between the two conjugate planes of a lens. Among other results, we show that the geometrical imaging relation between these two planes holds for the intensity distribution also within the Fresnel approximation. This result has been previously published by Yariv [2.9] and Kogelnik [2.10], but we will follow here a different approach.

The derivation is based on Figs. 2.4-1 and 2.4-2. Figure 2.4-1 shows the field of a plane wavefront $U_0(\rho_0)$ at $z_0 = 0$ focused by a thin lens with focal length f placed at $z_1 = +0$. The a_0 stands for the radius of the wave in the source plane at $z_0 = 0$ (e.g. the radius of an aperture or the waist of a Gaussian beam). The observation plane is placed at the distance $z_1 > 0$, where the field $U_1(\rho_1,z_1)$ is observed. We will use the expression 'converging wave' in order to refer to the propagating wave behind the lens in Fig. 2.4-1.

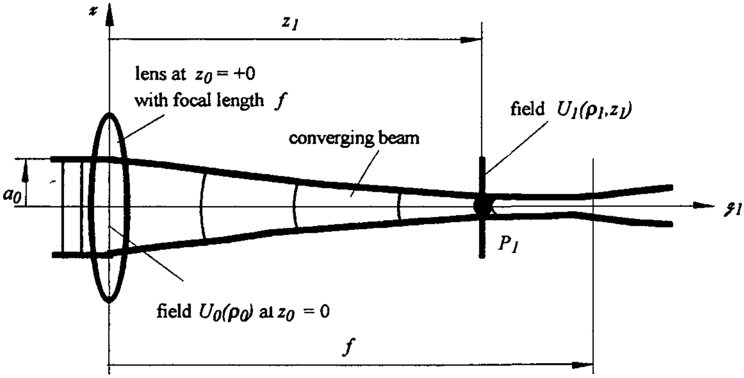


Figure 2.4-1: Propagation of a converging wave with $U_0(\rho_0)$ at $z_0 = 0$.

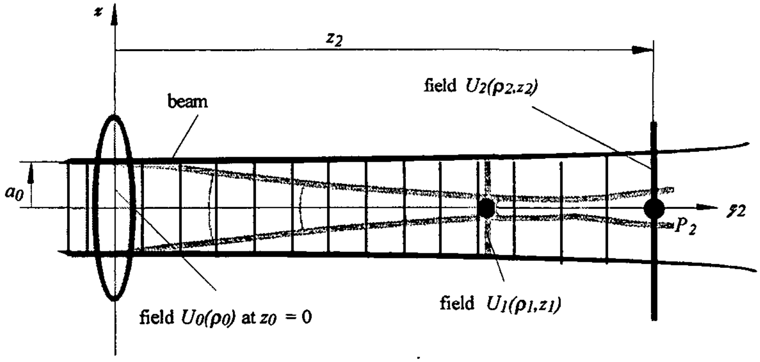


Figure 2.4-2: Propagation of a wave with $U_0(\rho_0)$ at $z_0 = 0$ without lens (solid line) and with lens (dashed line).

Figure 2.4-2 shows how the plane wavefront with field $U_0(\rho_0)$ of Fig. 2.4-1 would propagate freely into the half space $z_2 > 0$. We assume that the angular spectrum of the wave falls within a small solid angle. Such a wave is known as a (optical) beam [2.11] and we will use this expression in the following. An observation plane is placed at the distance z_2 ($z_2 > 0$), where the field $U_2(\rho_2,z_2)$ is observed.

We would like to point out that for the following we will keep strictly to the introduced symbols and indices. The index '*0*' refers always to the location $z = 0$ of both co-ordinate systems x_1, y_1, z_1 and x_2, y_2, z_2 of Figs. 2.4-1 and 2.4-2 respectively. The indices 1 and 2 refer to the corresponding co-ordinate system.

For convenience of derivation, we start the development with the beam $U_2(\rho_2,z_2)$ in Fig. 2.4-2 (bold lines). A beam propagating along the axis z_2 can be described by the Fresnel approximation of Eq. (2.3-19), i.e. by

$$U_2(\rho_2,z_2) = \frac{\exp\left(ikz_2 + \frac{ik}{2z_2}\rho_2^2\right)}{i\lambda z_2} \cdot FT\left\{U_0(\rho_0)\cdot \exp\left(\frac{ik}{2z_2}\rho_0^2\right)\right\}. \tag{2.4-1}$$

In Fig. 2.4-1 a thin aberration free lens with focal length f has been introduced. Therefore, the wave function $U_0(\rho_0)$ has to be multiplied by the phase function ϕ of the lens, which is

$$\phi(\rho_0) = -\frac{ik}{2f}\rho_0^2. \tag{2.4-2}$$

The converging wave $U_1(\rho_1, z_1)$ in Fig. 2.4-1 is then found to be given by

$$U_1(\rho_1, z_1) = \frac{\exp\left(ikz_1 + \frac{ik}{2z_1}\rho_1^2\right)}{i\lambda z_1} \cdot FT\left\{U_0(\rho_0) \cdot \exp\left(\frac{ik}{2} \cdot \rho_0^2 \cdot \left(\frac{1}{z_1} - \frac{1}{f}\right)\right)\right\}. \tag{2.4-3}$$

If the condition

$$\frac{1}{z_2} = \frac{1}{z_1} - \frac{1}{f} \tag{2.4-4}$$

is satisfied, the Fourier transforms of Eqs. (2.4-1) and (2.4-3) are equal. Equation (2.4-4) is recognized as the condition for geometric imaging, i.e. the Gaussian lens formula, and z_1 and z_2 are known as the positions of the conjugate planes of the lens. Furthermore, if we compare their Fourier kernels in the Eq. (2.4-1) and in Eq. (2.4-3) using the definition of the Fourier transform given in Eq. (2.3-17), we get for the conjugate planes

$$\frac{ik}{z_2} \cdot (\rho_2 \cdot \rho_0) = \frac{ik}{z_1} \cdot (\rho_1 \cdot \rho_0), \tag{2.4-5}$$

which leads to the magnification

$$|\beta| = \frac{|\rho_2|}{|\rho_1|} = \frac{|z_2|}{|z_1|} \tag{2.4-6}$$

well known from geometrical optics. Thus, if we neglect the phase terms kz_1 and kz_2 which are constant in the conjugate planes, the relation between the fields in Figs. 2.4-1 and 2.4-2 are given by

$$U_2(\rho_2, z_2) = \frac{\exp\left(ik \cdot (\beta - 1) \cdot \frac{\rho_1^2}{2z_1}\right)}{\beta} \cdot U_1(\rho_1, z_1). \tag{2.4-7}$$

It is seen from Eq. (2.4-7) that unless $\beta = 1$, the imaging relation between the fields in the two planes, as it would be expected from the geometrical optics, is perturbed by a quadratic phase term. This phase term is the difference of the phase exponents found in front of the integral in Eqs. (2.4-1) and (2.4-3). Their phase values depend on the actual distance between the lens and the conjugate planes and are independent on the field distributions in these planes.

The intensity distribution is given by the $I = |U|^2$. Consequently, the intensity distributions between the planes are related by

$$I_2(\rho_2, z_2) = \frac{1}{\beta^2} I_l(\rho_l, z_l), \tag{2.4-8}$$

which corresponds well to the imaging relation of conjugate planes in geometrical optics.

The relation of the complex angular spectra between the conjugate planes are given by the Fourier transform of the Eq. (2.4-7)

$$\hat{U}_2(q_2, z_2) = \frac{1}{\beta} \cdot FT_2\left\{\exp\left(ik \cdot (\beta - 1) \cdot \frac{\rho_l^2}{2z_l}\right)\right\} \otimes FT_2\{U_l(\rho_l, z_l)\}, \tag{2.4-9}$$

where the index 2 of the Fourier operator indicates that the Fourier kernel of the Eq. (2.4-1) is taken. Using the similarity theorem of Fourier transform [2.5], i.e.

$$FT\left\{U\left(\frac{x}{a}, \frac{y}{b}\right)\right\} = |ab| \cdot \hat{U}(aq_x, bq_y), \tag{2.4-10}$$

we find for the right hand side expression of the convolution sign in Eqs.(2.4-9)

$$FT_2\left\{U_l\left(\rho_l = \frac{\rho_2}{\beta}, z_l\right)\right\} = \beta^2 \cdot \hat{U}_l(q_l = q_2\beta, z_l). \tag{2.4-11}$$

Hence, the relation of the complex angular spectra between the conjugate planes is

$$\hat{U}_2(q_2, z_2) = \beta \cdot FT_2\left\{\exp\left(ik \cdot (\beta - 1) \cdot \frac{\rho_l^2}{2z_l}\right)\right\} \otimes \hat{U}_l(q_l, z_l) \tag{2.4-12}$$

where

$$|\beta| = \frac{|q_l|}{|q_2|} \tag{2.4-13}$$

is found. The equation (2.4-13) can be considered as the sine condition for the paraxial optics. From Eq. (2.4-12) it can be seen that only in the case where $\beta = 1$ (the left hand side in Eq. (2.4-12) is then a Dirac), the complex angular spectra of the scalar fields in the two conjugate planes are the same. The same is found for the spectral radiant intensity distribution

$$J_S(q) = \hat{U}^*(q) \cdot \hat{U}(q), \tag{2.4-14}$$

which describes the distribution of the radiation in the space as function of a spatial frequencies. When $\beta = 1$ we get

$$J_{S2}(q_2, z_2) = J_{S\,1}(q_1, z_1), \tag{2.4-15}$$

i.e. the same spectral radiant intensities in the conjugate planes. The *spectral* radiant intensity J_S has not to be confused with the radiant intensity J which describes the distribution of the radiation in free space as a function of angles [2.11]. However, in the paraxial approximation both become approximately equal. We may consider $J_S(q,z)$ as a function describing the angular distribution of the radiation in planes by the slopes of the rays given by $q\lambda$ (Eq. 2.3-16). Then, we find that the Eq. (2.4-15) is an exact description of geometrical rays in an optical imaging system. However, this agreement is restricted for the case where $\beta = 1$.

The linear imaging of the fields between the two conjugate planes can be enforced by placing a further lens in one of the observation planes, which compensates the perturbing quadratic phase term. If we would like to image the field on observation plane 2 to the observation plane 1 we have to mount the lens in the observation plane 1. The mounted lens must compensate the explicit phase exponent on the right hand side in Eq. (2.4-7). This is achieved if the focal length of the lens is

$$f = -\left(\frac{z_1}{\beta - 1}\right). \tag{2.4-16}$$

In the described situation of Figs. 2.4-1 and 2.4-2 we have $\beta > 0$ and $z_1 > 0$. Hence, $f < 0$ and therefore the lens must be negative. With this correcting lens and by neglecting the linear phase

kz_1 and kz_2 we get finally the linear imaging of the scalar fields in the space and frequency domain, i.e.

$$U_2(\rho_2,z_2)=\frac{1}{\beta}U_1(\rho_1,z_1) \text{ and } \hat{U}_2(q_2,z_2)=\beta\cdot\hat{U}_1(q_1,z_1). \qquad (2.4\text{-}17)$$

The focal length of the lens is a function of z_1 (Eq. (2.4-16)). Hence, phase compensation cannot be carried out by a single lens if a field imaging for different distances is required.

2.5 Conclusion

Starting with the electromagnetic vector field we have introduced the scalar field by the scalar form of the wave equation. We went on in describing the propagation by diffraction integrals. From the integral theorem of Helmholtz and Kirchhoff we derived the Rayleigh-Sommerfeld diffraction integral and the Fresnel and Fraunhofer diffraction integrals as approximations. Using the Fourier transform, other forms of Rayleigh-Sommerfeld and Fresnel diffraction integrals have been presented.

We have shown that within the approximation of Fresnel diffraction and for scalar deterministic and monochromatic waves, the intensity distributions in the two conjugate planes of a lens are related by the imaging law of the geometrical optics. We have found that this linear imaging relation cannot be extended to the field and the angular spectrum, unless the magnification is unity. However, we have shown that linear imaging of field and angular spectrum can be enforced by placing an additional lens in one of the conjugate planes.

So far we have only treated the case of deterministic fields. However, most sources are only partially coherent, emitting a field with random characteristics. For this type of fields the description in this chapter is not sufficient. We need to extend the theory to the cross-spectral density function, which will be discussed in the following chapter.

A further problem is the numerical evaluation of the diffraction integrals. For two-dimensional wavefronts these integrals take quite a long time be evaluated on computers. Using the fast Fourier transform, the computation time can be reduced significantly. However, in the case of wavefronts with small radii of curvature, the number of necessary sampling points may exceed the available memory space of most desktop computers. A way of reducing the number of required sample points will be presented in Chap. 7.

References

[2.1] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Oxford, Reprint, 1993), pp. 1-23.

[2.2] M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Wiley Series in Pure and Applied Optics, New York, 1991), pp. 184-207.

[2.3] J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill Series in Electrical and Computer Engineering, New York, 1996, 2nd edition), pp. 4-51.

[2.4] M. Abramowitz, I. Stegun, *Pocketbook of Mathematical Functions* (Verlag Harri Deutsch-Thun-Frankfurt/Main, 1984), pp. 28

[2.5] R. N. Bracewell, *The Fourier Transform and its Applications* (McGraw-Hill International Editions, New York, 1986, 2nd edition), pp. 1-176.

[2.6] G.W. Forbes, "Validity of the Fresnel approximation in the diffraction of collimated beams," J. Opt. Soc. Am. A **13**, 1818-1826 (1996).

[2.7] G. W. Forbes *et al.*, "Algebraic corrections for paraxial wave fields," J. Opt. Soc. Am. A **14**, 3300-3315 (1997).

[2.8] J. E. Harvey, " Fourier treatment of near-field scalar diffraction theory," Am. J. Phys. **47**, 974-980 (1979).

[2.9] A. Yariv, "Imaging of coherent fields through lens like systems," Opt. Lett. **19**, 1607-1608 (1994).

[2.10] H. Kogelnik, "Imaging of coherent of optical modes – resonators with internal lenses," Bell Sys. Tech. J. **44**, 455-494 (1965).

[2.11] L. Mandel and E. Wolf, *Optical coherence and Quantum Optics* (Cambridge University Press, New York, 1995), pp. 229-287.

CHAPTER 3

Propagation of random scalar wave fields

3. 1 Introduction

A random scalar wave field can be understood as a field consisting of many non-correlated deterministic fields. The superposition of the non-correlated deterministic fields will generate a field whose amplitude and phase do not depend anymore on the time *t* in a deterministic way. Therefore, such scalar fields can only be expressed by a random function describing the field as a probability of occurrence of amplitude and phase values.

In this chapter we describe mathematically the random field. We start with giving a short overview of the mathematical description of random processes. Then, we introduce the cross-spectral density function, which gives a full description of a random field. We shall show how this function can be measured and decomposed into deterministic fields. Finally, we describe the radiated random field of a source using the cross-spectral density function.

3.2 Description of a random process or a random function

We shall treat the scalar random wave field as the electromagnetic field of a random process. Therefore, the first section of this chapter will introduce the necessary expressions and equations which describe such a process in general [3.1, 3.2].

3.2.1 The ensemble and time average

A function $x(t)$, which does not depend on the time t on a deterministic way, is called a random process or a random function. Its values can only be described by a probability distribution or probability density function $p(x,t)$ which assigns to a value x over a small interval dx at a given time t a probability of occurrence. The integration of $p(x,t)$ over the entire interval of x is normalized to one indicating that the occurrence of a value within the interval is certain at any time t. The average or the expectation of x at time t is given by

$$E[x(t)] = \int x \cdot p(x,t) \cdot dx. \qquad (3.2\text{-}1)$$

Alternatively, the average of x may be evaluated by a collection of all possible realizations of $x(t)$. This countable collection of all possible realizations is known as the ensemble of $x(t)$. The average of x at time t may then be written as

$$E[x(t)] = \lim_{N \to \infty} \left(\frac{1}{N} \sum_{r=1}^{N} x_r(t) \right), \qquad (3.2\text{-}2)$$

where the possible realizations are labeled by the index r.

The time average of the r-th ensemble is given by

$$\langle x_r(t) \rangle = \lim_{T \to \infty} \left(\frac{1}{T} \cdot \int_{t-T/2}^{t+T/2} x_r(t') \cdot dt' \right). \qquad (3.2\text{-}3)$$

3.2.2 Stationarity

A random function is said to be strictly stationary if the probability density function does not change with time. A random process is said to be stationary in the wide sense if the following two conditions are satisfied:

$$E[x(t)] = constant \qquad (3.2\text{-}4)$$

and

$$E[x(t_2) \cdot x(t_2)] = E[x(t) \cdot x(t+\tau)]. \qquad (3.2\text{-}5)$$

A strictly stationary random process is always stationary in the wide sense. However, a wide-sense stationary process does not need to be strictly stationary.

3.2.3 Ergodicity

If every realization r of ensembles carries the same statistical information about the stationary process as every other realization, then the process is said to be ergodic. In such a case all time averages of possible ensembles $<x_r(t)>$ are equal to the ensemble average $E[x(t)]$. Hence, ergodicity can be described by

$$E[x(t)] = \langle x_r(t) \rangle \qquad \forall r. \tag{3.2-6}$$

An ergodic process is always strictly stationary.

3.3 Description of random fields

We assume that the scalar field $\mathcal{V}(\boldsymbol{r},t)$ is an ergodic random function of t and location $\boldsymbol{r}$. We can imagine such a field as superposition of many non-correlated deterministic scalar fields. Using the relations in Eqs. (3.2-5) and (3.2-6) the cross-correlation function of such a random field is given by [3.1]

$$\Gamma(\boldsymbol{r}_1,\boldsymbol{r}_2,\tau) = \langle \mathcal{V}^*(\boldsymbol{r}_1,t) \cdot \mathcal{V}(\boldsymbol{r}_2,t+\tau) \rangle, \tag{3.3-1}$$

which is known as the mutual coherence function of the complex scalar random field $\mathcal{V}(\boldsymbol{r},t)$ and describes the correlation of the field between two points at $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. If $\boldsymbol{r}_1 = \boldsymbol{r}_2$ and $\tau = 0$ the mutual coherence function becomes the optical intensity $I(\boldsymbol{r})$, which has been introduced by Eq. (2.2-3).

The cross-spectral density function

$$W(\boldsymbol{r}_1,\boldsymbol{r}_2,\nu) = \langle \mathcal{U}^*(\boldsymbol{r}_1,\nu) \cdot \mathcal{U}(\boldsymbol{r}_2,\nu) \rangle, \tag{3.3-2}$$

is defined as the correlation of the field between two points at $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ at a given temporal frequency ν. The random function $\mathcal{U}(\boldsymbol{r},\nu)$ has not to be confused with the $U(\boldsymbol{r},\nu)$, which is the function of a deterministic field. The functions $\mathcal{U}(\boldsymbol{r},\nu)$ and $\mathcal{V}(\boldsymbol{r},t)$ are related by a Fourier transform, as given by Eq. (2.2-6). Furthermore, according to the generalized Wiener-Khintchine theorem [3.1], the mutual coherence function and the cross-spectral density function form a Fourier transform pair:

$$W(\boldsymbol{r}_1,\boldsymbol{r}_2,\nu) = \int_{-\infty}^{+\infty} \Gamma(\boldsymbol{r}_1,\boldsymbol{r}_2,\tau) \cdot \exp(i2\pi\nu\tau) \cdot d\tau. \tag{3.3-3}$$

In the special case where $\boldsymbol{r}_1 = \boldsymbol{r}_2$, we have

$$S(\boldsymbol{r},\nu) = W(\boldsymbol{r},\boldsymbol{r},\nu), \tag{3.3-4}$$

which is the spectral density function. The normalized cross-spectral density function is given by

$$\mu(\mathbf{r}_1,\mathbf{r}_2,\nu)=\frac{W(\mathbf{r}_1,\mathbf{r}_2,\nu)}{\sqrt{S(\mathbf{r}_1,\nu)S(\mathbf{r}_2,\nu)}}, \qquad (3.3\text{-}5)$$

which is known as the spectral degree of coherence. It describes the correlation between two points of the field. For a deterministic field the spectral degree of coherence is $|\mu(\mathbf{r}_1,\mathbf{r}_2,\nu)|=1$, since the field in each point is defined by a function and hence, its value is precisely predictable at any time. Such fields are said to be coherent. In contrast, if $|\mu(\mathbf{r}_1,\mathbf{r}_2,\nu)|\neq 1$ the field contains random components and hence, the amplitude and phase of the field can be predicted only with a certain probability. These fields are said to be partially coherent. In the special case where in a domain or in a plane (e.g. of a source) $|\mu(\mathbf{r}_1,\mathbf{r}_2,\nu)|\rightarrow 0$ the field is considered as incoherent in that domain or plane. A spectral coherence length may be defined, which is the distance between two points $|\mathbf{r}_1-\mathbf{r}_2|$ having a given minimum of spectral degree of coherence [3.1, 3.2, 3.3].

We define the spectral radiant cross-intensity function as

$$\hat{W}_S(\mathbf{q}_1,\mathbf{q}_2,\nu)=\left\langle FT^*\{\mathcal{U}(\mathbf{r}_1,\nu)\}\cdot FT\{\mathcal{U}(\mathbf{r}_2,\nu)\}\right\rangle, \qquad (3.3\text{-}6)$$

which has to be understood as an analogy to the angular spectrum $\hat{U}(\mathbf{q})$ of the deterministic field. The spectral radiant intensity J_S, which quantifies the energy flow at the spatial frequency $\mathbf{q}$, is then given by

$$J_S(\mathbf{q},\nu)=\hat{W}_S(\mathbf{q},\mathbf{q},\nu), \qquad (3.3\text{-}7)$$

from which we can evaluate the total energy flow

$$\Phi(\nu)=\int_{-Q}^{+Q} J_S(\mathbf{q},\nu)\cdot d^2\mathbf{q} \qquad (3.3\text{-}8)$$

radiated into the space. The spectrum of the spatial frequency $\mathbf{q}$ has to be limited to

$$Q=\sqrt{q_x^2+q_y^2}=1/\lambda \qquad (3.3\text{-}9)$$

in order to suppress the evanescent waves, which do not contribute to the energy flow. From Eqs. (3.3-7) and (3.3-8) we see that only the diagonal elements of the cross-spectral density function contribute to the radiant intensity and to the energy flow.

Two possible set-ups to measure the cross-spectral density function are shown in Fig. 3.3-1. The set-ups are similar to the Young's interference experiment and to the Michelson stellar interferometer, respectively. Compared to the Young's interference experiment, the Michelson stellar interferometer has the advantage that the energy flow at the detector is higher, so that the cross-spectral density function of weaker fields can be measured.

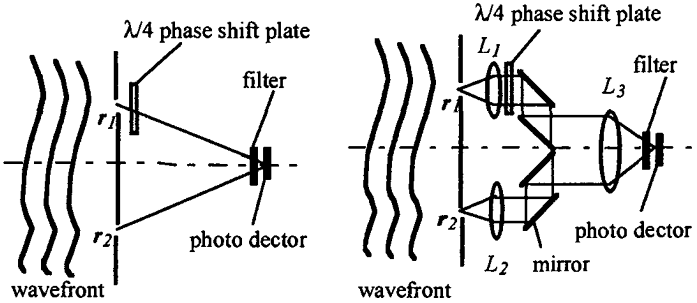


Figure 3.3-1: Two set-ups for measuring the cross-spectral density function. Left the Young's interference experiment and right the Michelson stellar interferometer. The filter in front of the photo detector selects the wavelength.

Let us now first calculate the intensities measured by photo detector depending on the condition (closed or open) at the holes 1 and 2 and then find out how $W(r_1,r_2)$ can measured. The incident wave has the form

$$\mathcal{U}(r) = a(r)\cdot \exp(i\tilde{\varphi}(r))\cdot \exp(-i2\pi\nu t), \tag{3.3-10}$$

where $a(r)$ and $\tilde{\varphi}(r)$ are the random amplitude and random phase functions, respectively. Then, the intensities measured by the photo detector are:

- hole 1 open and hole 2 closed

$$I_1 = \langle \mathcal{U}^*(r_1)\cdot \mathcal{U}(r_1)\rangle = W(r_1,r_1), \tag{3.3-11}$$

- hole 1 closed and hole 2 open

$$I_2 = \left\langle \mathcal{U}^*(r_2)\cdot \mathcal{U}(r_2)\right\rangle = W(r_2,r_2), \tag{3.3-12}$$

- both holes open, without the λ/4 phase shift plate

$$I_3 = \left\langle \left|\mathcal{U}(r_1)+\mathcal{U}(r_2)\right|^2\right\rangle = \left\langle \left|\mathcal{U}(r_1)\right|^2\right\rangle + \left\langle \left|\mathcal{U}(r_2)\right|^2\right\rangle + Re\left\{\left\langle \mathcal{U}^*(r_1)\mathcal{U}(r_2)+\mathcal{U}(r_1)\mathcal{U}^*(r_2)\right\rangle\right\} \tag{3.3-13}$$

where *Re* stands for real part. Using the Eqs. (3.3-11) and (3.3-12), Eq. (3.3-13) can brought into the form

$$\begin{aligned} I_3 &= W(r_1,r_1)+W(r_2,r_2)+2\cdot\left\langle a(r_1)\cdot a(r_2)\cdot\cos\left(\tilde{\varphi}(r_2)-\tilde{\varphi}(r_1)\right)\right\rangle \\ &= W(r_1,r_1)+W(r_2,r_2)+2\cdot Re\left\{W(r_1,r_2)\right\} \end{aligned}, \tag{3.3-14}$$

- both holes open and with the λ/4 phase shift plate; a phase shift of $\Delta\varphi = +\pi/2$ for $\tilde{\varphi}(r_1)$ is introduced causing the cosine function in Eq. (3.3-14) to become a sine function, thus

$$\begin{aligned} I_4 &= W(r_1,r_1)+W(r_2,r_2)+2\cdot\left\langle a(r_1)\cdot a(r_2)\cdot\sin\left(\tilde{\varphi}(r_2)-\tilde{\varphi}(r_1)\right)\right\rangle \\ &= W(r_1,r_1)+W(r_2,r_2)+2\cdot Im\left\{W(r_1,r_2)\right\} \end{aligned}, \tag{3.3-15}$$

where *Im* stands for imaginary part.

The function $W(r_1,r_2)$ is then obtained from

$$W(r_1,r_2) = \frac{1}{2}\left(\left(I_3-\left(I_1+I_2\right)\right)+i\cdot\left(I_4-\left(I_1+I_2\right)\right)\right) \tag{3.3-16}$$

and since the cross-spectral density function is Hermitian

$$W(r_2,r_1) = W^*(r_1,r_2). \tag{3.3-17}$$

$W(r_1,r_1)$ and $W(r_2,r_2)$ are directly obtained from I_1 and I_2 respectively.

Other procedures can be developed using different techniques of temporal phase measurement described in [3.4] and [3.5].

In the case where the polarization varies with $\boldsymbol{r}$ the measurement of I_1 to I_4 has to be repeated but with a polarization rotated about 90 degrees in one of the two locations given by r_1 and r_2.

3.4 Decomposition of random fields into orthogonal deterministic fields

We have mentioned in the beginning of the preceding section that a random field can be considered as a superposition of non-correlated deterministic fields. The decomposition of a random field into orthogonal deterministic fields can be achieved by solving the integral equation

$$\int W(r_1,r_2,\nu)\cdot\psi_k(r_1,\nu)\cdot d^3r_1=\alpha_k(\nu)\cdot\psi_k(r_2,\nu) \qquad (3.4\text{-}1)$$

known as the homogenous Fredholm integral equation of second kind [3.1]. By solving this equation the deterministic fields are presented by (orthonormalized) eigenfunctions $\psi_k(r,\nu)$ and eigenvalues $\alpha_k(\nu)$. Such deterministic fields are known as modes. The cross-spectral density function is then expressed by

$$W(r_1,r_2,\nu)=\sum_K\alpha_k(\nu)\cdot\psi_k^*(r_1,\nu)\cdot\psi_k(r_2,\nu), \qquad (3.4\text{-}2)$$

where the eigenfunctions have the property

$$|\mu_k(r_1,r_2,\nu)|=\left|\frac{\psi_k^*(r_1,\nu)\cdot\psi_k(r_2,\nu)}{\sqrt{\psi_k^*(r_1,\nu)\cdot\psi_k(r_1,\nu)}\cdot\sqrt{\psi_k^*(r_2,\nu)\cdot\psi_k(r_2,\nu)}}\right|=1, \qquad (3.4\text{-}3)$$

which is evident since they describe deterministic fields. In the publications [3.6, 3.7, 3.8, 3.9] different methods are presented to determine the deterministic fields once the cross-spectral density function or just the spectral density function is known.

3.5 Radiation from a planar secondary source

3.5.1 Cross-correlation density function of secondary planar sources

Many sources are planar secondary sources. A source of this kind is usually an aperture in an opaque planar screen, illuminated either directly or via an optical system by a primary source.

Laser diodes can be treated as a planar secondary sources if we consider the dielectric mirror of the front side at the chip as the aperture of a screen.

In the following we assume that the source emits different frequencies ν and that for each frequency the emitted field consists of many non-correlated deterministic fields. As discussed in Sec. 3.4, we may decompose the field by a linear superposition of eigenfunctions $\psi_k(\rho_0, \nu)$ with eigenvalues $\alpha_k(\nu)$. If we define $\mathcal{A}(\nu)$ as a complex random amplitude ($\mathcal{A}(\nu) = a(\nu) \cdot \exp(i\tilde{\varphi}(\nu))$) the emitted field of a source can then be described by

$$\mathcal{U}(\rho_0,\nu) = \sum_K u_k(\rho_0,\nu) = \sum_K \mathcal{A}_k(\nu) \cdot \psi_k(\rho_0,\nu), \tag{3.5-1}$$

where

$$\langle \mathcal{A}_k^*(\nu) \cdot \mathcal{A}_l(\nu) \rangle = \alpha_k(\nu) \cdot \delta_{k,l}. \tag{3.5-2}$$

The expression $u_k(\rho_0, \nu) = \mathcal{A}_k(\nu) \cdot \psi_k(\rho_0, \nu)$ may be considered as a sub-source emitting a single-mode wave with a random amplitude $\mathcal{A}_k(\nu)$. The eigenfunctions $\psi_k(\rho_0, \nu)$ with the condition of Eq. (3.5-2) describe the oscillating modes of the source which are known as natural modes of oscillation [3.1].

A frequently used model for the source is the Gaussian Schell-model, which has two principal properties. First, the spectral density function is given by

$$S_G(\rho_0,\nu) = \langle |\mathcal{A}(\nu)|^2 \rangle \cdot \exp\left(-\frac{\rho_0^2}{2\sigma_S^2(\nu)}\right), \tag{3.5-3}$$

where $\sigma_S(\nu)$ defines the extension of the source. Second, the spectral degree of coherence depends only of the difference $|\rho_{01} - \rho_{02}| = \rho_0'$, i.e. it has the form

$$\mu_G(\rho_0',\nu) = \exp\left(-\frac{\rho_0'^2}{2\sigma_\mu^2(\nu)}\right), \tag{3.5-4}$$

where $\sigma_\mu(\nu)$ defines the spectral coherence length of the source. In the limit $\sigma_\mu(\nu) \rightarrow \infty$ and $S_G(\rho_0, \nu) = S_G(\rho_0, \nu)\delta(\nu-\nu')$ the Gaussian Schell-model describes a Gaussian single-mode wave.

For a one dimensional Gaussian Schell-model source the normalized eigenfunctions, i.e. the natural modes of oscillation, are found to be [3.10]

$$\psi_k(x,\nu)=\left(\frac{2d}{\pi}\right)^{1/4}\frac{1}{\sqrt{\left(2^k k!\right)}}H_k\left[x\sqrt{(2d)}\right]\cdot\exp\left(-dx^2\right) \tag{3.5-5}$$

where $H_k(x)$ are Hermite polynomials and where

$$a(\nu)=\frac{1}{4\sigma_S^2(\nu)},\ b(\nu)=\frac{1}{2\sigma_\mu^2(\nu)},\ d(\nu)=\sqrt{a^2(\nu)+2a(\nu)b(\nu)}. \tag{3.5-6}$$

The eigenvalues are given by

$$\alpha_k(\nu)=\left\langle\left|A(\nu)\right|^2\right\rangle\sqrt{\left(\frac{\pi}{a+b+d}\right)}\cdot\left(\frac{b}{a+b+d}\right)^k. \tag{3.5-7}$$

The ratio $\alpha_k(\nu)/\alpha_0(\nu)$ as a function of k for different values of the normalized coherence length $\sigma_n(\nu)=\sigma_\mu(\nu)/\sigma_S(\nu)$ is shown in Fig. 3.5-1. It is seen that the coherence length decreases

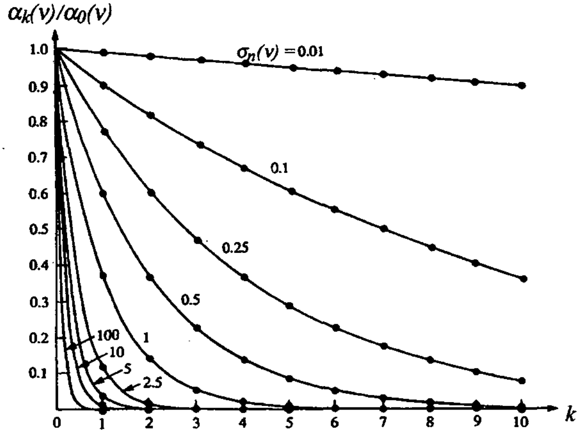


Figure 3.5-1: The ratio of the $(k+1)$th eigenvalue $\alpha_k(\nu)$ to the lowest-order eigenvalue $\alpha_0(\nu)$ of a one-dimensional, secondary Gaussian Schell-model source as a function of k for different values of the normalized coherence length $\sigma_n(\nu)=\sigma_\mu(\nu)/\sigma_S(\nu)$ [3.1].

with increasing contribution of higher natural modes of oscillation. Furthermore, the eigenvalues of the Gaussian Schell-model source are so that

$$\alpha(\nu)_0 \geq \alpha(\nu)_1 \geq \alpha(\nu)_2 \ldots \geq \alpha(\nu)_N, \qquad (3.5\text{-}8)$$

which is an important information about the source for the possibility of squeezing of the phase space volume (Chap.5).

3.5.2 Radiation from a secondary planar source

Let us derive the cross-correlation density function of a radiating secondary planar source. The derivation could be carried out by solving the Helmholtz equation for the cross-correlation density function [3.1], but we will base the derivation on the decomposition of the random field into the deterministic fields, since the propagation of deterministic fields has been discussed in Chap. 2. Let us assume that the secondary source is described by the Eq. (3.5-1). Using the first Rayleigh-Sommerfeld diffraction integral from Eq. (2.3-10) for the propagation of the deterministic fields we get

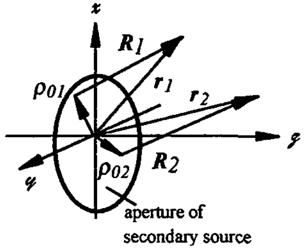


Figure 3.5-2: Radiation from a secondary source.

$$W(r_1,r_2,\nu) = \sum_K \left(\left\langle \left(\frac{1}{i\lambda} \iint_\Sigma a_k(\rho_{01},\nu) \cdot \frac{\exp(ikR_1)}{R_1} \cdot \cos(\theta_{01}) \cdot d^2\rho_{01} \right)^* \cdot \left(\frac{1}{i\lambda} \iint_\Sigma a_k(\rho_{02},\nu) \cdot \frac{\exp(ikR_2)}{R_2} \cdot \cos(\theta_{02}) \cdot d^2\rho_{02} \right) \right\rangle \right) \qquad (3.5\text{-}9)$$

The time averaging operator <> can be taken inside the integral, since integration is not carried out over the time. If we introduce

$$w_k(\rho_{01},\rho_{02},\nu)=\left\langle a_k^*(\rho_{01},\nu)\cdot a_k(\rho_{02},\nu)\right\rangle, \tag{3.5-10}$$

and recalling that

$$\left\langle a_k^*(\rho_1,\nu)\cdot a_l(\rho_2,\nu)\right\rangle=\delta_{k,l} \tag{3.5-11}$$

because $a_k(\rho,\nu)$ are orthogonal and non-correlated functions (see Eqs. (3.5-1) and (3.5-2)), we may write for Eq. (3.5-9)

$$W(\mathbf{r}_1,\mathbf{r}_2,\nu)=\frac{1}{\lambda^2}\cdot\int\limits_{\Sigma}\int\limits_{\Sigma}W(\rho_{01},\rho_{02},\nu)\cdot\cos(\theta_{01})\cdot\cos(\theta_{02})\cdot\frac{\exp(ik(R_1-R_2))}{R_1\cdot R_2}\cdot d^2\rho_{01}\cdot d^2\rho_{02} \tag{3.5-12}$$

where

$$W(\rho_{01},\rho_{02},\nu)=\sum_K w_k(\rho_{01},\rho_{02},\nu). \tag{3.5-13}$$

The cross-correlation density function $W(\mathbf{r}_1,\mathbf{r}_2,\nu)$ in Eq. (3.5-12) may be expressed by $\mathscr{A}_k(\nu)$ and $\Psi_k(\mathbf{r},\nu)$ where $\mathscr{A}_k(\nu)$ is the complex random amplitude of Eq. (3.5-2) and $\Psi_k(\mathbf{r},\nu)$ the Rayleigh-Sommerfeld integral of the function $\psi_k(\rho_0,\nu)$. Using superposition we can express Eq. (3.5 -12) by

$$W(\mathbf{r}_1,\mathbf{r}_2,\nu)=\sum_K\alpha_k(\nu)\cdot\Psi_k^*(\mathbf{r}_1,\nu)\cdot\Psi_k(\mathbf{r}_2,\nu). \tag{3.5-14}$$

In general, the functions $\Psi_k(\mathbf{r},\nu)$ in Eq. (3.5-14) do not form anymore a system of orthogonal functions, contrary to $\psi_k(\rho_0,\nu)$. The loss of orthogonality is due to the integration of the Rayleigh-Sommerfeld diffraction integral. However, since the random amplitudes $\mathscr{A}_k(\nu)$ are non-correlated (Eq. (3.5-2)), the individual contributions $\alpha(\nu)_k\cdot\Psi_k^*(\mathbf{r}_1,\nu)\cdot\Psi_k(\mathbf{r}_2,\nu)$ to the cross-correlation density function $W(\mathbf{r}_1,\mathbf{r}_2)$ in Eq. (3.5-14) are non-correlated too. Therefore, if $W(\mathbf{r}_1,\mathbf{r}_2,\nu)$ is considered as a new secondary source radiating in free space, the principle of superposition with $\mathscr{A}_k(\nu)\Psi_k(\mathbf{r},\nu)$ can still be applied in order to evaluated the cross-spectral density function at a further location.

The great advantage of the decomposition of a random field into deterministic fields becomes evident by comparing Eq. (3.5-9) with Eq. (3.5-12): in Eq. (3.5-12) a four dimensional integral has to be solved to get $W(\mathbf{r}_1,\mathbf{r}_2)$; from Eq. (3.5-9) we may obtain $W(\mathbf{r}_1,\mathbf{r}_2)$ by solving the Rayleigh-Sommerfeld integral which can be evaluated by a Fourier transform (Eq. (2.3-23)). In the case where the number of deterministic fields is reasonably small or can be approximated by a reasonably small number of fields, Eq. (3.5-9) is advantageous for the numerical evaluation of $W(\mathbf{r}_1,\mathbf{r}_2)$.

3.6 Conclusion

The cross-correlation density function is an appropriate tool to describe mathematically any random field of classical sources or waves. The decomposition into its deterministic fields allows treating a partially coherent source by a linear superposition of mathematically orthogonal and/or non-correlated, coherent sub-sources. The radiation of a source can then be evaluated by solving the two-dimensional diffraction integrals of the deterministic fields generated by the sub-sources contributing to the radiation. This procedure has the advantage that the four-dimensional integral of the cross-spectral density function is replaced by solving a limited number of two dimensional Fourier integrals.

References

[3.1] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, New York, 1995), pp. 41-332.

[3.2] J. W. Goodman, *Statistical Optics* (John Wiley and Sons, New York, 1985), pp 7-59.

[3.3] B. E. A. Saleh, M.C. Teich, *Fundamentals of Photonics* (Wiley Series in Pure and Applied Optics, New York, 1994), pp. 342-383.

[3.4] D.W. Robison, G.T. Reid, *Interferogram Analysis* (IOP Publishing 1993), pp. 94-140.

[3.5] H.J. Tiziani, "Optical methods for precision measurements," Opt. and Quantum Electron. **21**, 253-282 (1989).

[3.6] F. Gori *et al.*," Intensity-based modal analysis of partial coherent beams with Hermite-gaussian modes," Opt. Lett. **23,** 989-991 (1998).

[3.7] R. Borghi, M. Santarsiero, "Modal decomposition of partially coherent flat-topped beams produced by multimode lasers," Opt. Lett. **23**, 313-315 (1998).

[3.8] A. Cutolo *et al.*, "Transverse mode analysis of a laser beam by near- and far-field intensity measurements," Appl. Opt. **34**, 7974-7978 (1995).

[3.9] H. Gamo, "Matrix treatment of partial coherence," in *Progress in Optics III*, E. Wolf, ed. (North-Holland, Amsterdam, 1964), pp. 189-315.

[3.10] F. Gori, "Collett-Wolf sources and multimode lasers," Opt. Commun. **34**, 301-305 (1980).

CHAPTER 4

Entropy of radiated scalar fields

4.1 Introduction

The entropy is a well known quantity in thermodynamics. It describes the statistical state of a system and indicates whether and under which circumstances a transition from one state to an other is feasible. The concept of entropy of an electromagnetic field was first introduced by Laue [4.1, 4.2]. Later the concept was used by Gamo for the matrix treatment of partial coherent light [4.3].

The aim of this chapter is to present the concept of entropy of electromagnetic fields. Since this entropy uses the notation of photons emitted by deterministic fields or natural modes of oscillation, we will start with a first section introducing the electromagnetic field and its generation in terms of quantum mechanics. Then, we shall show how the entropy can be related to random fields.

4.2 Radiation from a source in terms of quantum mechanics

Quantum mechanics treats radiating electromagnetic fields as a harmonic oscillator confined inside a cavity [4.4]. In contrast to classical field theory, where the energy of a radiated field inside a cavity can have any positive value, the energy in quantum mechanics can take only the discrete energy levels

$$E_n = \left(n + \frac{1}{2}\right) \cdot h\nu \quad n = 0,1,2,3\ldots \ . \tag{4.2-1}$$

The energy levels E_n are known as eigenstates of the harmonic oscillator. E_0 is called the ground or vacuum state. The energy quantum $h\nu$ is known as photon and depends linearly on the frequency ν and $h = 6.63 \cdot 10^{-34}$Js is Planck's constant. We may differentiate three basic type of oscillators [4.4].

The first type is the number state oscillator. In this type of oscillator, exactly n photons are excited. The uncertainty of the number of photons is $\Delta n = 0$. Thus, the oscillation is in a pure eigenstate. These kinds of oscillation (i.e. fields) are not generated in practice.

Second, there is the single-mode coherent state oscillator. The state is a linear superposition of different number states of the oscillator and is expressed by a single complex number

$$|\alpha\rangle = \exp\left(-\frac{1}{2}|\alpha|^2\right)\cdot\sum_n \frac{\alpha^n}{(n!)^{1/2}}|n\rangle. \tag{4.2-2}$$

The mean number of photons of the oscillator is given by

$$\bar{n} = |\alpha|^2 \tag{4.2-3}$$

with a root-mean-square deviation of

$$\Delta n = |\alpha| = \sqrt{\bar{n}} \tag{4.2-4}$$

photons. The probability distribution function of the photons, i.e. the probability of finding n photons in the state α is given by

$$p(n,\alpha) = |\langle n|\alpha\rangle|^2 = \frac{|\alpha|^{2n}}{n!}\cdot\exp\left(-|\alpha|^2\right), \tag{4.2-5}$$

which is the Poisson probability distribution.

Single-mode coherent state oscillators are the closest quantum mechanical approach to the deterministic field of Chap. 2 or of the natural modes of oscillation of Chap. 3. They have an electromagnetic field variation that, in the limit of high excitation, approaches that of classical waves with stable amplitude and fixed phase. Typically single-mode lasers operating well above threshold generate an electromagnetic field corresponding to the coherent state oscillator.

Finally, there is the statistical mixture state oscillator. Its oscillation consists of a statistical combination of different number states and generates a random field. The photons of the number states are thermally excited in a cavity maintained at the temperature T. In practice this type of oscillation can be generated by filtering a small frequency band $d\nu$ of a chaotic source (e.g. gas discharge, incandescent lamp). The probability, that a single-mode (i.e. a natural mode of oscillation) of a cavity emitting at the frequency ν is in the number state n is given by the Planck's law and yields the probability function of Boltzmann

$$p_\nu(n) = \left\{1 - \exp\left(-\frac{h\nu}{k_B T}\right)\right\} \cdot \exp\left(-\frac{nh\nu}{k_B T}\right), \tag{4.2-6}$$

where k_B is the Boltzmann constant. The average photon number of the single-mode is given by

$$\bar{n}_\nu = \frac{1}{\exp\left(\frac{h\nu}{k_B T}\right) - 1} \tag{4.2-7}$$

with a root-mean-square deviation of

$$\Delta n_\nu = \bar{n}_\nu \tag{4.2-8}$$

if $n >> 1$.

It can be shown that the density of the number of single-modes in the volume dV and the frequency interval $d\nu$ is given by

$$K = \frac{8\pi\nu^2 dV d\nu}{c^3}. \tag{4.2-9}$$

Hence, by multiplication of Eqs. (4.2-7) and (4.2-9) we find for the average energy

$$d\bar{E} = K \cdot (\bar{n}_\nu \cdot h \cdot \nu) = \frac{8\pi\nu^2}{c^3} \cdot \frac{h \cdot \nu}{\exp\left(\frac{h \cdot \nu}{k_B \cdot T}\right) - 1} \cdot dV \cdot d\nu, \tag{4.2-10}$$

which is Planck's radiation formula.

In Chap. 3 we have introduced the spectral degree of coherence and the natural modes of oscillation. For this function we may associate incoherence to the statistical mixture state oscillator and coherence to the single-mode coherent state oscillator. The partial coherence may be considered as superposition of a finite number of uncoupled (non-correlated) single-mode coherent state oscillators. However, the spectral degree of coherence is not able to distinguish between the field of a number state oscillator and of a single-mode coherent state oscillator [4.4]. For both oscillators the absolute value of the spectral degree of coherence is unity. They can only be distinguished with the help of higher order correlation functions. Since we restrict this work to classical sources and fields, the single-mode number state oscillator is not further considered.

4.3 Entropy of a radiated field

The entropy S_E of thermodynamics is a powerful tool to investigate statistically complex systems consisting of microscopic particles [4.5]. It can be interpreted as a quantity of disorder where a high value of S_E indicates a high disorder in the system. The entropy is an extensive quantity; i.e. when a system is composed of two sub-systems with $S_{E,1}$ and $S_{E,2}$, the entropy of the total system is $S_{E,1} + S_{E,2}$. In a closed system, where energy exchange with another system (including environment) is not possible, the entropy may remain constant ($\Delta S_E = 0$) or increases ($\Delta S_E > 0$). This is known as the second law of thermodynamics.

Generally, the entropy of a closed system increases when energy is transformed into heat (lost energy) which is absorbed by an internal heat reservoir at lower and constant temperature (e.g. environment which is included to the closed system). Only in an open system, where an energy exchange with external systems is possible, the entropy may decrease ($\Delta S_E < 0$).

The concept of the entropy of an electromagnetic field was used by Gamo for the matrix treatment of partial coherent light [4.3]. He defined the entropy of a (polarised) electromagnetic field as

$$S_E = -\sum_K p_k \cdot \ln(p_k), \tag{4.3-1}$$

where p_k may be considered as the probability that a detected photon belongs to the k-th single-mode coherent state oscillator. This entropy is zero if the electromagnetic field or source consists of one single-mode coherent state oscillator. Every supplementary single-mode coherent state oscillator will increase the entropy.

It has been mentioned above that the electromagnetic wave emitted by a laser can be considered as a single-mode coherent state oscillator. This would imply that a wave emitted by a laser has zero entropy. However, a natural mode of oscillation of a real laser, although very close, will never be in a single-mode coherent state because of phase fluctuations due to spontaneous emission. The natural mode of a laser resonator is therefore in a mixed single-mode coherent state (a temporally varying complex α) and will not have a non-vanishing entropy due to the lack of information on its phase. Assuming that the phase is statistically uniformly distributed over the range of [0...2π] it can be shown [4.6] that the entropy of the field emitted by a laser is

$$S_{E,l} = |\alpha|^2 - |\alpha|^2 \ln\left(|\alpha|^2\right) + \exp\left(-|\alpha|^2\right) \sum_{n=0}^{\infty} \frac{\ln(n!)}{n!} |\alpha|^{2n}. \tag{4.3-2}$$

If the mean number of photons is large ($\bar{n} >> 1$), Eq. (4.3-2) can be approximated by

$$S_{E,l} = \frac{1}{2} \cdot \ln(\bar{n}). \tag{4.3-3}$$

The entropy of a thermal source can be obtained from the statistical mixture state oscillator in Eq. (4.2-6). It can be evaluated by introducing Eqs. (4.2-6) into Eq. (4.3-1) and making use of the relation in Eq. (4.2-7). This yields [4.6]

$$S_{E,th}(\bar{n}_\nu) = (\bar{n}_\nu + 1) \cdot \ln(\bar{n}_\nu + 1) - \bar{n}_\nu \cdot \ln(\bar{n}_\nu) \tag{4.3-4}$$

and if $\bar{n} >> 1$ we can approximate Eq. (4.3-4) by

$$S_{E,th}(\bar{n}_\nu) = \ln(\bar{n}_\nu). \tag{4.3-5}$$

Comparing Eqs. (4.3-3) and (4.3-5) it can be seen that the entropy of a field emitted by a thermal source is twice that of a laser. This difference may be explained by the mean square deviation of the probability distribution function of the photons of these two sources. The thermal source has a Boltzmann probability distribution function with a mean square deviation (fluctuation of photons) equal to the mean number of photons (Eq. 4.2-8). However, for a laser, where the probability is a Poisson distribution function, the mean square deviation is the square *root* of the mean number of photons (Eq. 4.2-4), resulting in a entropy which is one half of that of a thermal source. If these fields are used as a signal carrier, the signal to noise ratio of the field generated by a thermal source will be lower than that of the field generated by a laser source. Since entropy is an extensive quantity, the entropy S_E of a field generated by many uncoupled sources is the sum

$$S_E = \sum_K S_{E,k} \tag{4.3-6}$$

of the entropy $S_{E,k}$ of each source.

We have introduced the entropy as a function of p_k, where p_k is considered as the probability that the detected photon belongs to the k-th quantum mechanical single-mode coherent state oscillator (Eq. (4.3-1)). We have seen that the quantum mechanical single-mode coherent state oscillator can be associated to the classical field with stable amplitude and fixed phase, which is in fact a deterministic field of Chap. 2 or a natural mode of Chap. 3 in the special case if the deterministic fields of the source form a system of orthogonal functions. Furthermore, it can be shown that the energy flow $\Phi_k(\nu)$ (Eq. (3.3-8)) of the deterministic field k is proportional to the time averaged number of photons [4.4]. Hence, there is a linear relationship between the energy flow $\Phi_k(\nu)$ and the probability $p_k(\nu)$. Normalising $\Phi_k(\nu)$ by the total power flow $\Phi(\nu)$ of the wave we get $\Phi_k(\nu)/\Phi(\nu) = p_k(\nu)$.

The decomposition of random fields into K the deterministic fields has been described in Chap. 3. Since the functions $\psi_k(r, \nu)$ of this decomposition are normalised, the values $\alpha_k(\nu)$ in Eq. (3.5-2) (or in Eq. (3.4-2)) are related directly to $\Phi_k(\nu)$. Thus, by using the linear relationship

$$\alpha_k(\nu) = \gamma(\nu) \cdot p_k(\nu) \tag{4.3-7}$$

with $\gamma(\nu)$ chosen so that

$$\sum_K p_k(\nu) = 1 \tag{4.3-8}$$

and introducing it into Eq. (3.4-2)

$$W(r_1,r_2,\nu)=\gamma(\nu)\cdot\sum_K p_k(\nu)\cdot\psi_k^*(r_1,\nu)\cdot\psi_k(r_2,\nu) \tag{4.3-9}$$

we are able to determine the probabilities $p_k(\nu)$ by the (eigen-)values α_k of the cross-spectral density function. The entropy of the radiated field is then given by

$$S_E(\nu)=-\sum_K p_k(\nu)\cdot\ln(p_k(\nu)) \tag{4.3-10}$$

as a function of frequency. The field described by Eq. (4.3-9) has zero entropy if it consists of only one deterministic field or natural mode ($K = 1$). The entropy is maximum when all K deterministic fields or natural modes radiate the same energy flow.

In Ref. [4.7] the coupling of modes is discussed in terms of the entropy. The resulting field from the coupled modes can be considered as a new unique natural mode, since all values of the fields at any point are correlated with each other. The entropy of the field emitted by the new natural mode is lower than the entropy of the field of the uncoupled natural modes. Because of the second law of entropy, this decrease of entropy must be compensated somehow. For passive coupling the compensation is achieved by loss in form of heat in the saturable absorber, whereas for active coupling the compensation is achieved in form of energy supplied by the modulator.

4.4 Conclusion

In this chapter we have treated the electromagnetic field as photons of a quantum mechanical harmonic oscillator. This oscillator can take only discrete energy states (eigenstates) which are given by the number of photons. Based on the harmonic oscillator we have shown that an electromagnetic field can be described as a linear superposition of the eigenstates of a harmonic oscillator.

The description of the electromagnetic field by photons, i.e. countable microscopic gas-like particles, made it possible to introduce the entropy. By combining the concepts of the deterministic field or natural mode, the quantum mechanical harmonic oscillator and the entropy of statistical mechanics, we were able to introduce the entropy for electromagnetic radiation. In the next chapter we shall show that this entropy gives an important information whether the radiation property of a source or a beam of light can be improved or not.

References

[4.1] M. Laue, "Die Entropie von partiell kohärenten Strahlenbündeln," Ann. d. Phys. **23**, 1-43 (1907); or SPIE Milestoene Series, **19**, part I, 37-79.

[4.2] M. Laue, "Die Entropie von partiell kohärenten Strahlenbündeln; Nachtrag" Ann. d. Phys. **23**, 795-797 (1907); or SPIE Milestoene Series, **19**, part I, 80-82.

[4.3] H. Gamo, "Matrix treatment of partial coherence," in *Progress in Optics III*, E. Wolf, ed. (North-Holland, Amsterdam, 1964), pp. 189-315.

[4.4] R. Loundon, *The Quantum Theory of Light* (Oxford Science Publications, Clarendon Press, Oxford, 1983, 2nd edition), pp. 1-161.

[4.5] K. Huang, *Statistical Mechanics* (John Wiley and Sons, New York, 1987, 2nd edition), pp. 130-134.

[4.6] Th. Graf, J. E. Balmer, "Laser beam quality, entropy and limits of beam shaping," Opt. Commun. **131**, 77-83 (1996).

[4.7] Th. Graf, J. E. Balmer, "Entropy balance of optically pumped cw lasers," Opt. Commun. **148**, 256-260 (1998).

CHAPTER 5

Phase space volume of radiated scalar fields

5.1 Introduction

The phase space volume is another quantity, which describes a scalar field. Principally, it describes the necessary volume for free propagation of a wave. The quantity is of interest for optical system design, where maximum energy flow is of major concern, because it gives a first estimation of the necessary size of the system.

This chapter introduces the phase space and the related M^2-factor by means of the uncertainty principle. In order to avoid solving the integrals when just an estimation is necessary, an approximation for the calculation of the phase space volume is suggested. Finally, the possibility of phase space volume matching and squeezing is discussed.

5.2 Uncertainty principle, phase space and M^2-factor

In chapter 2, we have introduced light as an electromagnetic field. However, as we have mentioned in Chap. 4, quantum theory treats light as particles, which are called photons, having a quantum energy $E = h \cdot \nu$. These photons can be considered as moving microscopic particles in space which are described by their momentum $\boldsymbol{p}$ and location $\boldsymbol{r}$. Although in classical mechanics the two parameters can be determined with any desirable precision, this is impossible in quantum mechanics, since momentum and location are conjugate Fourier pairs [5.1]. The increase of accuracy of one parameter can only obtained at the cost of reduced accuracy of the other parameter. This fact is expressed by the uncertainty principle

$$V_P = \Delta x \cdot \Delta y \cdot \Delta z \cdot \Delta p_x \cdot \Delta p_y \cdot \Delta p_z \geq \left(\frac{h}{4\pi}\right)^3 \qquad (5.2\text{-}1)$$

of quantum mechanics [5.2]. Generally, the product on the left hand side of Eq. (5.2-1) is known as the phase space volume V_P (in a six dimensional space). The uncertainty principle imposes that this volume can never be zero.

Let us find in a first order approximation how the phase space volume, uncertainty and coherence are linked together. For the sake of simplicity we restrict the investigation to a one dimensional wavefront. Using $\Delta p_x = p \cdot \sin(\Delta\theta) \approx (h/\lambda)\cdot\Delta\theta$ and Eq. (2.3-16) we can express the uncertainty principle

$$V_P = \Delta x \cdot \Delta q_x \geq \frac{1}{4\pi} \tag{5.2-2}$$

with the spatial frequency instead of the momentum. The source which we are going to use for the discussion is shown in Fig. 5.2-1.

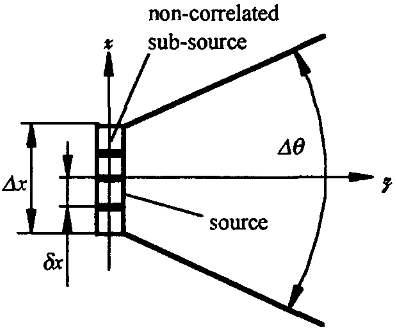


Figure 5.2-1. Model of a source consisting of a number of non-correlated, coherent sub-sources.

The source has a length of Δx, a coherence length of δx and it radiates its energy within an angle of $\Delta\theta$. The coherence length of $\delta x < \Delta x$ corresponds in fact to the extension of one non-correlated sub-source emitting a deterministic field. In Fig. 5.2-1 the number of sub-sources is then given by $m = \Delta x/\delta x$. The phase space volume where all photons are found is about $V_P \approx \Delta x\Delta\theta/\lambda$. From diffraction optics we known that the angle $\Delta\theta$ depends indirect proportionally on the extension of the sub-source, i.e. $\Delta\theta \propto 1/dx$ and hence,

$$V_P \propto \frac{\Delta x}{\delta x}. \tag{5.2-3}$$

This equation is very important, since it shows that with increasing coherence length of the field in the source plane the phase space volume V_p decreases. However, V_p will never become zero as it is given by Eq. (5.2-1) and as it can be shown by the following considerations. Think of the source in Fig. 5.2-1 with m sub-sources. If we cover all sub-sources except one we are able to locate the photon within a region of δx and an angle $\Delta\theta$ yielding a phase space volume of $V_{p,\delta x} = \delta x \Delta\theta/\lambda$. If we want to locate the photon along Δx with higher precision we need also to cover partially the remaining sub-source. But the more we cover this sub-source the grater becomes the angle $\Delta\theta$ because of diffraction. Thus, $V_{p,\delta x}$ remains approximately constant and therefore it is not possible to distinguish the photons within $V_{p,\delta x}$ by decreasing this volume further. We can imagine $V_{p,\delta x}$ as cloud consisting of all photons emitted by a coherent sub-source where the principle of uncertainty does not allow us to get view of its inside. Hence, the photons of a coherent sub-source are intrinsically indistinguishable [5.2].

The quantities Δx and Δq for an arbitrary propagating wave are defined as [5.3]

$$\Delta x(\nu)\cdot\Delta y(\nu)=\sqrt{\frac{\int_{-\infty}^{+\infty}\rho^2\cdot S(\rho-\rho_g,\nu)\cdot d^2\rho}{\int_{-\infty}^{+\infty}S(\rho-\rho_g,\nu)\cdot d^2\rho}} \qquad (5.2\text{-}4)$$

and

$$\Delta q_x(\nu)\cdot\Delta q_y(\nu)=\sqrt{\frac{\int_{-\infty}^{+\infty}q^2\cdot J_S(q-q_g,\nu)\cdot d^2q}{\int_{-\infty}^{+\infty}J_S(q-q_g,\nu)\cdot d^2q}}, \qquad (5.2\text{-}5)$$

where S is the spectral density function with $S(\rho,\nu) = W(\rho,\rho,\nu)$ (Eq. 3.3-4) and J_S the spectral radiant intensity with $J_S(q,\nu) = \hat{W}(q,q,\nu)$ (Eq. 3.3-6). The index g refers to the center of gravity in both equation.

It can be shown [5.4] that for all possible fields with the same degree of coherence, the Gaussian Schell-model (Sec. 3.5.1) yields a minimum for the phase space volume $\Delta x\Delta y\Delta q_x\Delta q_y$. However, the value of the minimum depends on the degree of coherence. It

decreases with increasing coherence and is smallest in the limit where the Gaussian Schell-model describes a (coherent) Gaussian single-mode wave (i.e. TEM_{00}). This global minimum is then $1/4\pi$, which is in agreement with the quantum mechanics of smallest phase space volume. Thus, for the coherent Gaussian beam the phase space volume is smallest.
A quality parameter, the M^2-factor, which describes the deviation of the field from the global minimum of the phase space volume has been introduced by Siegman [5.5]. It is defined by

$$\Delta x \cdot \Delta q_x = \frac{M_x^2}{4\pi} \geq \frac{1}{4\pi} \tag{5.2-6}$$

and similar for the axis y

$$\Delta y \cdot \Delta q_y = \frac{M_y^2}{4\pi} \geq \frac{1}{4\pi} \tag{5.2-7}$$

and therefore

$$\Delta x \cdot \Delta y \cdot \Delta p_x \cdot \Delta p_y = M^4 = M_x^2 \cdot M_y^2 . \tag{5.2-8}$$

For $M^4 = 1$ the field is a coherent Gaussian beam and the smallest possible phase space volume of the field is achieved. If the eigenfunctions of the cross-spectral density function are Gaussian Hermitian ($TEM_{k,l}$-modes), it can be shown [5.5] that

$$M^4 = \frac{\sum_{k,l} \alpha_{k.l}(2k+1)(2l+1)}{\sum_{k,l} \alpha_{k.l}} . \tag{5.2-9}$$

For $k = 0$ and $l = 0$ (TEM_{00}-mode) we get $M^4 = 1$, as expected for a coherent Gaussian beam. Any optical obstacle having sharp discontinuities will generate higher spatial frequencies due to diffraction causing an increase of the phase space volume or of the M^2-factor. It has been shown [5.6] that, if calculated directly, these discontinuities may yield an infinite M^2-factor. However, ignoring the evanescent waves in Eqs. (5.2-4) and (5.2-5) the M^2-factor remains finite.

5.3 First order approximation of phase space volume and M^2-factor

In many situations an approximation can already give an information about the limits of a design. Instead of solving the integrals of Eqs. (5.2-4) and (5.2-5) we may define an estimation of the phase space volume given by

$$V_P = \frac{w \cdot \zeta}{\lambda}, \tag{5.3-1}$$

which is a one-dimensional source with a uniform intensity distribution in its exit pupil of width w and radiating uniformly into the space within an angle ζ. This situation is depicted in Fig. 5.3-1.

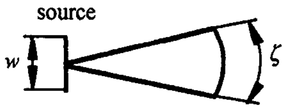


Figure 5.3-1: Approximation of the phase space volume.

The M^2-factor for this approximation shall yield unity when the field has smallest phase space volume, i.e. when the field is a Gaussian beam (TEM$_{00}$). The half angle ϑ of a Gaussian beam is given by [5.7]

$$\vartheta = \frac{\lambda}{\pi \cdot w_0} \tag{5.3-2}$$

where w_0 is the half width. Introducing Eq. (5.3-2) into Eq. (5.3-1) yields

$$V_{P,Gauss} = \frac{2w_0 \cdot 2\vartheta}{\lambda} = \frac{4}{\pi} \tag{5.3-3}$$

and imposing the condition of unity we find for the approximation of the M^2-factor

$$\mathcal{M}^2 = \frac{V_P}{V_{P,Gauss}} = \frac{w \cdot \zeta \cdot \pi}{4 \cdot \lambda}. \tag{5.3-4}$$

We may use the same approximation in order to define the phase space volume of a lens like system. In this case, *w* would be the width of a limiting stop somewhere in front of the lens

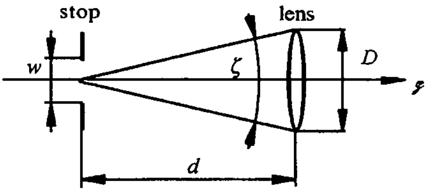


Figure 5.3-2: First order phase space volume of an optical system.

and ζ the angle subtended by the lens and the aperture *D* at the distance *d* from the stop (Fig. 5.3-2). In the paraxial approximation, where $\zeta \approx D/d$, the approximation of the phase space volume is the same if the angle is taken to be subtended by the stop, i.e. $\zeta' = D'/d$

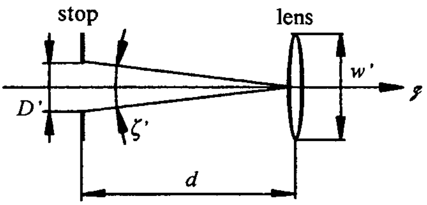


Figure 5.3-3: The phase space volume given in this figure has the same value as in Fig. 5.3-2. Therefore, the definition of *D* (from which the angle $\zeta = D/d$ is evaluated) and *w* may be interchanged.

(Fig. 5.3-3), and the width *w* is given by the aperture of the lens. Hence, we can write

$$V_P = \frac{4 \cdot \mathcal{M}^2}{\pi} = \frac{w \cdot \zeta}{\lambda} = \frac{\zeta' \cdot w'}{\lambda}. \tag{5.3-5}$$

In a cascade of stops and lenses it is evident that in this approximation the phase space volume of the total optical system is given by the smallest value of the individual phase space volume in the cascade. In a loss less imaging system (phase space volume of the optical system is greater than the phase space volume of the object) the demagnification of angle ζ equals the magnification β of the object size. Therefore, we can conclude from Eq. (5.3-1) that the phase space volume in the defined approximation remains constant in a loss less optical system. This is equal to saying that the radiance of a loss less optical system remains constant.

For an optical design where in a first order approximation the effects of diffraction can be neglected, the Eq. (5.3-1) with the imaging equation of the geometrical optics will give a first estimation of the necessary volume of the optical system (Fig. 5.3-5). The imaging equation will define the length z_1 and Eq. (5.3-1) the necessary cross-sectional diameter of the optical system i.e. $z_1\zeta$.

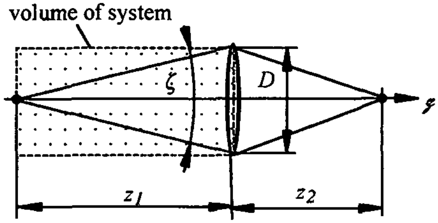


Figure 5.3-4: Volume estimation of an optical system.

5.4 Matching the phase space

To avoid losses, the optical system must have at least the phase space volume of the source. In the case of rotational symmetric optical system, i.e. phase space volume of the two axes are $M_x = M_y$ (or $\mathcal{M}_x = \mathcal{M}_y$), with a source whose phase space volume is not rotational symmetric, i.e. $M_x \neq M_y$ (or $\mathcal{M}_x \neq \mathcal{M}_y$), the design must be based for the larger phase space volume of the two axes. This may result in a design where most of the volume of the system is not used of the propagating field. In such a case it may be tried to match the phase space volume of the source to the desired requirements of the optical system. Matching is possible if the non-correlated, coherent sub-sources can be separated in the source plane or somewhere else (e.g. Fourier plane). It is carried out by rearranging the distribution of the sub-sources in the plane (eventually an imaging step is necessary) with the aim that the phase space volume of the source fits better the phase space volume of the optical system (Fig. 5.4-1). Since the matching consists principally of shifting the sub-sources in a plane it does not change the entropy of the field and hence, it can be realized by passive optical elements. A typical case for phase space matching is the pulsed laser diode which can be considered as a line source radiating within angles $\zeta_x \approx \zeta_y$ but having an extension of $w_x >> w_y$ (or $w_y >> w_x$). This source can be made more suitable for rotational symmetric optical systems if it is folded so that $w_x' \approx w_y'$ yielding $\mathcal{M}_x^2 \approx \mathcal{M}_y^2$ (Fig. 5.4-1).

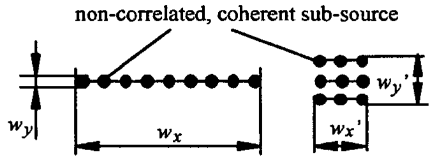


Figure 5.4-1: Original line source (left) and folded line source (right) to match better the phase space of a rotational symmetric optical system.

5.5 Squeezing phase space volume

In this section we show how the phase space volume, as defined by the Eqs. (5.2-4) ... (5.2-6) can be reduced or squeezed, yielding a smaller M^2-factor, which is important in optical pointing systems, where the optical system and the divergence angle of the emitted beam have to be small (small $\Delta x\Delta y$ and $\Delta q_x\Delta q_y$). An example of squeezing the phase space volume is shown in Fig. 5.5-1 with a laser array consisting of three *correlated* sub-sources (Fig. 5.5-1) each having a uniform intensity distribution in the source plane and emitting at a wavelength of 633nm. If the phase shift between the sub-sources is compensated by a phase shift plate (on the right hand side of Fig. 5.5-1) the energy flow in the far field becomes concentrated about the zero point. A numerical calculation shows that the phase

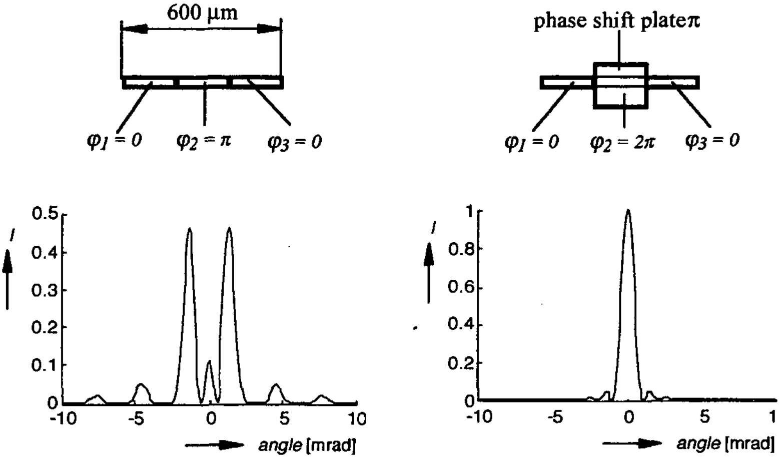


Figure 5.5-1: Left: Description of the original line source (top) and resulting intensity distribution of the far field (bottom). Right: Phase shift compensation between the sub-sources (top) yields an intensity distribution in the far field which is restricted to a smaller region (bottom).

space volume of the source has been halved. Hence, the M^2-factor has been decreased by a factor of 2. The general case of N sub-sources is treated in [5.8, 5.9]. This method of phase shifting between the sub-sources has been widely used in the antenna techniques of telecommunication and radio astronomy in order to achieve a desired directivity of emitted or received signal [5.10]. A further example has been mentioned in [5.11] where the phase space volume is squeezed by transforming a TEM_{10} mode into a TEM_{00}. This transformation is achieved by evanescent coupling of a TEM_{10} waveguide (input) with a TEM_{00} waveguide (output). An incident TEM_{10} wave would then leave the waveguide system as TEM_{00} wave. According to the Eq. (5.2-9) the phase space volume of the field is reduced by a factor of 3. These two examples lead us to the question whether the phase space of any field can be squeezed. A general answer is found by using the entropy introduced in the last chapter: in the case where only passive optical elements (no supplementary energy source) are used any modification of the field is possible as long as the entropy is not decreased.

In the two preceding examples the fields before and after the volume squeezing consist of a single deterministic field and hence, following to Eq. (4.3-7) and (4.3-10), the entropy has not been decreased. In the case of a general monochromatic random field we may try to get access to the individual deterministic fields and transform them into lower phase space volumes. Furthermore, we could think of redistributing the energy flow between the deterministic fields. However, if the probabilities satisfy the inequality

$$p_0(v) \geq p_1(v) \geq p(v)_2 \ldots \geq p(v)_K, \tag{5.5-1}$$

where the numbering is chosen so that their corresponding M^2-factor of the deterministic fields are ordered fulfilling

$$M_0^2 \leq M_1^2 \leq M_2^2 \ldots M_K^2, \tag{5.5-2}$$

the phase space volume of the random field cannot be further reduced by redistributing the energy flow. This can be easily seen with Eqs. (4.3-7) and (4.3-10) in case where the deterministic fields are modes. In order to decrease the phase space volume (Eqs. (5.2-6) and (5.2-7)) the probability p_k or the eigenvalue α_k of a higher mode has to be reduced by the same amount as a lower order is increased (conservation of energy flow). However, if the inequality in Eq. (5.5-1) is satisfied, this modification would result in a lower entropy, since the probability that a photon is generated by the modified lower mode has been further increased at the cost of the modified higher mode. Thus, the entropy of the squeezed field would be smaller, which is in contradiction to the fact that passive optical elements cannot decrease the entropy. An example of a source which satisfy Eq. (5.5-1) with the condition in Eq. (5.5-2) is the Gaussian Schell-model. In the Fig. 3.5-1 in Sec. 3.5.2 it can be seen that the eigenvalues $\alpha_\kappa(v)$ monotonely decrease with increasing mode number k. Following Eq. (4.3-7) $\alpha_k(v) = \gamma(v) p_k(v)$, Eq. (5.5-1) must be therefore satisfied and consequently any change of the intensity distribution yields an increased phase space volume of the source. This is in agreement with the fact that the Gaussian Schell-model source has the smallest phase space volume of all possible fields having the same degree of coherence.

If the M^2-factors and the probabilities of a real existing source are not known we may obtain them by measuring the deterministic fields using the measurement methods listed as references in Chap. 3. Then, we may evaluate the probabilities $p(\nu)_k$ by Eqs. (4.3-7) and (4.3-8) and the M^2-factors by Eqs. (5.2-4...5.2-8).

Even though the above analysis yields that the phase space volume can be squeezed, it might not be technically possible to achieve a reduction. This is the case when the deterministic fields cannot be accessed separately in order to perform the modification. We can distinguish between three different cases where the deterministic fields can be accessed separately.

1) The emitted field consists of only two non-correlated deterministic fields with orthogonal polarizations. In this case the fields can be separated by a polarisator, then transformed to fields of smaller or smallest phase space volume and finally recombined by a second polarisator. Evidently, the separation of the fields by a polarisator is not necessary if the fields are emitted at different locations.

2) The emitted random field consists of a few non-correlated deterministic fields, which have different wavelengths and can be separated by dichroic mirrors. This case is very similar to the previous one. The fields are separated by dichroic mirrors, then transformed to fields of smaller or smallest phase space volume and finally recombined by dichroic mirrors. The separation of the fields by dichroic mirrors is not necessary if the fields are emitted at different locations.

3) The random field consists of a few non-correlated deterministic fields whose wavelengths are so close that they cannot be separated by dichroic mirrors. However, in the source plane (or somewhere else, e.g. Fourier plane), they are locally well separated. In this case we transform the fields directly where they are well separated and redistribute their energy flow by beamsplitters and prisms so that the M^2-factor of the entire source is a minimum.

Additionally, phase space squeezing can be carried out, at least theoretically, by coupling the different non-correlated, deterministic fields and so increasing the coherence length of the entire source (see Sec. 5.2). However, due to the laws of entropy, Sec. 4.3, the coupling of non-correlated, deterministic fields can only be performed by an active optical element (i.e. element needs an external energy source) or by a passive optical element at the cost of loss in energy flow (loss in form of heat radiation at coupling element).

5.6 Conclusion

We have introduced a phase space concept for radiated fields. We have shown that the size of the phase space volume can be characterized by the M^2-factor. It is unity (smallest possible phase space volume), if the radiated field is a Gaussian beam and it increases with increasing phase space volume of the field. An approximation of the phase space volume and the M^2-factor has been introduced and combined with the geometrical image equations. The result is a simple mathematical tool to get a first order estimation of the necessary volume of an optical system.

We have shown the possibility of phase space transformation by phase space matching and volume squeezing. We have seen that matching and squeezing are only technically possible if the deterministic fields of the radiated field can be locally separated. Furthermore, we have shown that squeezing is possible by transforming the individual deterministic fields to smaller phase space volume or by redistribution of the energy flow between these deterministic fields. However, the transformation of an individual deterministic field into a smaller phase space volume is only possible if the aimed phase space volume at the same location is not occupied by another field radiating at higher energy flow. Hence, a smaller phase space volume by redistribution of the energy flow between the deterministic fields is only achieved if the energy flow of the individual deterministic field *does not* decreases monotonely with increasing M^2-factor.

References

[5.1] C. Cohen-Tannoudji, B. Diu, F. Laloë, *Méchanique Quantique I* (Collection ensignement des sciences Hermann, Editeurs des Sciences et des Arts, Paris, 1995), pp. 213-255.

[5.2] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, New York, 1995), pp. 147-159.

[5.3] M. J. Bastiaans, "New class of uncertainty relations for partially coherent light," J. Opt. Soc. Am. A. **1**, 711-715 (1984).

[5.4] M. J. Bastiaans, "Uncertainty principal and informational entropy for partially coherent light," J. Opt. Soc. Am. A. **3**, 1243-1246 (1986).

[5.5] A. E. Siegman, "New developments in laser resonators," SPIE **1224,** Optical Resonators, 2-14 (1990).

[5.6] P. Bölanger *et al.*, "Paré Beam propagation factor of diffracted laser beams," Opt. Commun. **105**, 233-242 (1994).

[5.7] T. Johnston, "M^2 concept characterises beam quality," Laser Focus World, May, 173-183 (1990).

[5.8] P. Ehbets *et al.*, "Beam shaping of high-power laser diode array by continuous surface-relief elements," J. Mod. Opt. **40**, 637-645 (1993).

[5.9] J. Morel, *Study of Coherent Coupling of Fiber Lasers by Use of an Intracavity Phase Grating*, (Thesis, Institute of Microtechnology, Unversity of Neuchâtel, Switzerland, 1994).

[5.10] R. Colin, *Antennas and Radiowave Propagation* (MacGraw-Hill Book Company, New York, 1987), pp. 87-163.

[5.11] Th. Graf, J. E. Balmer, "Laser beam quality, entropy and limits of beam shaping," Opt. Commun. **131**, 77-83 (1996).

CHAPTER 6

Combining scalar fields

6.1 Introduction

There may be cases where a single source does not fulfil the requirements of the specified minimum of emitted energy flow. In such a situation it will be necessary to combine the field of several sources in order to fulfil the specification. Therefore, we shall discuss in this chapter the possibilities of combining fields. We shall restrict the discussion to combiner elements which do not decrease the entropy in the combined field and which do not need be driven by an energy source (loss less and passive combining). We have to distinguish between combining of coupled and uncoupled sources.

6.2 Combining fields of coupled sources

We shall limit us to coupled coherent sources radiating at the same wavelength. However, there is no restriction for the polarization and intensity distribution of the emitted field of the individual sources. Due to the coupling, the individual emitted fields of the sources are correlated with each other. The total field which is radiated by all sources into the space can therefore be considered as an extended deterministic field. The combining will not change this property of the field. Combining just modifies the field distribution of this extended deterministic field. The entropy of the individual fields or of the extended field and the entropy of the combined field is the same. Thus, the combining of coupled sources can be carried out by passive optical elements without losses, neglecting losses due to absorption, scattering or imperfection of the combiner elements.

We shall show four different types of combining: 1) combining by polarization, 2) combining by interference, 3) combining by evanescent coupled waveguides or fibers, 4) combining by pupil filling.

1) For the combining by polarization let us assume that radiated field of the coupled sources is linear polarized and deterministic or that the polarization has been transformed into a linear one. The combining can then be carried out by polarizing beamsplitters. Two incident waves on the polarizing beamsplitter have their linear polarization perpendicular to each other. At the output of the polarizing beamsplitter the wave is the sum of the two input waves. The polarization of the output wave depends on the phase shift between the two input waves. At phase shift zero the output wave is linearly polarized at 45 degree. If the phase shift of the two waves is 90 degrees wave at the output is circularly polarized. Nevertheless, for the next step of combining the polarization of the wave must be transformed into a linear polarization and orthogonal to the next wave to be combined. This transformation can be carried out by passive optical elements, i.e. by polarizers. Because the entropy is not changed, the procedure of combining and polarization transformation can be repeated in a cascade of polarizing beamsplitters and polarization transformers to combine the desired number of sources. Great care must be taken in the mechanical stability of this combiner. Small displacements of the optical elements due to mechanical stress or temperature variations, even in fractions of wavelength, will affect the polarization in each stage of the cascade resulting in loss of energy flow at the polarization transformers.

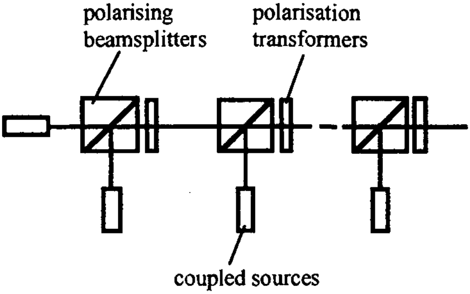


Figure 6.2-1: Combining fields by polarization of coupled sources.

2) An alternative to the previous combiner is obtained when the polarizing layer in the beamsplitters is replaced by a semi-transparent mirror splitting the wave into two parts. This combiner is shown in Fig. 6.2-2. The beamsplitters have a transmission of $T = |\tau|^2$ and a reflectivity of $R = |\rho|^2$ where τ and ρ are complex and $R+T = 1$. At the output 3 the field is given by $U_3 = U_{31} + U_{32}$ with $U_{32} = \rho U_1$ and $U_{31} = \tau U_2$. The field U_3 becomes zero if $U_{31} = -U_{32}$. This is achieved introducing the necessary phase shift and using a semi-transparent mirror with a transmission of $T = |U_1|^2/(|U_1|^2+|U_2|^2)$. In this case the energy flow at output 4 must be equal to the sum of the energy flows of the two input waves and therefore, the two input waves are combined at the output 4 of the beamsplitter.

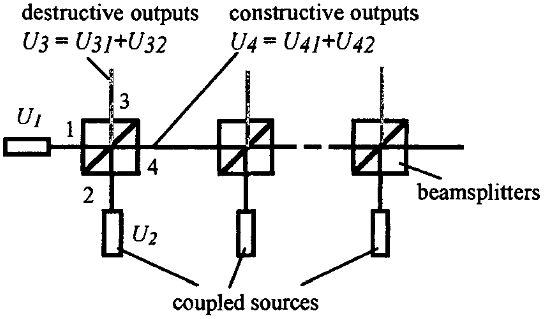


Figure 6.2-2: Combining fields by interference of coupled sources.

As for the previous combiner, great care must be taken in the mechanical stability. Small displacements of the optical elements due to mechanical stress or temperature variations, even in fractions of wavelength, prevents a full suppression of the field at output 3 ($U_{31} \neq -U_{32}$), so that complete combination of the two input waves is not achieved.

3) A different kind is the combining by evanescent wave coupling between fibers or waveguides. If two waveguides are sufficiently close, so that their evanescent fields overlap, light can be coupled from one to the other as the waves propagate. If the field in each waveguide propagates at the same velocity a coupling efficiency of 100% can theoretically be achieved (phase matching). However, the efficiency is less than 100% if the propagation velocities of the fields in the two waveguides differ [6.1].

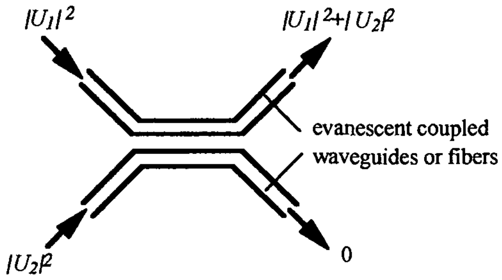


Figure 6.2-3: Combining fields by evanescent coupling.

For this combiner as well, great care must be taken for the mechanical stability. Variations in the propagation distances even in fractions of wavelength and changes of the relative phases of the input fields will reduce the coupling efficiency.

4) A further possibility is source combining by pupil filling. Pupil filling simply consists of

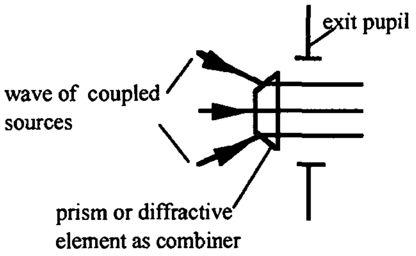


Figure 6.2-4: Combining fields by pupil filling of coupled sources. Shown example is with prism as combiner (mirrors as combiner see Fig. 6.3-5).

juxtaposing the fields of the different sources. However, the number of possible sources to be combined is limited by the size of the aperture in the plane. Small variations in the propagation distances between the individual sources and exit pupil will modify the divergence angle of the combined beam (cf. Fig. 5.5-1). Therefore, high mechanical stability of this combiner is needed too.

We have seen that the combining of coupled sources imposes a mechanical tolerance range which is smaller than the emitting wavelength. Therefore, great care must be taken for mechanical stress and temperature variations in the design of such combiners. Common work shop manufacturing will never achieve the required precision. Therefore, newer manufacturing procedures such as micro- or nano-technology are necessary to produce these combiners.

6.3 Combining fields of uncoupled sources

The case of combining fields of uncoupled sources is more complex than of coupled sources. In the previous section we could base the discussion on the fact that combining will conserve the entropy of the field. This is different for combining uncoupled sources. For each kind of proposed combining we have to find out how the combining affects the entropy and then decide whether a realization with passive elements is possible or not according to the change of entropy in the combined field.
Let us define for this section the straight line which intersects with the center of gravities of the intensity distribution in the near and far field of a propagating wave as the axis of propagation. Then, we may distinguish two main groups of combining uncoupled sources: First combining by superposition of the fields (Fig. 6.3-1, left). In this case the axis of propagation $\mathfrak{z}_1$,$\mathfrak{z}_2$... of the individual fields coincides with the axis of propagation of the combined field, $\mathfrak{z}_0$. Second, combining by juxtaposing of the fields (in the near or far field). In this case the axes of propagation of the individual fields $\mathfrak{z}_1$,$\mathfrak{z}_2$... do not coincide with the axis of propagation of the combined field, $\mathfrak{z}_0$. Combining by juxtaposing yields a larger phase space volume than the largest phase space volume of the individual fields. Furthermore, combining by juxtaposing results in a larger the phase space volume than combining by superposition

(this can be seen intuitively in Fig. 6.3-1 but can be proven as well by Eqs. (5.2-4)...(5.2-7)). The same is true for the M^2, since M^2 and phase space volume are related by the Eqs. (5.2-6) and (5.2-7).

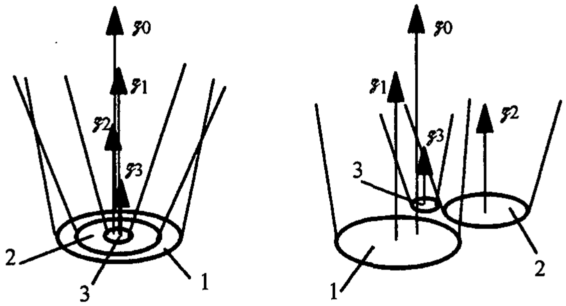


Figure 6.3-1: Combining of fields of uncoupled sources. Left superposition of the fields, right juxtaposing of the fields. The axis of propagation of the combined field is z_0.

In order to carry out the combining with loss less and passive optical elements, the entropy of the combined beam must be at least the sum of the entropies of the individual fields at the same wavelength (entropy is an extensive quantity and is a function of wavelength). Based on this fact let us find the possibilities of combining.

6.3.1 Combining by superposition

We begin with the assumption that all fields have the same wavelength and are unpolarized. In this case, the combining by superposition would yield a new field in which the individual fields could not be distinguished anymore, i.e. all photons seem to be emitted by a unique source. According to the laws of thermodynamics, the entropy of such a field would be smaller than the sum of the entropies of the individual fields. Hence, we may conclude that under these conditions the combining is not possible without loss in the energy flow. A combiner of this type is shown in Fig. 6.3-2. It is realized by injecting the field of the sources into a rare-earth-doped silica fiber (pumping a fiber laser) [6.1]. The field leaving the fiber is coherent and thus its entropy is lower than that of the field at the input. If N lasers, emitting the power P_{in} at the wavelength λ_{in}, are combined, the maximum output power P_{out} at the

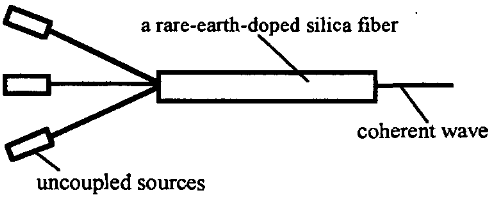


Figure 6.3-2: Combining by injecting the field of the sources into a rare-earth-doped silica fiber.

wavelength λ_{out} of the combined field is [6.1, 6.2]

$$P_{out} = \left(\frac{\lambda_{in}}{\lambda_{out}} \cdot \sum_{N} P_{in} \right) - L_e - L_c$$

where L_e compensates the decrease of entropy of the field at the output of the fiber in form of heat radiation and where L_c takes the losses due the coupling of the field into the fiber into account. In case where the sources are uncoupled lasers the value of L_e has been shown as neglecting small [6.2] (about 10^{-6}% of the total power). This is due to the fact that the pumping lasers have already low entropy. However, L_c can be relatively high, especially when the pumped fiber is a single mode fiber, because only one mode from the emitted field is coupled into the fiber. The loss L_c can be decreased by using double-clad fibers [6.3]. The active fiber core is surrounded by a second waveguide for the pump light. This second waveguide is multi-mode, which results in a lower L_c. A relative L_c of less than 55% has been achieved experimentally [6.3]. The maximum number of lasers, which can be combined, is limited by the available phase space volume of the fiber.

Next, let us suppose that the fields are polarized and that we combine two individual fields. In the combined field we are still able to distinguish the photons of the individual fields. The separation of the photons can be carried out by a polarizer. Therefore, the entropy in the combined field corresponds to the sum of the entropy of the individual fields and combining should be possible by passive optical elements. A combiner which exploits the polarization of the fields is shown in Fig. 6.3-3. Combining two fields is carried out by a polarizing

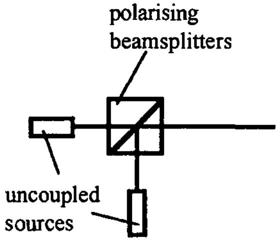
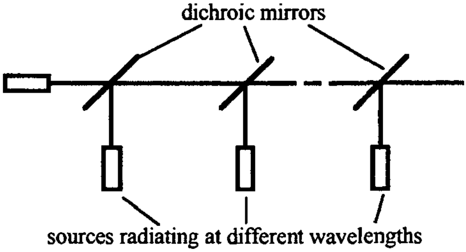


Figure 6.3-3: Combining fields by polarization of uncoupled sources. Only the fields of only two sources can be combined.

beamsplitter. If we would like to cascade the combining, as it has been done for coupled sources, we would need first to transform the leaving unpolarized field of the first stage of the cascade into a linearly polarized field. This transformation is necessary to perform the combination in the next stage of the cascade. However, after such a transformation the photons cannot be distinguished anymore according to their origin source. They seem to be generated by a single source, which shows that this transformation decrease the entropy of the field. Hence, in order to carry out the transformation an active optical element is necessary. Consequently, combining is restricted to two individual fields.

Last let us assume that the individual fields have different wavelengths. The entropy has been introduced as function of wavelength and therefore, as long as the individual fields have different wavelengths, combining will not change the entropy at these wavelengths. Consequently, combining field of different wavelengths can be performed by a passive element. The combining can be realized by wavelength sensitive optical elements such as dichroic mirrors (using wavelength sensitive reflecting layers), or diffractive and refractive elements (using the dispersion of the material), as shown in Fig. 6.3-4. As long as each source radiates at a different wavelength which are separated at least by their line widths, there is theoretically no limit for the number of sources to be combined. However, technically, the limits are given by available wavelength range of the sources and maximum spectral selectivity of dichroic mirrors or wavelength multiplexers [6.4].



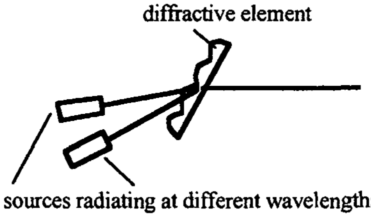


Figure 6.3-4: Combining fields of different wavelengths from uncoupled sources. Top: combining with dichroic mirrors; bottom: combining with diffractive element (grating).

6.3.2 Combining by juxtaposition

The principle of combining by juxtaposition is shown in Fig. (6.3-5). Because the combined field can be locally decomposed into the individual fields (e.g. image plane of the combiner element) the entropy of the combined field corresponds to the sum of the entropies of

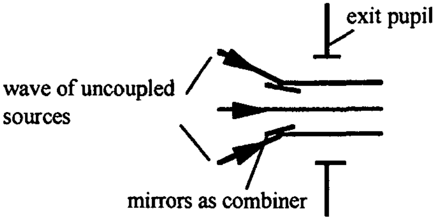


Figure 6.3-5: Combining by pupil filling fields of uncoupled sources. The combiner element consists in this example of mirrors (for prism as combiner elements see Fig. 6.2-4).

individual fields. Consequently, combining can be performed with passive elements. In principle, combining could be carried out just by juxtaposing the housing of the different sources (Fig. 6.3-6, top). However, in most cases the size of the housing is large compared to the effective emitting area of the source, which causes a waste of phase space volume. Squeezing of the phase space volume can be achieved if the radiating surfaces of the sources are juxtaposed in a common imaging plane (Fig. 6.3-6, bottom). Assuming that all sources emit a field of equal phase space volume, the ratio of phase space volume of the optical system and of an individual source gives the maximum number of sources which can be combined.

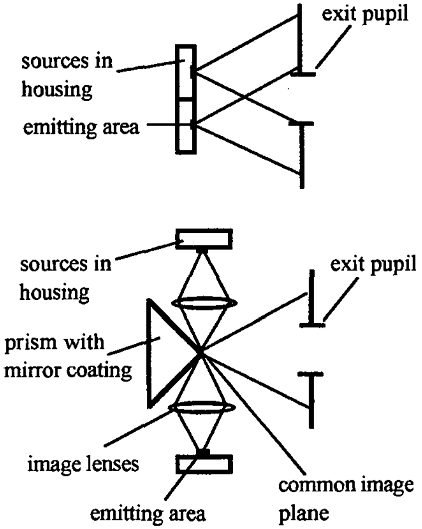


Figure 6.3-6: Combining fields by pupil filling of uncoupled sources. Top: Combining fields by juxtaposing the housing of the sources results in a large phase space volume. Bottom: Combining fields of two sources in a common image plane reduces the phase space volume significantly and hence, more energy will flow through the exit pupil.

In contrast to the combining of coupled sources, there is no high demand for mechanical tolerances of the combiner, because the relative phase between the uncoupled sources do not affect the efficiency of combining.

6.4 Conclusion

We have seen that the fields of sources can be combined without losses and by passive optical elements in order to achieve a higher energy flow at the exit pupil of an optical system. For coupled sources we have found that combining can be considered as a redistribution of the field between the correlated individual sources, which conserves the entropy. The situation is different for uncoupled sources. Combining without losses by passive optical elements is only possible if the entropy is not decreased. Under this condition, we have found that for uncoupled sources the individual fields can be combined if they have orthogonal polarizations (restricted to two fields) or different wavelengths, or if they can be juxtaposed (pupil filling). Juxtaposing yields a phase space volume of the combined field which is larger than the largest phase space volume of the individual fields.

The mechanical tolerances of the combiner of coupled sources must be within the fraction of the wavelength, since the efficiency of the combining depends on the relative phases between the individual fields. Therefore, common work shop tolerances of >0.01 mm are not sufficient to manufacture such type of combiners. Novel technologies used in micro- or nano-mechanics and optics are necessary to manufacture these combiners. The situation is different for the combining of uncoupled sources, where the efficiency of the combining does not depend on the relative phases. The tolerances are given by the mechanical necessities of the design of the combiner, which are usually in the range of 0.01 mm and 0.1 mm.

References

[6.1] B. E. A. Saleh, M.C. Teich, *Fundamentals of Photonics* (Wiley Series in Pure and Applied Optics, New York, 1994), pp. 264-269, 477-480.

[6.2] Th. Graf, J. E. Balmer, "Entropy balance of optically pumped cw lasers," Opt. Commun. **148**, 77-83 (1998).

[6.3] H. Zellmer *et al.*, "Fiber lasers – compact laser light sources for the near infrared spectral range," Laser und Optoelektronik, **29**, 53-59 (1997).

[6.4] G. S. Mecherle, “Laser diode combining for free space optical communication,”SPIE, **16**, Optical Technologies for Communication Satellite Applications 281-291 (1986).

CHAPTER 7

Evaluating numerically the Fresnel diffraction integrals

7.1 Introduction

In cases where the degree of coherence of a source becomes significant, ray optics may not be the appropriate method to evaluate intensity distribution in the transverse direction of an optical system. This is especially the case when the cross-sectional intensity distribution of a coherent (collimated) beam has to be evaluated. Because of diffraction, the cross-sectional intensity distribution will vary with distance. Since geometrical optics is based on the assumption that diffraction can be neglected, these variations cannot be calculated by ray tracing. Hence, the need to use the diffraction integrals introduced in Chap. 2 becomes evident. Unfortunately, all diffraction integrals are two dimensional integrals (two dimensional wavefronts are considered), which results in long calculation time for numerical evaluation unless the discrete Fourier transform (DFT) and the fast algorithms (FFT) are exploited. Fortunately, most of the integrals are Fourier integrals or convolutions. Convolutions can be evaluated by Fourier transforms.

The power of the FFT can be shown by an example. Evaluating the Fresnel diffraction integral of a function described by 512x512 equi-spaced sample points by FFT takes about 7000 times less time than evaluating it by numerical integration (on MATLAB using the functions ‘fft2’ for Fourier transform and ‘sum’ for numerical integration). Furthermore, due to the Shannon theorem of bounded functions the integral is not only known at the equi-spaced sampling points, but can be evaluated as well between them when the DTF of the (sampled) function is known. However, there is an inherent problem of the DTF, which is known as the sampling theorem. This theorem imposes a minimum number of sampling points of the wavefront in order to get the correct result. This number can easily reach millions of necessary sampling points and so, can exceed the available memory of any desktop computer. For the Fresnel diffraction integrals this is just the case if the equations are implemented as they appear in Sec. 2. Using a coordinate transformation and the properties of Fourier transforms an algorithm can be found for which the Fresnel diffraction integral can be evaluated independently of the size of the available memory of the computer.

7.2 Notion of the numerical Fourier transform

Applying the comb-function [7.1] to the Fourier transform integrals in Eqs. (2.3-17) and (2.3-18), the discrete Fourier transform (DFT)

$$\hat{u}(p_x,p_y)=\frac{1}{M\cdot N}\sum_{x=-\frac{N}{2}\cdot\Delta x}^{+\frac{N-1}{2}\cdot\Delta x}\ \sum_{y=-\frac{M}{2}\cdot\Delta y}^{+\frac{M-1}{2}\cdot\Delta y} u(x,y)\cdot\exp\left(-i2\pi\cdot\left(\frac{p_x x}{N}+\frac{p_y y}{M}\right)\right)\equiv DTF\{u(x,y)\} \quad (7.2\text{-}1)$$

and its inverse transform (DTF^{-1})

$$u(x,y)=\sum_{p_x=-\frac{N}{2}\cdot\Delta p_x}^{+\frac{N-1}{2}\cdot\Delta p_x}\ \sum_{p_y=-\frac{M}{2}\cdot\Delta p_y}^{+\frac{M-1}{2}\cdot\Delta p_y} \hat{u}(p_x,p_y)\cdot\exp\left(i2\pi\cdot\left(\frac{p_x x}{N}+\frac{p_y y}{M}\right)\right)\equiv DTF^{-1}\{u(p_x,p_y)\} \quad (7.2\text{-}2)$$

can be developed for a sampled function $u(x,y)$ which is represented by a NxM array of N equi-spaced samples in x and M equi-spaced samples in y. The distance between two

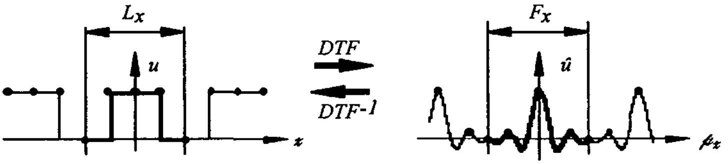


Figure 7.2-1: The discrete Fourier transform in the one-dimensional case.

sampling points is Δx and Δy in the x and y directions of the space domain, and Δp_x and Δp_y in the p_x and p_y directions of the frequency domain. Because the Fourier transform conserves the number of sampling points, the side lengths of the array in the space domain are

$$L_x = N \cdot \Delta x \quad L_y = M \cdot \Delta y \tag{7.2-3}$$

and in the frequency domain

$$F_x = N \cdot \Delta p_x \quad F_y = M \cdot \Delta p_y . \tag{7.2-4}$$

Furthermore, from the Fourier pair of the comb function [7.1] we can straightforward deduce the relations

$$\Delta x = \frac{1}{F_x} = \frac{1}{N \cdot \Delta p_x} \quad \Delta y = \frac{1}{F_y} = \frac{1}{N \cdot \Delta p_y} . \tag{7.2-5}$$

The DFT is cyclic having a period length of N and M samples, i.e. $u(x,y) = u(\Delta x(n+\alpha N), \Delta y(m+\beta M))$, and $\hat{u}(p_x,p_y) = \hat{u}(\Delta p_x(n+\alpha N), \Delta p_y(m+\beta M))$, where m, n, α and β are integers in the interval -∞…+∞. Hence, the DTF in Eq. (7.2-1) yields for a sampled function $u(x,y)$ a sampled spectrum function $\hat{u}(p_x,p_y)$ which is periodically repeated, and vice versa for the inverse transform. An undersampling (too small number of sampling points) of $u(x,y)$ causes aliasing (Fig. 7.2-2). The spectrum $\hat{u}(p_x,p_y)$ extends over the side length F_x (and F_y) of the array and the angular spectrum of one array will penetrate into the neighboring arrays of the repeated angular spectrum and vice versa. As seen from Fig. 7.2-2 aliasing is avoided if the

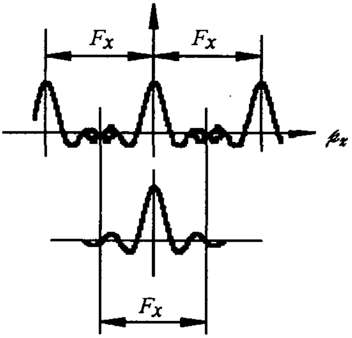


Figure 7.2-2: The figure shows a spectrum, which is larger than the array length F_x. The higher frequencies are perturbed by the neighboring repeated spectrum.

side lengths of the spectrum $\hat{u}(p_x,p_y)$ with maximum frequencies $p_{x,max}$ and $p_{y,max}$ satisfy the inequalities

$$F_x > |2p_{x,\max}| \quad F_y > |2p_{y,\max}|, \tag{7.2-6}$$

which is known as the sampling theorem with F as the sampling frequency.
Introducing Eq. (7.2-6) into Eq. (7.2-5), the condition for the distance between two sampling points becomes

$$\Delta x < \frac{1}{|2p_{x,\max}|} \quad \Delta y < \frac{1}{|2p_{y,\max}|}. \tag{7.2-7}$$

With similar reasoning we obtain the conditions

$$\Delta p_x < \frac{1}{|2x_{\max}|} \quad \Delta p_y < \frac{1}{|2y_{\max}|} \tag{7.2-8}$$

for the distance between two sampling points for the inverse Fourier transform. All theorems of the continuous Fourier transforms (e.g. the convolution theorem) are also valid for the DTF.
A fast algorithm, known as fast Fourier transform and commonly abbreviated by FFT, exists for the numerical evaluation of the DFT and its inverse for the case where the number of sample points (M, N) are equal to 2^n, where n is a positive integer [7.2].

7.3 Minimum number of sampling points for the Fresnel diffraction integral

Let us derive an equation to estimate the necessary number of sampling points for the two Fresnel diffraction integrals given in Eqs. (2.3-19) and (2.3-25). The maximum spatial frequency q_{max} of the wavefront can be estimated by the Debye approximation [7.3, 7.4]. In this approximation it is assumed that a uniform spherical wavefront has no spherical aberration and that the diffraction caused by the limiting aperture can be neglected. Under these conditions it can be shown that the wavevectors of the angular spectrum, which are perpendicular to the wavefront, meet in the focal point (geometrical optics of a converging spherical wavefront). Therefore, the maximum frequency in the Debye approximation is given by the maximum slope of the phase function.

Let us assume, that the spherical wavefront U_0 with uniform intensity distribution in Eq. (2.3-19) has a curvature radius of R (Fig. 7.3-1). Then, the resulting spherical phase term in the argument of the Fourier operator is given by the paraxial approximation

$$\varphi(x_0, z) = \frac{k \cdot x_0^2}{2} \cdot \left(\frac{1}{z} - \frac{1}{R}\right). \tag{7.3-1}$$

Applying the Debye approximation, as shown in Fig. 7.3-1 to Eq. (7.3-1), and using the relation of the spatial frequency and angle in Eq. (2.3-16) it can be shown that the spatial frequency of the spherical phase is given by

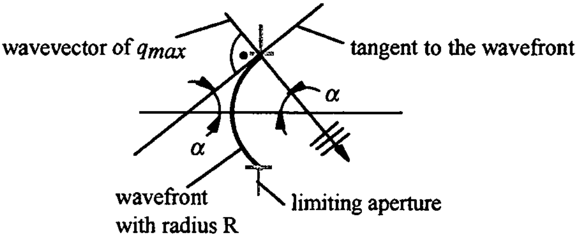


Figure 7.3-1: Finding the maximum spatial frequency q_{max} by Debye approximation.

$$q_x = \frac{1}{2 \cdot \pi} \cdot \frac{d\varphi(x)}{dx} = \frac{x}{\lambda} \cdot \left(\frac{1}{z} - \frac{1}{R}\right). \tag{7.3-2}$$

If the limiting aperture at $z = 0$ has the radius r, then the maximum spatial frequency is

$$q_{x,\max} = \frac{r}{\lambda} \cdot \left(\frac{1}{z} - \frac{1}{R}\right). \tag{7.3-3}$$

Introducing Eq. (7.3-3) into the equation of the sampling theorem of Eq. (7.2-7) yields for the distance between two sampling points

$$\Delta x(z) < \left| \frac{\lambda}{2 \cdot r} \left(\frac{z \cdot R}{R - z} \right) \right|. \tag{7.3-4}$$

From Eq. (7.2-3) the number of sampling points is found to be

$$N(z) > \left| \frac{2 \cdot r \cdot L_x}{\lambda} \cdot \left(\frac{1}{z} - \frac{1}{R} \right) \right|. \tag{7.3-5}$$

If we choose $L_x \approx 2r$ and an incident plane wave ($R = \infty$), Eq. (7.3-5) becomes

$$N(z) > \frac{L_x^2}{\lambda \cdot z}. \tag{7.3-6}$$

The minimum distance Δp between the sampling points in the frequency domain for the Fresnel diffraction integral in Eq. (2.3-25) is derived in a similar way. Let us assume again that U_0 is a spherical wave with a uniform intensity distribution and with curvature radius R. From Eq. (2.3-26) its Fourier transform is found to be

$$FT\left\{ \exp\left(-\frac{ik}{R} \cdot x_0^2 \right) \right\} = i\lambda z \cdot \exp\left(i\pi\lambda R \cdot q_x^2 \right) \tag{7.3-7}$$

and hence, the phase function in the argument of the inverse Fourier transform operator in Eq. (2.3.5-4) becomes

$$\hat{\varphi}(q_x, z) = \pi\lambda q_x^2 \cdot (R - z). \tag{7.3-8}$$

Determining the derivative yields

$$x = \frac{1}{2 \cdot \pi} \cdot \frac{d\varphi(q, z)}{dq} = \lambda \cdot q_x \cdot (R - z). \tag{7.3-9}$$

With $q_{max} \approx F_x/2$ we get from Eq. (7.3-9) $x_{max} \approx \lambda F_x(R-z)/2$. Using Eqs. (7.2-8) and (7.2-4) we finally obtain the condition

$$N(z) > F_x^2 \cdot \lambda \cdot |(R - z)| \tag{7.3-10}$$

for the number of sampling points in the frequency domain.

If U_0 is a plane wave ($R = \infty$) propagating along the axis z and diffracted at an aperture at $z = 0$ $FT\{U_0\}$ becomes a sinc-function (or a Bessel function if a two dimensional wavefront is considered). The maximum slope of the phase in the frequency domain is then mainly given by the propagator $exp(-i\pi\lambda z q_x)$ in Eq. (2.3-25) so that R in Eq. (7.3-10) can be neglected. Using Eqs. (7.2-3) and (7.2-5) we see that $F_x = N(z)/L_x$ and therefore we get form Eq. (7.3-10)

$$N(z) < \frac{L_x^2}{\lambda \cdot |z|}, \tag{7.3-11}$$

which is the sampling condition for the inverse Fourier transform in function of the side length L_x of the array U_0 in the space domain. In contrast to Eq. (7.3-6), Eq. (7.3-11) gives an upper limit of the number of sample points. The direct explanation is found in the representation of the Fresnel integrals: in Eq. (2.3-19), for which the condition of Eq. (7.3-6) is valid, the quadratic phase function in the argument of the Fourier operator depends on the propagation distance by $1/z$ where as in Eq. (2.3-25) for which the condition of Eq. (7.3-11) is valid, the dependence on z is linear.

Now let us estimate the number of sampling points for a given situation. As an example we take a plane wavefront which is transformed into a divergent wave by a concave aberration free lens of focal length $f < 0$ and having an aperture of radius r. The lens is placed at $z = 0$. At the distance of $z = -f$ we determine the field $U(x,z = -f)$. The extension of the wavefront at the exit pupil of the lens is about twice the radius r and therefore, we choose $L \approx 2r$. Using Eq. (7.3-5) the number of sample points is given by

$$N > \frac{8 \cdot r^2}{|f| \cdot \lambda}. \tag{7.3-12}$$

Taking the example of a negative lens with focal length $f = -10$ mm, a pupil radius $r = 2.5$mm, and a wavelength of 633 nm Eq. (7.3-12) we get $N \approx 8000$, which means that for the two dimensional wavefront about $8000^2 = 64$ million samples points are necessary. For instance, MATLAB uses about 16 bytes for a complex number. The required memory space for mathematical operations with complex numbers is about five times the needed array space. Hence, for a matrix of 64 million elements a memory of $64 \cdot 10^6 \cdot 16 \cdot 5 = 5 \cdot 10^9$ bytes is necessary. This example shows that the evaluation of the Fresnel integral by DFT or FFT exceeds the available memory of a common desktop computers. Therefore, we are looking for a strategy to reduce the required number of sampling points.

7.4 Reducing the number of sampling points

The crude method of evaluating the diffraction integral by Fourier transform in the example of Sec. 7.3 has shown that the required number of sampling points is large or even too large to be handled by desktop computers. In this section we will show that by transformation of the Fresnel integrals in Eqs. (2.3-19) and (2.3-25), so that tilt and sphere in the phase function of U_0 are eliminated, the propagation of a wavefront can be evaluated independent of the available size of computer memory, at least in the case where the spherical aberrations of the wavefront are zero.

Such a transformation has already been presented by Sziklas [7.5, 7.6], based on the propagation of a Gaussian wavefront and limited to near and intermediate distances. However, the procedure presented here is based on a general wavefront and can be applied for any distance, as long as the spherical aberrations are not significant.

For the sake of simplicity we carry out the derivation for the one-dimensional case. Let us assume a wave

$$U_0(x_0) = A_0(x_0)\cdot \exp\left(i\varphi_0(x_0)\right) \tag{7.4-1}$$

at the exit pupil of an optical system propagating along the axis z. The phase $\varphi_0(x_0)$ can be expanded as

$$\varphi_0(x_0) = k\cdot\left(a_0 + b_0\cdot x_0 - \frac{1}{2c_0}\cdot x_0^2\right) + \sigma(x_0), \tag{7.4-2}$$

where a_0 is the coefficient for piston, b_0 for tilt, and c_0 is the radius of the sphere. The coefficient c_0 is positive for a converging and negative for a diverging. The function $\sigma(x_0)$ describes the spherical aberration of the phase. The coefficients in Eq. (7.4-2) of an arbitrary phase function $\varphi_0(x_0)$ could be determined by a least square fit [7.7]. Using Eqs. (7.4-1) and (7.4-2) the Fresnel integral of Eq. (2.3-19) becomes

$$U_1(x_1,z_1) = \frac{\exp\left(ik\left(z_1 + \frac{x_1^2}{2z_1}\right)\right)}{i\lambda z_1} \cdot \int_{-\infty}^{+\infty} A_0(x_0)\cdot\exp\left(ka_0 + kb_0x_0 + kx_0^2\left(\frac{1}{2z_1} - \frac{1}{2c_0}\right)\right)\cdot\exp\left(-ik\frac{x_0x_1}{z_1}\right)\cdot dx_0 \quad . \tag{7.4-3}$$

We introduce the following transformation of the variables:

$$\frac{1}{\xi} = \frac{1}{z_1} - \frac{1}{c_0} \tag{7.4-4}$$

$$\beta = \frac{\xi}{z_1} \quad \chi = x_1 \cdot \beta . \tag{7.4-5}$$

From Eq. (7.4-3) and using the shift theorem we get

$$u(\mathrm{X} = \chi - b_0\xi) = \exp(ika_0)\int_{-\infty}^{+\infty} \mathcal{A}_0(x_0)\cdot\exp\left(ik\frac{x_0^2}{2\xi}\right)\cdot\exp\left(-ik\frac{x_0\mathrm{X}}{\xi}\right)\ dx_0, \tag{7.4-6}$$

where $\mathcal{A}_0(x_0) = A_0(x_0)\cdot\exp(i\sigma(x_0))$. Applying the discrete Fourier transform from Eq. (7.4-6) yields

$$u_D(\mathrm{X} = \chi - b_0\xi) = \exp(ika_0)\cdot DFT\left\{\mathcal{A}_0(x_0)\cdot\exp\left(ik\frac{x_0^2}{2\xi}\right)\right\} \tag{7.4-7}$$

and thus the field $U_1(x_1,z_1)$ is given by

$$U_1(x_1,z_1) = \frac{\exp\left(ik\left(z_1 + \frac{x_1^2}{2z_1}\right)\right)}{i\lambda z_1}\cdot u_D(x_1\beta - z_1b_0\beta) . \tag{7.4-8}$$

In the case where $\sigma(x_0) \approx 0$ Eq. (7.4-7) corresponds to the Fresnel integral of a wave having constant phase. If furthermore the wave has a uniform intensity distribution, the Eqs. (7.4-4) and (7.4-5) transform the Fresnel integral of a spherical wave into the Fresnel integral of a plane wave.

An alternative expression for Eq. (7.4-7) is found if Eq. (7.4-6) is brought into the form

$$u(\mathrm{X}=\chi-b_0\xi)=\exp\left(ik\cdot\left(a_0-\frac{\mathrm{X}^2}{2\xi}\right)\right)\int_{-\infty}^{+\infty}\mathcal{A}_0(x_0)\cdot\exp\left(ik\frac{(x_0-\mathrm{X})^2}{2\xi}\right)\cdot dx_0. \qquad (7.4\text{-}9)$$

Using the theorem of convolution of the Fourier transform, the relation of Eq. (2.3-26) and the discrete Fourier transform will yield

$$\begin{aligned}u_D(\mathrm{X}=\chi-b_0\xi)=\exp\left(ik\cdot\left(a_0-\frac{\mathrm{X}^2}{2\xi}\right)\right)\cdot DFT^{-1}\{DFT\{\mathcal{A}_0(x_0)\}&\\ \cdot\exp\left(-i\cdot\pi\cdot\xi\cdot\lambda\cdot q^2\right)\}&\end{aligned}\;. \qquad (7.4\text{-}10)$$

In contrast to Eq. (7.4-7) we have to use twice the DFT (or FFT) in Eq. (7.4-10). However, this is still faster and more precise than to evaluate the convolution by numerical integration. It can be shown that the Eq. (7.4-8) with u_D of Eq. (7.4-10) is in fact Eq. (2.3-25) with the transformed variables from Eqs. (7.4-4) and (7.4-5).

We have found the two transformed Fresnel integrals presented in Eqs. (7.4-8) and (7.4-10). In order to show the procedure of choosing one of them let us assume that the spherical aberration of the wavefront are not significant and that the available computer limits memory the size of the array to N_C. The range of distances ξ in the Eqs. (7.4-7) and (7.4-10) for which no aliasing occurs, are then given by Eqs. (7.3-6) and (7.3-11). Hence, for Eq. (7.4-7) the range corresponds to

$$|\xi|>\frac{L_x^2}{N_C\cdot\lambda} \qquad (7.4\text{-}11)$$

and for Eq. (7.4-10) to

$$|\xi|<\frac{L_x^2}{N_C\cdot\lambda}. \qquad (7.4\text{-}12)$$

We see from these two equations that using the Eqs. (7.4-7) and (7.4-10) the field u_D (and thus U_1) in a transverse plane can be calculated at any distance ξ (and thus at any distance z_1) independent of the available maximum array size N_C. The maximum possible size of array N_C just determines at what distance ξ (or z_1) the program has to switch from Eq. (7.4-7) to Eq. (7.4-10) or vice versa. However, a small sampling number (i.e. N_C) will result in a low resolution of the calculated wave. Depending on required resolution a certain minimum of sampling points (N_C) is therefore necessary. The algorithm can also be used for the improved Fresnel approximation in Eq. (2.3-20).

A closer look to the transformation reveals that the Eqs. (7.4-4) and (7.4-5) correspond to the lens formulae of geometrical optics. In fact, physically, we image the converging or diverging wave into a propagating plane wave having the same amplitude distribution $\mathcal{A}_0(x_0)$ at $z_0 = 0$. At the 'image' distance ξ we find the same amplitude distribution as at the image distance z_l but in a co-ordinate system magnified by β.

For the two-dimensional case the Eq. (7.4-8) takes the form

$$U_l(x_l, y_l, z_l) = \frac{\exp\left(ik\left(z_l + \frac{x_l^2 + y_l^2}{2z_l}\right)\right)}{i\lambda z_l} \cdot u_D\left(x_l\beta_x - z_l b_x\beta_x, y_l\beta_y - z_l b_y\beta_y\right). \quad (7.4\text{-}13)$$

If $\beta_x \neq \beta_y$ the wavefront is astigmatic.

7.5 Examples

The equations of Sec. 7.4 have been implemented as a MATLAB function. Apart from the transformed Fresnel integrals, additional functions have been implemented; namely thin paraxial lenses and apertures which are necessary for the design in Chap. 8. These functions enables to calculate an optical system by simply listing in a script file the functions of the optical elements as they appear in the design. An example of a collimation system consisting of a Gaussian source and a collimation lens of a focal length of 0.1 m, for which the wavefront is evaluated at a distance of 100 m, is shown in appendix I.

We shall use the established relations to demonstrate first the significance of the procedure developed in the Sec. 7.4. Second, we evaluate the diffraction pattern of a plane wave passing through a slit and a focusing lens. Finally, we shall discuss the focal shift of a Gaussian beam. The general set-up is shown Fig. 7.5-1. The source is in the transverse plane 0. The emitted plane wavefront has a wavelength of $\lambda = 633$ nm and infinite extension. The lens is placed in the plane 1 and the aperture (pupil or slit) is placed in the plane 2. The plane 3 is the observation plane. The distances are: $z_{01} = z_{12} = 0$ and $z_{23} > 0$.

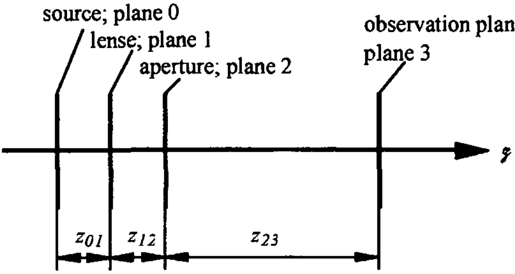


Fig. 7.5-1: General se-up for he simulations. The distances are: $z_{01} = z_{12} = 0$ and $z_{23} > 0$.

7.5.1 Significance of the transformation

Let us place in the plane 1 a thin lens with focal length of $f_0 = 0.5$ m and in the plane 2 a pupil with diameter $d = 0.02$ m. For both planes we choose $L_x = L_y = 2d = 0.04$ m. The observation plane 3 is at $z_{23} = 1.2$ m.

Let us determine $/\xi_C/$, the distance at which the evaluation of the intensity distribution is switched between the Eqs. (7.4-7) and Eqs. (7.4-10). The memory size of the computer allows $N_C^2 = 512^2$ sampling points. Introducing $N_C = 512$ into Eq. (7.4-11) yields, with $L_x = 0.04$ m, a distance $/\xi_C/ = 5$ m. Let us now evaluate $/\xi/$ for the propagating wave. The radius c_0 is given by the focal length f of the lens, i.e. $c_0 = f$. According to Eq. (7.4-4) $/\xi/ = 0.86$ m. Hence, $/\xi/ < /\xi_C/$ and therefore, we are obliged to use the transformed Fresnel integral for near and intermediate distances, i.e. Eq. (7.4-8) with Eq. (7.4-10). The resulting intensity distribution is shown in Fig. 7.5-2. In contrast, if we use the non-transformed Fresnel integral, i.e. Eq. (2.3-25) of Sec. 2.3.5, and the same $N_C = 512$, the evaluated intensity distribution becomes as shown in Fig. 7.5-3. It is clearly seen that the distribution is rather random indicating that the sampling condition is not fulfilled.

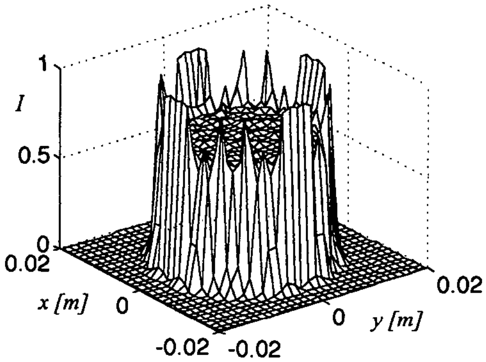


Figure 7.5-2: Intensity distribution I in the observation plane 3 calculated with 128x128 sample points and using the transformed Fresnel diffraction integral Eq. (7.4-8) with Eq. (7.4-10) (diagram is a 32x32 grid). Clearly the sampling theorem is satisfied.

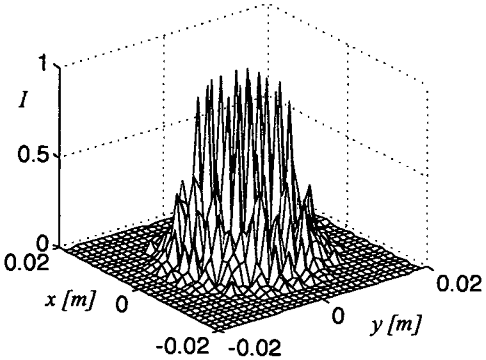


Figure 7.5-3: Intensity distribution I in the observation plane 3 calculated with 512x512 sample points and using the non-transformed Fresnel diffraction integral Eq. (2.3-25) (diagram is a 32x32 grid). The random intensity distribution indicates that the sampling theorem has not been satisfied.

7.5.2 Diffraction by a slit

The diffraction pattern of a slit of 0.01 m width illuminated by a converging wave is evaluated. A Fourier lens of focal length $f = 0.576$ m is placed in plane 1 of Fig. 7.5-1. The slit is put into plane 2. The diffraction pattern in plane 3 is calculated at distances z_{23} ={0.530 m, 0.550 m, 0.576 m}. The program switches automatically between the two forms of the transformed Fresnel diffraction integrals. The calculated diffraction patterns are verified with the measurement set-up presented at the EOS topical meeting on 'Optoelectronic Distance/Displacement Measurements and Applications' in Nantes (France) [7.8]. The results are presented in Figs. 7.5-4 to 7.5-6. They show good agreement between calculations and measurements.

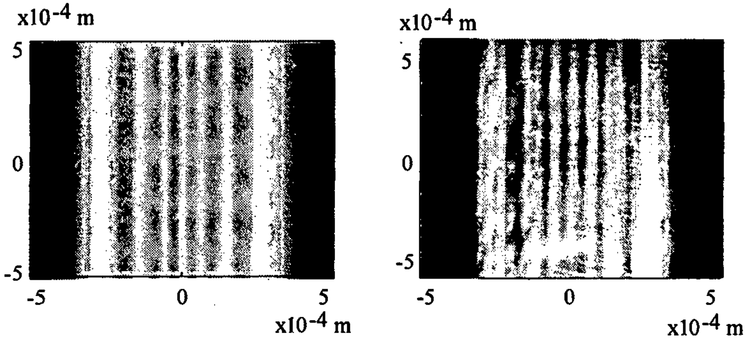


Fig. 7.5-4: Intensity pattern at z_{23} = 0.530 m. Left calculated pattern, right measured pattern.

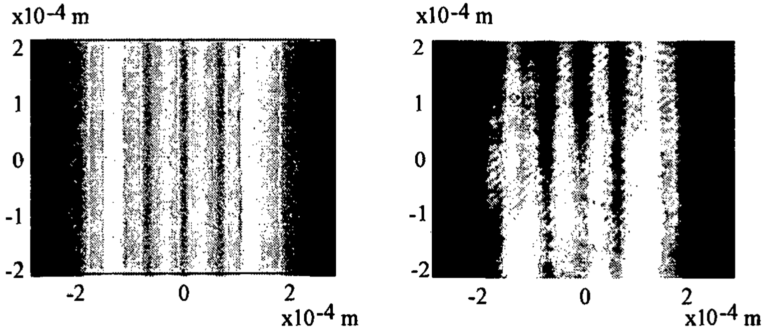


Fig. 7.5-5: Intensity pattern at z_{23} = 0.550 m. Left calculated pattern, right measured pattern.

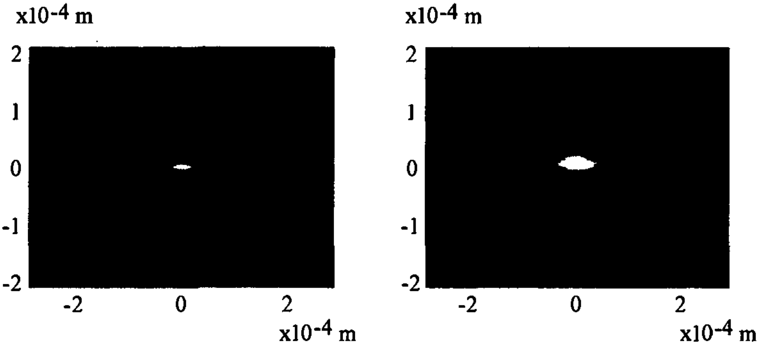


Fig. 7.5-6: Intensity pattern at z_{23} = 0.576 m. Left calculated pattern, right measured pattern.

7.5.3 Focal shift

Contrary to the prediction of geometric optics the point of highest intensity of a focused plane wave is not in the focal plane of the focusing lens but in the front of it. This effect is known as focal shift and mostly evaluated by solving the Fresnel integral of the converging wave [7.9]. However, using the derivations of the Sec. 2.4, it can be shown that if the field of the collimated wave is known, the focal shift can be evaluated by the law of geometrical optics [7.10].

Let us demonstrate the focal shift with the program. We use a Gaussian laser beam having a waist radius of $w_o = 5\cdot 10^{-4}$ m and a thin lens with focal length of $f = 0.2$ m. The focal shift δ of the Gaussian beam is given by [7.11]

$$\delta = -\frac{f}{1+\left(\frac{\pi\cdot w_o^2}{\lambda\cdot f}\right)^2}, \qquad (7.5\text{-}1)$$

where the negative sign indicates that the shift is from the focal plane towards to the lens. With the given values the focal shift is $\delta = -5.05\cdot 10^{-3}$ m or $z_{23} = f + \delta = 0.19495$ m. It can be shown that at δ the wavefront is plane and the radius of the sphere c_0 becomes infinite and that for $0 < z_{23} < (f + \delta)$ the radius c_0 is positive (the wave converges) whereas for $z_{23} > (f + \delta)$ the radius c_0 is negative (the wave diverges). The simulation calculates the radius c_0. However, as c_0 approaches infinity the results of the calculation become very inaccurate due to the limited accuracy (number of digits) in the MATLAB program. Therefore, we will not look for z_{23} for which $c_0 = \infty$ but for a small interval of $[z_{23,1}\ldots z_{23,2}]$ for which $|c_0|$ is larger than 50 m. This approach yields for the left hand side of the interval a $c_0 = 93$ m at $z_{23,1} = 0.19494$ m and for the right hand side of the interval a $c_0 = -105$ m at $z_{23,2} = 1.9496$ m (the print out of the program is shown in appendix II). We see that the evaluated focal shift δ by Eq. (7.5-1) is well within the interval given by the focal shifts $\delta_1 = f - z_{23,1} = -5.06 \cdot 10^{-3}$ m and $\delta_2 = f - z_{23,2} = -5.04 \cdot 10^{-3}$ m, i.e. $\delta_1 < \delta < \delta_2$.

7.6 Conclusion

We have introduced the discrete Fourier transform in order to solve the Fresnel diffraction integral numerically. However, we have seen that the sampling theorem imposes a minimum number of sampling points, which depends on the initial curvature of the wavefront and on the propagation distance. A simple example has shown that this number can easily exceed the available memory size of any available desktop computers. However, we were able to overcome this problem for wavefronts with relative small spherical aberrations. We have introduced a transformation, which transforms the Fresnel diffraction integral of a spherical wavefront into a Fresnel diffraction integral of a plane wavefront. Using this transformation, we have derived an algorithm which allows evaluating the field of any propagating spherical wavefront at any distance independent of the available memory size of the computer. Making use of this algorithm we have shown the effect of the sampling theorem and the performance of the introduced algorithm by calculating the intensity distribution of a propagating wave. Thereafter, we have calculated the intensity distribution of a uniform converging wave after having passed through a slit. The intensity distribution was compared with measurements. We have found good agreement between the calculated and measured intensity distributions. Finally, we have used the algorithm to investigate the focal shift of a Gaussian beam and have found the same results as predicted by the focal shift formula. Hence, we have shown that a program for wave propagation based on the derived algorithm of Sec. 7.4 works accurately and efficiently within the approximation of scalar paraxial wave optics.

References

[7.1] J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill Series in Electrical and Computer Engineering, New York, 1996, 2nd edition), pp. 4-27.

[7.2] R. N. Bracewell, *The Fourier Transform and its Applications* (McGraw-Hill International Editions, New York, 1986, 2nd edition), pp. 356-410.

[7.3] M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Wiley Series in Pure and Applied Optics, New York, 1991), pp. 217-219.

[7.4] J. J. Stamnes, *Waves in Focal Region* (Adam Hilger, Bristol and Boston, 1986), pp. 378-387.

[7.5] E.A. Sziklas, A. E. Siegman, "Diffraction calculations using fast Fourier transform modes," Proc. IEEE **62**, 410-412 (1974).

[7.6] E.A. Sziklas, A. E. Siegman, "Mode calculations in unstable resonators with flowing saturable gain. 2: Fast Fourier transform method," Appl. Opt. **14**, 1874-1889 (1975).

[7.7] W. H. Press *et al.*, *Numerical Recipes in Pascal, Threat of Scientific Computing* (Cambridge Press, 1989), pp. 558-571.

[7.8] Urs Vokinger, Hans Peter Herzig, "Intensity distribution measurement at intermediate and far distance of nearly collimated beams," EOS, Topical Meetings Digests Series **14** (1987).

[7.9] Y. Li, E. Wolf, "Focal shifts in diffracted converging spherical waves," Opt. Commun. **39**, 211-215 (1981).

[7.10] U. Vokinger, R. Dändliker, P. Blattner, H. P. Herzig., "A different view of focal shift," Opt. Commun. **157**, 218-224 (1998).

[7.11] A. G. van Nie, "Rigorous calculation of electromagnetic field of wave beam," Philips Res. Repts. **19**, 378-394 (1964).

Appendix I

Example of a MATLAB script file of a collimation system with Gaussian source whose beam is evaluated at 100 m distance of its exit pupil (Sec. 7.5)

```
% Start of program collimator
% ____________________

clear all;
Init_HappyII;

% Creat source
PlaneNumber = 0;
VacuumWaveLength = 633e-9;
Aperture = 40e-3;
Waists = [1e-6,1e-6];
PosData = [0,0];
TotalFlux = 1;
RefracIndex = 'air';
PlaneSize =[0.5e-3,0.5e-3];
            gaussian_source( PlaneNumber, Aperture, PlaneSize, ...
                               Waists, PosData, TotalFlux, RefracIndex);

% Propagate to plane 1
PlaneNumber = 1;
Distance = 100e-3;
RefracIndex = 'air';
            propagate (PlaneNumber, Distance, RefracIndex)

% Collimator lens f = 100mm in plane 1
PlaneNumber = 1;
FocalLength = 100e-3;
Aperture = 5e-3;
PosData = [0,0];
            thin_paraxial_lens (PlaneNumber, FocalLength, Aperture,PosData)

% Propagate to plane 2
PlaneNumber = 2;
Distance = 100;
RefracIndex = 'air';
            propagate (PlaneNumber, Distance, RefracIndex)
```

Appendix II

Print out for the focal shift of a Gaussian beam with waist radius of $5 \cdot 10^{-4}$ m focused by a lens with $f = 0.2$ m (Sec. 7.5.3)

First print out: z_1 = z =0.19494 m; c_{01} = X-Radius = Y-Radius =92.7854 m

```
Data of the wavefront in plane number 3:
   Location of the plane:
       x = 0 units of length
       y = 0 units of length
       z = 0.19494 units of length
   Relative flux of the wavefront:
       Flux = 1 (relative).
   Optical Path = 0.19494 units of length
   Tilt of Wavefront @ n=1:
      X_Tilt: -2.2973e-19 rad.
      Y_Tilt: 1.1789e-18 rad.
   Radius of Wavefront @ n=1:
      X_Radius: 92.7854 units of length.
      Y_Radius: 92.7854 units of length.
      Resulting Astigmatism ( 1 - Y_Radius/X_Radius ): -3.5601e-1
```

Second print out: z_2 = z =0.19496 m; c_{01} = X-Radius = Y-Radius = -104.824 m

```
Data of the wavefront in plane number 3:
   Location of the plane:
       x = 0 units of length
       y = 0 units of length
       z = 0.19496 units of length
   Relative flux of the wavefront:
       Flux = 1 (relative).
   Optical Path = 0.19496 units of length
   Tilt of Wavefront @ n=1:
      X_Tilt: -1.6517e-19 rad.
      Y_Tilt: -3.5129e-19 rad.
   Radius of Wavefront @ n=1:
      X_Radius: -104.8244 units of length.
      Y_Radius: -104.8244 units of length.
      Resulting Astigmatism ( 1 - Y_Radius/X_Radius ): -1.1265e-11
```

Third print out: z_2 = z =0.2000 m (focal plane); c_{01} = X-Radius = Y-Radius = -0.19985 m

```
Data of the wavefront in plane number 3:
  Location of the plane:
     x = 0 units of length
     y = 0 units of length
     z = 0.2 units of length
  Relative flux of the wavefront:
     Flux = 1 (relative).
  Optical Path = 0.2 units of length
  Tilt of Wavefront @ n=1:
     X_Tilt: -3.3772e-19 rad.
     Y_Tilt: -9.7409e-19 rad.
  Radius of Wavefront @ n=1:
     X_Radius: -0.19985 units of length.
     Y_Radius: -0.19985 units of length.
     Resulting Astigmatism ( 1 - Y_Radius/X_Radius ): 5.6621e-15
```

CHAPTER 8

Realizing a pointing system

8.1 Introduction

In this chapter we shall describe the realization of an optical pointing system. We define as optical pointing system an optical device emitting the energy flow into the free space within a small solid angle. The fundamental application of such a system is to point a distant target by hitting it with the emitted beam. Such devices are commonly used in range finders and laser pointers. The design will be based on the theory of the previous chapters.

The system will be assembled by the newly developed Optical Surface Mounted Device technique (OSMD) [8.1]. This assembling system was developed at the Swiss Federal Institute of Technology in Lausanne (EPFL) with the goal to find an automated assembling system for mass production of optical devices. The OSMD technique consists of an assembling automate, of universal holders (Fig. 8.1-1) for optical elements of maximum 9 mm diameter, and a base plate on which the optical system will be assembled. At the beginning of the assembly, the universal holder is brought to its intended position by the assembling automate so that the three flexible and spherical shaped legs touch the surface of the base plate. During the on-line-alignment procedure of the assembling automate, the legs keep mechanical contact with the base plate. After the alignment of the six degree of freedom the

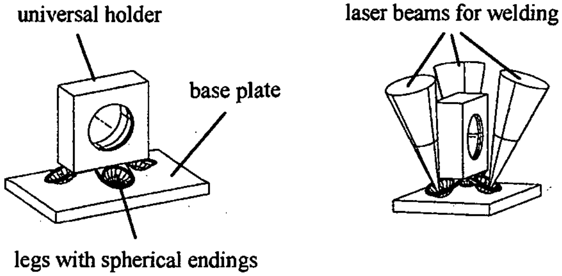


Figure 8.1-1: Left: universal holder (10 mm x 10mm x 4 mm) with flexible legs on a base plate. Right: welding of a universal holder on the base plate with three laser beams.

universal holder is fixed on the base plate by simultaneous laser point welding of the three legs.

The chapter will begin with the specification of the pointing system. In the next section, we shall design the system, based on the theory of the previous chapters. The proposed layout of the design will then be checked with a simulation program, which has been developed according to Chap. 7. The last sections will present the measurements compared with the results of the simulations and conclusions.

8.2 Specification for the design

The pointing system is designed as a module, able to be incorporated in different optical equipment. It consists of two sub-modules. The source module contains four lasers, a beam combiner and a fiber into which the combined beam is coupled. The collimator module consists of a collimation lens (the principal collimation lens), which collimates the output beam of the fiber (Fig.8.2-1). The source module shall be as compact as possible. Since mass production, and hence, low assembling cost is the goal of the developed product the source module shall be assembled by automated OSMD.

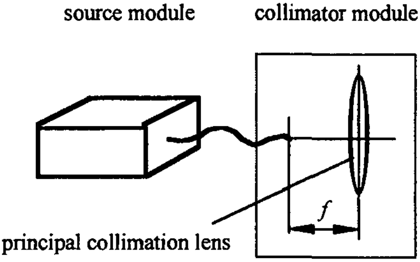


Figure 8.2-1: Concept of the pointing system. It consists of two modules: the source module and the collimator module.

The use of a fiber instead of simply joining the two modules together will certainly introduce losses, however, the fiber reduces the design constrains for the equipment in which the module will be installed: the source module can be mounted anywhere in the equipment independent on the principal collimation lens.

The basic specifications of the pointing system are:

wavelength $\lambda = 1550$ nm
diameter of exit pupil of the principal collimation lens $d_P = 50$ mm
optimized for target size H x H = 2.3 m x 2.3 m at distance of $z_T = 4$ km
module size as small as possible
suitable for mass production (i.e. assembling by OSMD)

Some elements have been specified by Leica. These are

pulsed laser diode: C86091E (EG&E, Quebec, Canada)

peak wavelength: $\lambda = 1550$ nm
spectral bandwidth: $\Delta\lambda = 30$ nm
emitting size: length: $l_x = 150$ µm (on axis x)
width: $h_y = 1$ µm (on axis y)

measured intensity function along the emitting length l_x

$$I(x) \approx I_0 \cdot rect\left(\frac{x}{l_x}\right) \tag{8.2-1}$$

measured radiant intensity function expressed in angles

$$J(\theta_x, \theta_y) = J_0 \exp\left(-2\left(\frac{\tan(\theta_x)}{\tan(\zeta_x/2)}\right)^2\right) \cdot \exp\left(-2\left(\frac{\tan(\theta_y)}{\tan(\zeta_y/2)}\right)^2\right) \tag{8.2-2}$$

with diffraction angles

$\zeta_x = 0.54$ rad (full slow angle @ $1/e^2$)
$\zeta_y = 1.4$ rad (full fast angle @ $1/e^2$)

TE polarization (parallel to the junction): $p_{TE} = 98\%$

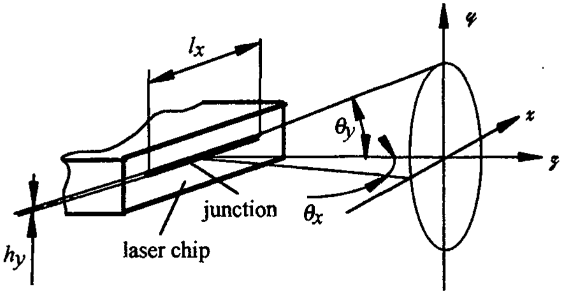


Figure 8.2-2: Notation for the laser.

multi-mode fiber: AF 100/140 IRA (Fiberware GmbH, Berlin, Germany)

core diameter: $d_c = 100\ \mu\text{m}$

acceptance function of fiber measured at Leica:

$$T_l(\theta) = \exp\left(-2 \cdot \left(\frac{\tan(\theta)}{\tan(\kappa/2)}\right)^2\right) \tag{8.2-3}$$

with acceptance angle $\kappa = \kappa_x = \kappa_y = 0.42$ rad (full angle @ $1/e^2$)

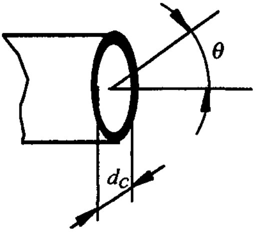


Figure 8.2-3: Notation for the fiber.

Further specification

optimization parameter: maximum energy flow on target at 4 km

prototype cost optimized.

8.3 Design

In this section the design of the pointing system will be presented. We start with an investigation of the necessary phase space volume of the system. As mentioned in Sec. 5.3 the phase space volume defines the size of an optical system when energy flow is a major concern. Hence, in order to have a small source module with low losses we need to answer the question whether the defining element of phase space volume, i.e. the laser diode, can be squeezed and transformed. Having found an answer to these questions, we shall carrying on with the coupling of the beam into the fiber. The optimization of the coupling will give us a first information about the optical imaging system in the source module. In the next step we shall choose the type of combiner. In this part we shall pay attention to the number of necessary optical elements. Since each supplementary element will increase the size of the module, we have to find a combining system which consists of a minimum number of optical elements. Having found a convincing solution, the design of the pointing system has reached a stage where the parameters of the lenses can be evaluated and the layout of the source module can be drawn. With this layout the procedure of the assembling can be studied. In a short section we shall briefly explain the assembly strategy. The feasibility and tolerances of the assembling will be checked by a simulation program based on Chap. 7. The simulation shall show the performance of the pointing system and reveal the points of highest lost in energy flow.

8.3.1 Consideration on squeezing phase space volume of the source

For the desired compactness of the source module a small phase space volume of the emitted field is important. Therefore, let us investigate whether the phase space volume of the specified laser source can be reduced. For that purpose we have to decompose the radiated field of the pulsed laser diode into its natural modes of oscillation.

The specified small spectral bandwidth of the source enables to consider the laser diode as a source emitting at a single wavelength. The measurements show that the intensity distribution in the far field of the source is nearly Gaussian (Eq. (8.2-2)). One of the major characteristics of a propagating Gaussian beam is that the transverse intensity distribution remains Gaussian

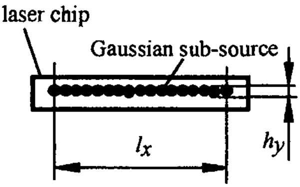


Figure 8.3-1: Model of the laser diode C86091E (EG&E).

at all distances (paraxial approximation assumed). Therefore, the intensity distribution in the source plane should also be a Gaussian. However, the measured intensity distribution in the source plane parallel to the junction (x-direction) is approximately a rectangular function. A solution to this contradiction is found if the measured rectangular function is considered as composed of many small Gaussian sub-sources juxtaposed along the junction of the laser diode (Fig. 8.3-1). The size of the assumed sub-sources is given by the waist

$$w_0 = \frac{2 \cdot \lambda}{\zeta \cdot \pi}, \tag{8.3-1}$$

where ζ is the full diffraction angle of the source [8.2]. With the specified diffraction angles in Sec. 8.2 the waists of the sub-sources are $w_{0,x} = 1.8$ µm and $w_{0,y} = 0.7$ µm. Hence, the waist $w_{0,x}$ is quite small compared with the emitting length l_x so that it is justified to consider the measured rectangular function as a superposition of non-correlated Gaussian sub-sources distributed along the junction of the laser chip. Using the notation introduced in Sec. 3.5.1, we may therefore express the radiant field of the source by

$$\mathcal{U}(\rho_0) = \sum_N \mathcal{A}_k \cdot \psi_k(\rho_0) \tag{8.3-2}$$

with

$$\langle \mathcal{A}_k^* \cdot \mathcal{A}_l \rangle = \alpha_k \cdot \delta_{kl}, \tag{8.3-3}$$

to express that the sub-sources are non-correlated, and with

$$\psi_k(\rho_0) = \sqrt{\frac{2}{\pi \cdot w_{0,x} \cdot w_{0,y}}} \cdot \exp\left(-\left(\frac{x_0 - x_{0,k}}{w_{0,x}} \right)^2 - \left(\frac{y_0}{w_{0,y}} \right)^2 \right) \tag{8.3-4}$$

describing the non-correlated sub-sources as Gaussian functions normalized to 1. In the case where the distribution of the Gaussian sub-sources is not discretely, as it is assumed here, but continuos, $x_{0,k}$ in Eq. (8.3-4) has to be replaced by the corresponding distribution function and hence, the Eq (8.3-2) will result into an integral.

Furthermore, the intensity distribution along the junction is uniform (specifications of Sec. 8.2) and hence, the amplitudes $\mathcal{A}_k$ in Eq. (8.3-2) must satisfy

$$\mathcal{A}_0^2 = \mathcal{A}_1^2 = \mathcal{A}_2^2 \ldots = \mathcal{A}_{N-1}^2 = \mathcal{A}_N^2. \tag{8.3-5}$$

With this description of the source, the possibility of phase space squeezing can be answered straightforwardly. First, the Eq. (8.3-2) describes the sub-sources as coherent Gaussian and thus, they emit the light in the smallest possible phase space volume. Therefore, the phase space volume of the entire source cannot be squeezed by modifying the phase space volume of the individual sub-sources. Second, the phase space volume could be squeezed by modifying the amplitude distribution $\mathcal{A}_k$ of the non-correlated sub-sources so that the intensity in the center of the junction is maximum and decreases towards the edges. However, because of Eq. (8.3-5) the source has highest entropy and hence, under the condition that the modification must be loss less and carried out by passive optical elements (e.g. micro-lenses and micro-prism), any change in the distribution of $\mathcal{A}_k$ violates the second law of thermodynamics. We have therefore to conclude that the phase space volume of the emitted field of the used laser diode cannot be squeezed.

8.3.2 Matching the phase space volume of source and optical system

Let us use the concept of the approximated phase space volume of Sec. 5.3 in order to check whether the optical system with the given specification does match the phase space volume of the source. We will carry out this check for the source module and the collimation module separately and according to the obtained results discuss the limits of matching. We will start with the source module.

The approximated phase spaces volume of the laser diode expressed by the $\mathcal{M}^2$ factor is (Eq. (5.3-4)) in x-direction

$$\mathcal{M}^2_{x,B} = \frac{\zeta_x \cdot l_x \cdot \pi}{4 \cdot \lambda} \approx 40 \tag{8.3-6}$$

and in y-direction

$$\mathcal{M}^2_{y,B} = \frac{\zeta_y \cdot h_y \cdot \pi}{4 \cdot \lambda} \approx 0.7. \tag{8.3-7}$$

The phase space volume in Eq. (8.3-7) is smaller than unity which is theoretically not possible. We assume that the junction width $h_y = 1$ µm is larger than specified by the manufacturer. Nevertheless, we consider this disagreement as an indication that the sub-sources are Gaussian and that the field in the source plane is coherent in y-direction. The entire approximated phase space volume $\mathcal{M}_L^4$ of the combined beams of the four laser diodes in the source module is

$$\mathcal{M}^4_B = 4 \cdot \mathcal{M}^2_{x,B} \cdot \mathcal{M}^2_{y,B} \approx 160. \tag{8.3-8}$$

The approximated phase space volume in the two axis of the fiber are

$$\mathcal{M}^2_{x;F} = \mathcal{M}^2_{y,F} = \frac{\kappa_x \cdot d_c \cdot \pi}{4 \cdot \lambda} \approx 20. \tag{8.3-9}$$

However, in order to obtain the total phase space volume the simply multiplication of $\mathcal{M}_{x,F}^2$ and $\mathcal{M}_{y,F}^2$ would lead to a too large phase space volume due to the rotational symmetry of input pupil and acceptance function (Eq. 8.2-3) of the fiber. Therefore, we introduce a form factor F which is the ratio of a circular area with diameter d to a quadratic area with side length d, i.e. $F = (\pi/4)$. This factor has to be squared since near and far field have rotational symmetry. Hence the entire phase space of the fiber is

$$\mathcal{M}^4_F = F^2 \cdot \mathcal{M}^2_{x,F} \cdot \mathcal{M}^2_{y,F} \approx 240. \tag{8.3-10}$$

In order to avoid losses in the source module the phase space volume of the combined beams must be smaller than the phase space volume of the fiber into which the beams are coupled.

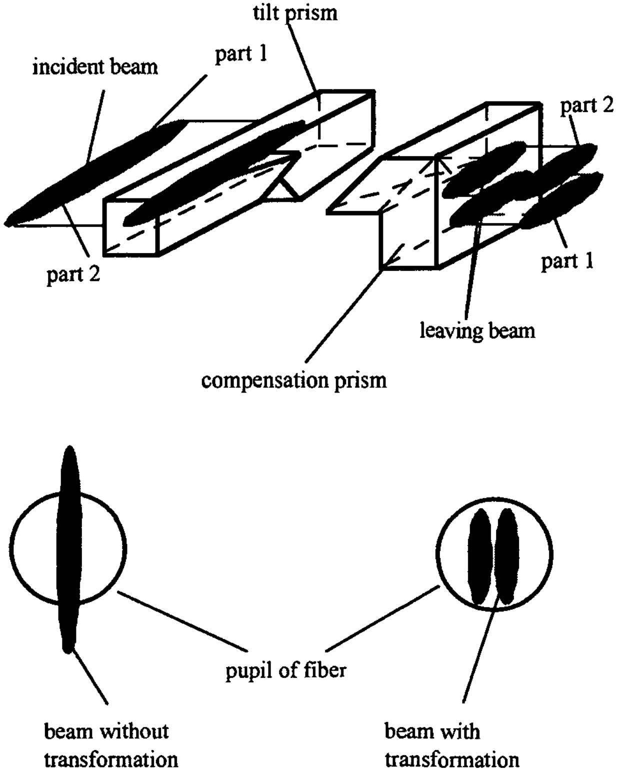


Figure 8.3-2: The principle of source transformation is shown. The tilt prism introduces two different tilts to the wavefront so that after a distance of propagation the wavefront is divided into two parts where one part is upon the other. The compensation prism compensates the introduced tilt so that the two parts of the beam remain together as they continue to propagate. The effect of the transformation is shown at the bottom: before, only about half of the beam is coupled into the fiber; after, the entire beam will be coupled into the fiber.

Comparing the result of Eq. (8.3-8) and Eq. (8.3-10) it can be seen that $\mathcal{M}_F^4 > \mathcal{M}_B^4$ and hence, in a first estimation, it should be possible to couple the whole beam into the fiber. However, comparing the value of Eq. (8.3-6) with the value of Eq. (8.3-9) it results that $\mathcal{M}_{x,B}^2 > \mathcal{M}_{x,F}^2$, and therefore, the phase space volume in x-direction cannot be coupled into the fiber. However, because the values of Eqs. (8.3-7) and (8.3-9) are such that $\mathcal{M}_{y,B}^2 << \mathcal{M}_{y,F}^2$ a transformation as described in Sec. 5.4 can be used to match the phase space volume of the laser to the phase space volume of the fiber. This transformation could be carried out as shown in Fig. 8.3-2 where the laser source is folded by a set of prisms so that it fits well to the circular aperture of the fiber.

However, in order to reduce costs for the first prototype, we decided that the phase space volume of the source will not be matched to the phase space volume of the fiber.

Finally, let us compare the phases space volume of the fiber with the phase space volume formed by the collimator module and the target. The approximated phase space volume of the exit pupil of the collimator module with target is

$$\mathcal{M}_{x;P}^2 = \mathcal{M}_{y,P}^2 = \frac{\zeta_P \cdot H \cdot \pi}{4 \cdot \lambda} = \frac{\frac{d_P}{z_T} \cdot H \cdot \pi}{4 \cdot \lambda} \approx 15 \qquad (8.3\text{-}11)$$

and hence, resulting in a total phase space volume of

$$\mathcal{M}_P^4 = F \cdot \mathcal{M}_{x,P}^2 \cdot \mathcal{M}_{y,P}^2 \approx 176, \qquad (8.3\text{-}12)$$

where the form factor F has to be taken only once, since the target is a square surface. Comparing the value of Eq. (8.3-11) with the value of Eq. (8.3-9) we see that $\mathcal{M}_{x,P}{}^2 < \mathcal{M}_{x,F}{}^2$ and $\mathcal{M}_{y,P}{}^2 < \mathcal{M}_{y,F}{}^2$ and hence, the phase space volumes of fiber and the exit pupil with the target are mismatched. Since the phase space volume of the collimator module with the target is smaller in x-direction and y-direction than that of the fiber, a transformation is not possible. This mismatch introduces a loss of about $1\text{-}(\mathcal{M}_P{}^4 / \mathcal{M}_F{}^4) \approx 0.3$ for the efficiency of the pointing system.

8.3.3 Coupling into the fiber

The coupling is carried out by imaging the junction of the laser chip onto the core of the fiber. In the previous section we have found that due to the asymmetry of the laser diode the x-component of the phase space volume of the laser diode and of the fiber do not match. Furthermore, a source transformer, which would match theses two volumes, is not planed for the first prototype. However, losses due to this mismatch can be minimized if an optimal image of the source onto the core of the fiber is formed. We have mentioned in Sec. 5.3 that in conjugate planes of a loss less optical imaging system the phase space volumes are equal. Hence, it results from this property that by magnifying the image of the emitting length l_x on the fiber the imaged diffraction angle ζ_x will be reduced by the same amount. Thus, there may be a magnification for which the coupling efficiency is maximized. In order to find this maximum we have to find an analytical expression for the coupling efficiency. Such an expression can be found by multiplying the x-component of the radiant intensity function of the laser diode with acceptance function of the fiber, then integrating it over the angle of the half space and finally normalizing it by the total emitted energy flow of the laser diode. This yields the following expression

$$T_1 = \frac{\int_{-\pi/2}^{-\pi/2} J_0 \cdot \exp\left(-2 \cdot \left(\tan(\theta_x)\right)^2 \cdot \left(\left(\frac{1}{\tan(\kappa/2)}\right)^2 + \left(\frac{1}{\tan(\zeta_x/(2 \cdot \beta))}\right)^2\right)\right) \cdot d\theta_x}{\int_{-\pi/2}^{-\pi/2} J_0 \cdot \exp\left(-2 \cdot \left(\tan(\theta_x)\right)^2 \cdot \left(\frac{1}{\tan(\zeta_x/(2 \cdot \beta))}\right)^2\right) \cdot d\theta_x} \tag{8.3-13}$$

if the image is completely within the core of the fiber, and the magnification is β. Using the paraxial approximation $\tan(\theta) \approx \theta$ and carry out the integration from -∞ to +∞, the above expression becomes

$$T_l \approx \left[1 + \left(\frac{\zeta_x}{\kappa \cdot \beta}\right)^2\right]^{-\frac{1}{2}}. \tag{8.3-14}$$

In the case where the imaging of the emitting length l_x is longer than the core diameter d_c of the fiber, we have to multiply Eq. (8.3-14) with a factor T_2. This factor expresses the relative quantity of the energy flow which enters into the pupil of the fiber. If the intensity distribution along l_x is a rectangular function and the emitting length l_x is aligned to the center of the (circular) core of the fiber (Fig. 8.3-3), the factor T_2 is simply given by

$$T_2 = \begin{cases} \dfrac{d_c}{l_x \cdot \beta} & l_x \cdot \beta > d_c \\ 1 & l_x \cdot \beta \le d_c \end{cases} \tag{8.3-15}$$

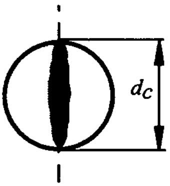


Figure 8.3-3: Best coupling is achieved if the image of the emitting length is in the center of the core of the fiber.

and therefore, the ratio of the power flow coupled into the fiber is

$$T = T_1 \cdot T_2 = \begin{cases} \dfrac{d_c}{l_x \cdot \beta} \cdot \left[1 + \left(\dfrac{\zeta_x}{\kappa \cdot \beta}\right)^2\right]^{-\frac{1}{2}} & l_x \cdot \beta > d_c \\ \left[1 + \left(\dfrac{\zeta_x}{\kappa \cdot \beta}\right)^2\right]^{-\frac{1}{2}} & l_x \cdot \beta \leq d_c \end{cases} . \tag{8.3-16}$$

In Eq. (8.3-16) the only parameter which is not specified is β. Therefore, we search for the condition of β for which T is a maximum. This condition is found just by inspection of the two equations and is

$$\beta = \frac{d_c}{l_x}, \tag{8.3-17}$$

yielding with the specified values a magnification of $\beta = 0.67$ and a coupling efficiency of $T \approx 0.5$.

8.3.4 Combining the fields of the sources

Let us find in this section the design for the field combiner by a trial and error approach. The specifications define how it has to be: small and low cost, i.e. the number of elements has to be as small as possible; furthermore, the combining of the field is limited to pupil filling and combining by polarization since the sources are non-correlated and polarized (Chap 6). We shall start with the simplest design and ask then whether it is feasible. If we have to reject it, we will suggest a new one with an increased minimum number of elements.

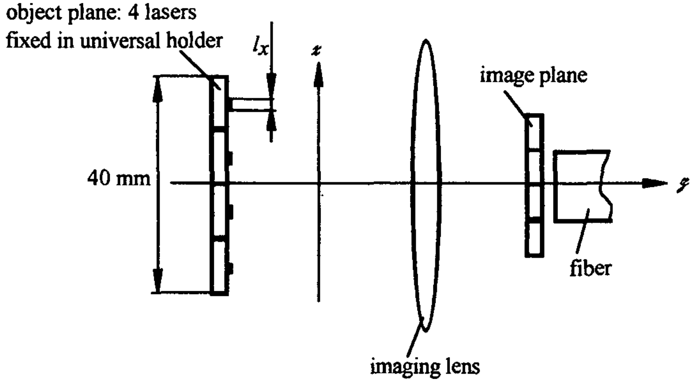


Figure 8.3-4: The simplest layout of coupling the beam of four lasers into the fiber.

It is easy to find the simplest design. In fact it is implicitly given by the preceding section: the sources has to be imaged by magnification with $\beta = 0.67$ (Sec. 8.3.3) into the core of the fiber. The most basic design is shown in Fig. 8.3-4 and consists of the four laser diodes, a single

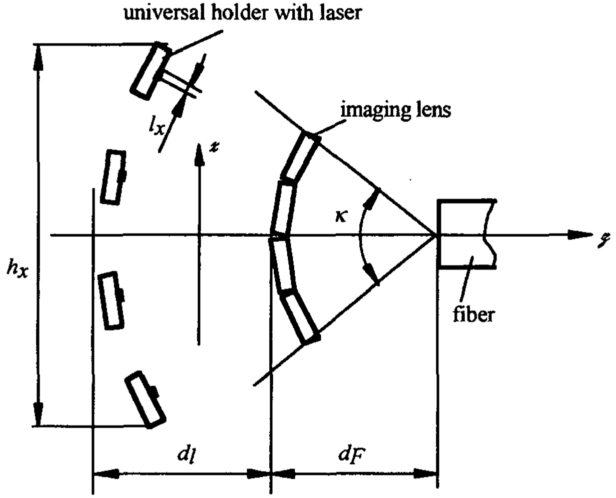


Figure 8.3-5: The second layout consists of four lasers and four lenses. Combining is carried out at the input the of the fiber.

lens, and the fiber. The laser diodes are juxtaposed; the lens images the junctions of the lasers with the necessary magnification onto the core of the fiber. The beam combining is carried out by pupil filling at the imaging lens. This design is simple, however we have to reject it: the universal holders for the automated assembling are relatively large so that the distance between two laser junctions is about $d_h = 10$ mm. This results in a phase space volumes of $\mathcal{M}_x^2 \approx 10^4$ and $\mathcal{M}_y^2 \approx 1$, which is many times larger than the phase space of the fiber. The next possible design (Fig. 8.3-5) is that each laser diode gets its own imaging lens. We have to increase the number of lenses by a factor of four. Each lens has to be mounted in a universal holder, i.e. the distance between two lenses is about the width $d_h = 10$mm of the universal holders. In contrast to the previous suggestion combining in this design is carried out at input of the fiber. In order to get a good coupling efficiency, the universal holders of the imaging lenses must be within the acceptance angle κ of the fiber. Hence, we are able to estimate the distance between lenses and fiber which is $d_F \approx 4d_h/(\kappa/2) \approx 200$ mm. The distance between laser diodes and lenses is given by the magnification $d_l \approx d_F/\beta \approx 300$ mm. The resulting length of this module is about $d_{tot} = d_F + d_l = 500$ mm and the width is $h_x = d_{tot}\kappa = 200$ mm. This design has to be rejected due to its size.

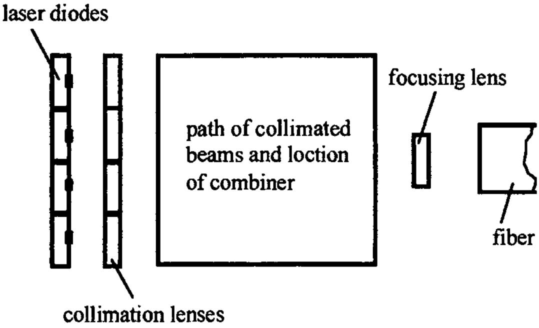


Figure 8.3-6: Draft of third layout. Imaging is carried out by collimation and the focusing lens. The combiner is shown as a black box.

Imaging can be carried out by collimating the laser beam and focussing it by an additionally lens. Hence, we increase the number of lenses by 1 for the imaging part of the design and try to find a convenient combining for this suggestion. The principle of this suggestion is shown in Fig. 8.3-6. The four lenses (the collimation lens) next to the laser diode collimate the beams. In the collimated optical path, the beams are combined and then focused by the fifth lens (the focussing lens) onto the core of the fiber. The sources are not correlated and hence, combining can be carried out by polarization or pupil filling. Pupil filling would be the cheapest solution but will probably yield high losses since the outer beams will be limited by the aperture of the focusing lens (Fig. 8.3-7). Hence, we exploit the high polarization degree of the emitted beam: We combine two by two beams by pupil filling using mirrors and combine these combined

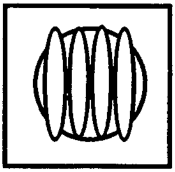


Figure 8.3-7: This figure shows the problem of combining by pupil filling. The outer beams are limited by the pupil, which causes losses.

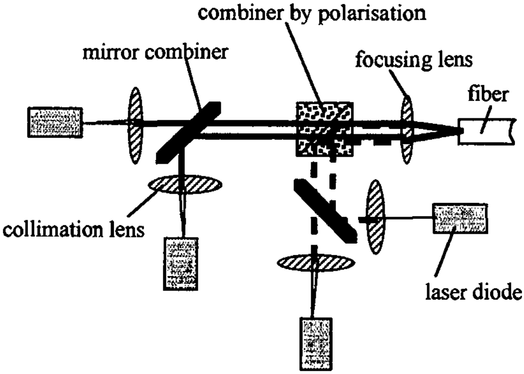


Figure 8.3-8: Final concept of the design for the source module.

beams using a polarizing beamsplitter. Both combining procedures are carried out in the collimated part of the optical path. The conceptual optical design is shown in Fig. 8.3-8. The mirror which combines two beams by pupil filling is shown in Fig. 8.3-9. It is a glass plate where one half is coated with a reflective layer (e.g. aluminum). The mirror coating on the glass plate yields the necessary sharp edge between the transmitting and reflecting part of the combiner.

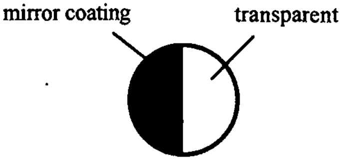


Figure 8.3-9: Mirror combiner. The mirror is coated on a glass plate yielding the necessary sharp edge between reflective and transmitting part.

8.3.5 Choosing the optics

The concept of the layout is now fixed and shown in Fig.8.3-8. We shall take the next step and carry out a quantitative calculation of the lenses. As shown in Chap. 2, the intensity distribution in conjugate planes of a lens are related by the imaging law of geometrical optics also for coherent waves. Therefore, we can use this law not only to evaluate the imaging of the entire source but also to evaluate the imaging of a single Gaussian sub-source.

8.3.5.1 Reducing the fast diffraction angle

The fast diffraction angle of laser diode corresponds to a numerical aperture of about $NA_{fast} \approx \zeta_y/2 \approx 0.7$. This is rather high compared to common lenses having typically numerical apertures of $NA < 0.3$. Therefore, we show in this section how $NA_{fast} \approx 0.7$ can be reduced to a value smaller than $NA_{max} \approx 0.3$.

The principle is presented in Fig. 8.3-10. It is based on the fact that by increasing the waist of Gaussian beam the diffraction angle is reduced. This is achieved by collimating the beam

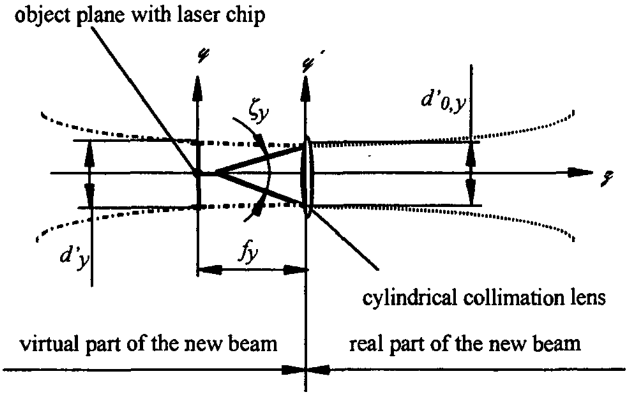


Figure 8.3-10: Collimating the fast angle of the laser beam. The dashed lines shows the new beam consisting of a virtual and a real part.

(bold) with a cylindrical lens in the $\mathfrak{y}'$-$\mathfrak{z}$ plane. The resulting beam is shown by the dashed lines in Fig. 8.3-10. The diameter d'_y of the new beam in the object plane is given by

$$d'_y \approx \frac{\zeta_y}{2} \cdot f_y \qquad (8.3\text{-}18)$$

if the effects of diffraction are neglected (assuming that the object plane is well within the Rayleigh zone of the new beam, i.e. $d'_y \approx d'_{y,0}$). This diameter is imaged by the magnification factor β on the core of the fiber. For an efficient coupling into the fiber with a core diameter d_c the condition

$$d'_y \cdot \beta < d_c \qquad (8.3\text{-}19)$$

must be satisfied, yielding an upper limit of the focal length

$$f_{y,up} = \frac{2 \cdot d_c}{\zeta_y \cdot \beta} \approx 210 \mu m. \qquad (8.3\text{-}20)$$

The lower limit of f_y is obtained by the diffraction angle ϑ of the new beam. This angle is given for a Gaussian beam by

$$\vartheta \approx \frac{2 \cdot \lambda}{\pi \cdot d'_y}. \qquad (8.3\text{-}21)$$

Introducing $\vartheta \approx NA_{max} = 0.3$ and the Eq. (8.3-18) into Eq. (8.3-21) yields for the lower limit

$$f_{y,low} = \frac{4 \cdot \lambda}{\pi \cdot 0.3} \approx 7\mu m \quad . \tag{8.3-22}$$

Thus, the focal length of the cylindrical lens must satisfy the condition

$$7\mu m < f_y < 210\mu m . \tag{8.3-23}$$

We have searched the market for small cylindrical lenses having a focal length of micrometers operating at a wavelength of 1550 nm. There are many types of lenses offered by different manufacturer [8.3…8.6]. One cylindrical lens is offered by Doric. It is a doped fiber and the aberrations are diffraction limited up to a numerical aperture of 0.5 (the total aperture is about 0.7). Although the diffraction limited numerical aperture is smaller than of other products, the Doric lens has advantages which outweigh the eventually loss introduced by the smaller aperture. In contrast to other lenses, the Doric lens has a rotational symmetry along the axis z. Thus, the alignment is reduced by one degree of freedom, resulting in an automated alignment procedure which is cheaper than for other lenses. The cylindrical lens is available at different focal lengths, starting at 41 μm and ending at 1.712 mm. The lens can be easily cut by the user to different lengths, which relaxes the constraints of the design. Furthermore, the price per unit length is quit low compared to other lens types. Summarizing, the three points cost, easy to manipulate, and the available assortment were the convincing points to choose this lens. For the design we choose the lens with the smallest focal length. This lens satisfies the condition of Eq. (8.3-23). Furthermore, the image of the beam diameter d'_y becomes much smaller than the core diameter d_c (Eq. (8.3-19)), resulting in low coupling losses. The cylindrical lens reduces the fast angle to $\zeta_y' = 34$ mrad.

The short focal length of the cylindrical lens results in tight alignment tolerances, which are within micrometers. The automate of the OSMD, however, has only an alignment precision of 5 μm. Therefore, laser diode and cylindrical lens have to be mounted on a special alignment set-up. The cylindrical lens is fixed to the heat sink of the laser chip with a UV sensitive adhesive (Fig. 8.3-11).

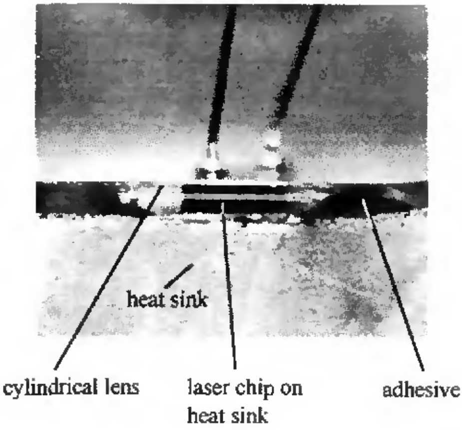


Figure 8.3-11: Cylindrical lens with 60 μm ($f = 41$ μm) diameter mounted in front of the laser chip.

8.3.5.2 Focal length of the collimation and focus lens

For the choice of the focal length of the two lenses we have to take the following data and facts into account.

- The determining numerical aperture in the source module is given by the acceptance angle of the fiber, i.e. $\kappa = 0.42$ @ $1/e^{-2}$. Hence, the focusing lens must have a $NA > 0.21$.
- The magnification of the system must be $\beta = 0.67$.
- Between the two lenses an estimated distance of $l_c \approx 50$ mm is necessary to place the mirror combiner and polarization beam splitter (for combining by polarization, Fig. 8.3-8).
- For a compact module the focal lengths of the lenses have to be as small as possible. However, the minimum focal length is imposed by the minimum distance between two universal holders, which is about 0.5 mm.
- The cost of the lenses must be as low as possible, i.e. we have to choose available lenses from a catalogue.

These specifications are summarized in Fig. 8.3-12. Applying a trial and error procedure, we have found the following lenses from the Spindler & Hoyer catalogue; collimation lens:

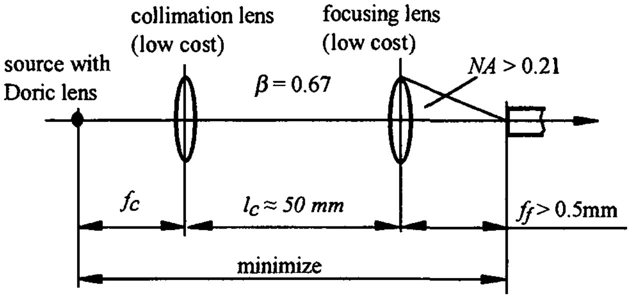


Figure 8.3-12: Specifications necessary to chose the focal lengths f_c and f_f.

positive plane-convex lens with focal length f_c = 5 mm and a pupil diameter of 5 mm; focusing lens: positive plane-convex lens with focal length f_f= 3 mm and a pupil diameter of 3 mm.

8.3.5.3 Focal length of the principal collimation lens

The aim is to get the most emitted power at the exit pupil of the fiber to the target at 4 km. The diameter of the principal collimation lens has been specified and hence, the focal length is the only parameter, which can be optimized to achieve that goal. In order to find the optimum focal length we approximate the radiation at the exit of the fiber as a superposition of non-correlated Gaussian beams, having a diffraction angle equal to the measured beam acceptance angle κ of the fiber. With this approximation we are able to calculate numerically the amount of energy flow in the target as a function of the focal length, using the Fresnel integral and superposition (Chap. 2 and Chap. 3).

The Fig. 8.3 -13 shows the diagram of the normalized energy flow in the target in function of the focal length of the principal collimation lens. Normalization has been carried out with the energy flow at the exit pupil of the fiber. The target has a side length of 2.3 m and is at a distance of 4 km. The maximum energy flow is about 0.6 and is found at a focal length of 0.2 m.

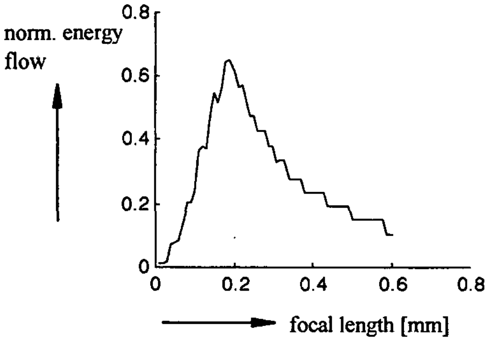


Figure 8.3-13: Normalized energy flow from the exit pupil of the fiber to the target (side length 2.3 m) in function of the focal length of the principal collimation lens. An optimum is achieved if the focal length is about 0.2 m.

8.3.6 Layout

The resulting layout from the above discussion is shown in Fig.8.3-14. The four laser diodes with the glued cylindrical lens are mounted in the universal holders 1 to 4. The junction length l_x of the laser diodes in the universal holders 1 and 2 is parallel to the base plate contrary to these in holders 3 and 4, which are perpendicular to the base plate. The beams from the universal holders 1 and 2 are collimated by the lenses in the universal holders 5 and 6, respectively, and finally combined by the mirror combiner in the universal holder 9. Similarly for the laser diodes with cylindrical lenses in the universal holders 2 and 4. They are collimated by the collimation lenses in the universal holders 7 and 8, respectively, and then combined by the mirror combiner in universal holder 10. The two beams coming from the mirror combiner are combined by the polarization beamsplitter in universal holder 11. The leaving beam is focused by the focusing lens in universal holder 12 and finally coupled into fiber fixed in the universal holder 13.

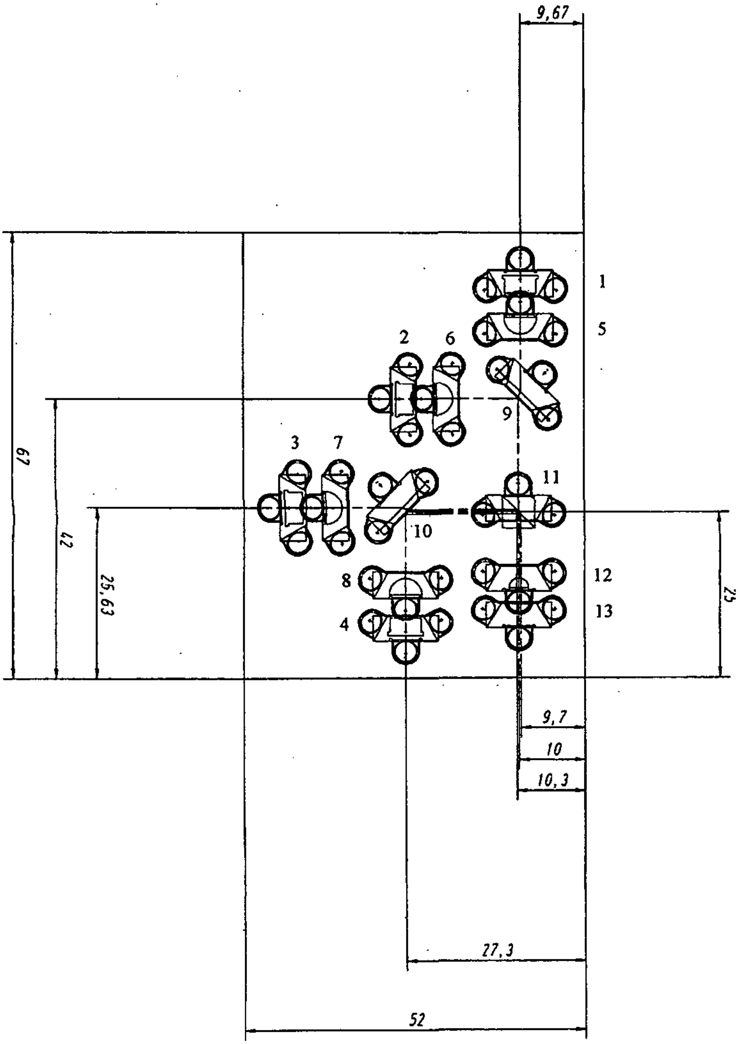


Figure 8.3-14. Layout of the source module.

8.3.7 Assembling

The assembling automate of the OSMD, with which our source module shall be assembled, is a first prototype for which the feasibility of automated assembling has be shown successfully. However, the main deficiency of the automate is the little available space to mount the necessary alignment equipment for the assembling. Especially far infrared (1550nm) cameras are bulky and cannot be fixed onto the mounting plate of the automate. Therefore, we decided to carry out most part of the assembling with visible light at 633nm. The alignment is carried out with the help of four small collimated sources fixed on the mounting plate and aligned to the optical axis of the infrared lasers (Fig. 8.3-15). These sources emit collimated and uniform beams at 633nm with 0.6mm diameter. The CCD camera is also fixed on the mounting plate in the optical axis of the fiber. The first step of the alignment consists of finding the location of the optical axis of the fiber on the CCD array of the camera. This can be achieved by the collimated source No.1. However, due to the shift introduced by the mirror combiner the beam of this collimator does not coincide with the optical axis of the fiber. A plane parallel glass plate is used to compensate this shift. Then the collimator No. 1 is switched on and the beam position on the CCD array is measured. This reference point on the CCD will define the optical axis of the combined beams and the fiber.

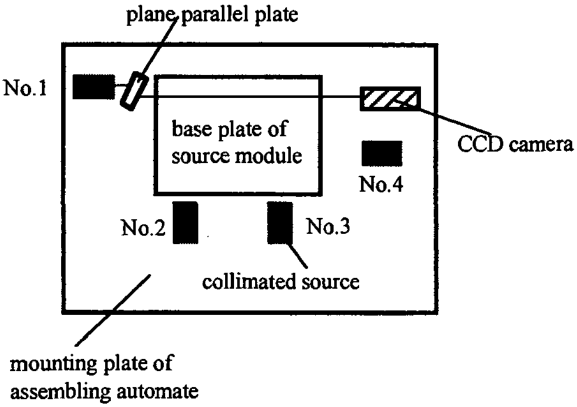


Figure 8.3-15. Mounting plate of the assembling automate. The step of finding the reference point on the CCD array of the camera is shown.

In the following steps all the optical elements, except the laser diodes and the fiber are aligned with respect to that reference point with the help of the four collimated beams: first the polarization beamsplitter (11 in Fig. 8.3-14), second the mirror combiners (9 and 10), third the collimation lenses (5, 6, 7 and 8) and finally the focusing lens (12). Next, the lasers (1, 2, 3 and 4) are aligned with the help of a cross hair in the focal plane of the focusing lens. The cross hair is imaged onto the infrared camera. It acts as a target on which the imaged spots of the lasers have to be centered. The laser diodes are in the object plane when the imaged spot has minimum size. Finally, the fiber will be mounted. The alignment is controlled by measuring the output power.

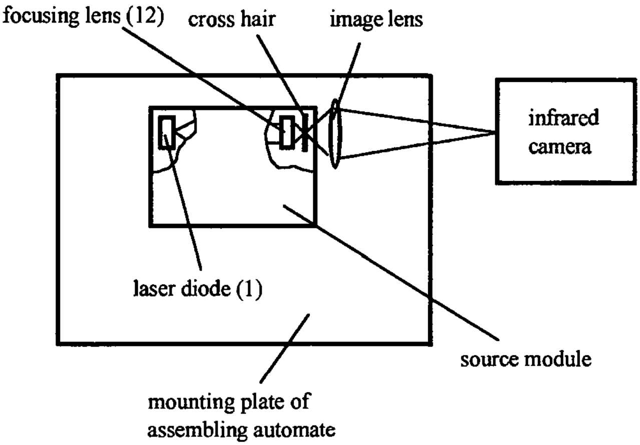


Figure 8.3-16: Set-up for the alignment of the laser diodes 1, 2, 3 and 4.

8.3.8 Simulation

The performance of the layout in Fig. 8.3-14 has been evaluated by the simulation program. The simulation program is based on the mathematical development of Chap. 7. We used the Eqs. (8.3-2) to (8.3-4) to describe the laser sources analytically. Due to the restricted available memory, the simulation has been carried out for aberration free lenses (surfaces anti-reflection coated) and therefore, the results correspond to an ideal case. The relative energy flow calculated at different position is shown in Tab. 8.3-1. Hence, only about 40% of the emitted

position	relative energy flow	relative loss between consecutive positions
exit pupil laser	1.00	-
exit pupil cylindrical lens	0.90	0.10
exit pupil collimation lens	0.90	0.00
exit mirror combiner	0.89	0.01
exit focusing lens	0.82	0.08
exit fiber	0.41	0.50
exit principal collimation lens	0.24	0.41
target	0.24	0.00

Table 8.3-1: Relative energy flow and relative transmission of the elements evaluated by the program.

power of the laser diode gets coupled into the fiber. The highest loss of about 50% occurs as expected (Eq. (8.3-16)) at the fiber input because of the mismatch of the phase space volume. The loss in energy flow due to limiting pupils is relatively low and does not exceed 20% (difference between energy flow at the exit the pupil of laser and of the focusing lens). About 24% of the emitted energy flow of the lasers reaches the target.

Because of the long target distance z_T, the exit pupil of the fiber will be imaged on the surface of the target. The resulting magnification is given by the geometrical magnification β (Eqs. (2.4-4) and (2.4-6)), which can be shown to be approximately $\beta \approx z_T/f_{PC}$. With the given value a magnification of $\beta \approx 4$ km / 0.2 m $= 20 \cdot 10^3$ is obtained, so that the imaged core diameter is about $d_c \cdot \beta = 2$ m, corresponding approximately to the side length of the target (2.3 m).

Each step of the assembling procedure has been verified with the simulation program. The simulation has revealed some weak points of the assembling procedure, which were than eliminated. The principal steps of the assembling are as presented in Sec. 8.3.7. The figure 8.3-17 shows the expected image at the CCD camera (Fig. 8.3-15) for the aligned collimation lens. This lens has to be aligned about 0.3 mm on the right hand side of the optical axis in order to compensate the lateral shift caused by the substrate thickness of the mirror combiner.

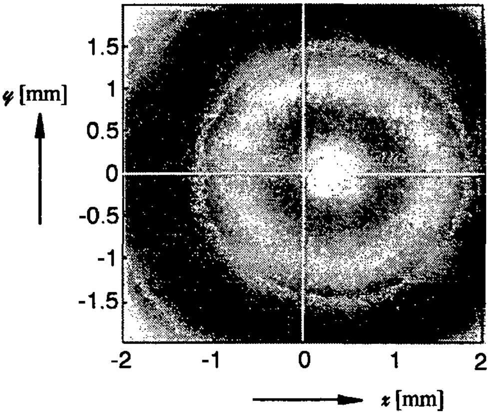


Figure 8.3-17: Simulated diffracion pattern of the alignment of the collimation lens.

8.4 Measurements

We present in this section the measurements which have been carried out. Due to the operating wavelength of 1550 nm it was difficult to find equipment to carry out precise measurements. Furthermore, the separation between the optical elements was so small that it was not possible to place probes to measure the energy flow after each element. However, the measurements give sufficient information about the performance and errors of the assembled source module.

We start with the picture of the diffraction pattern taken during the assembling of the source module. Then, in the following section, we will present the measured performance of the assembled source module. Finally, we shall present the measurements of the cross-sectional intensity distribution of the propagating wave at the exit of the principal lens.

8.4.1 Alignment

The simulated diffraction pattern of Fig. 8.3-17 did correspond well with the observed pattern during the assembling. Figure 8.4-1 shows the pattern taken by the CCD camera (Fig. 8.3-15).

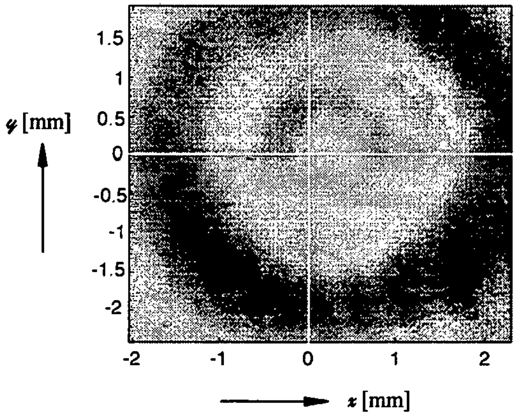


Figure 8.4-1: Measured diffracion pattern of Fig. 8.3-17 (alignment of the collimation lens).

However, it was observed, that with increasing number of optical interfaces, the diffraction pattern became noisy and the contrast decreased significantly. This was probably due to dirty surfaces and the missing anti-refection coatings on the optical elements.

8.4.2 Measurement of the relative energy flow

The most important parameter, which defines the performance of the source module, is the energy flow at the output of the fiber. The relative energy flow of the individual laser diode measured at the output of the fiber are shown in Tab. 8.4-1.

source	relative energy flow at output of fiber
laser diode 1	0 (not measurable)
laser diode 2	0.14
laser diode 3	0.02
laser diode 4	0.08
overall performance of the prototype (average of the four measured values)	0.06

Table 8.4-1: Measured energy flow at the output of the fiber.

The intensity distribution at the input of the fiber is shown in Fig. 8.4-2. It is seen that the beams are not well superposed and that the image of the emitting areas of the laser diode are blurred and longer than the expected 0.1 mm. The bad superposition of the beams can be explained by the alignment problems we experienced during the assembling. These problems were the evaluation of the exact position of the optical element with respect to the optical axes and the alignment precision and reliability of the automate. Furthermore, we have observed that the Doric lenses on the heat sink changed its position up to 5 µm along the optical axis during the curing process of the adhesive. This was specially the case for the laser diodes number 1 and 3.

Comparing the above Tab. 8.4-1 with the picture in Fig. 8.4-2, we can conclude that the fiber is alignment to the beam of the laser diode No. 2. In the following we shall show that the relative energy flow of 0.14 of this laser can be explained by including the losses at the optical interfaces and the aberration introduced by spherical lenses.
The losses at the optical interfaces are:

- About 7% and 4% for the polarization parallel and perpendicular to the incident plane, respectively, for the coating of the mirror combiner.
- About 4% for the Fresnel losses at each optical surface.

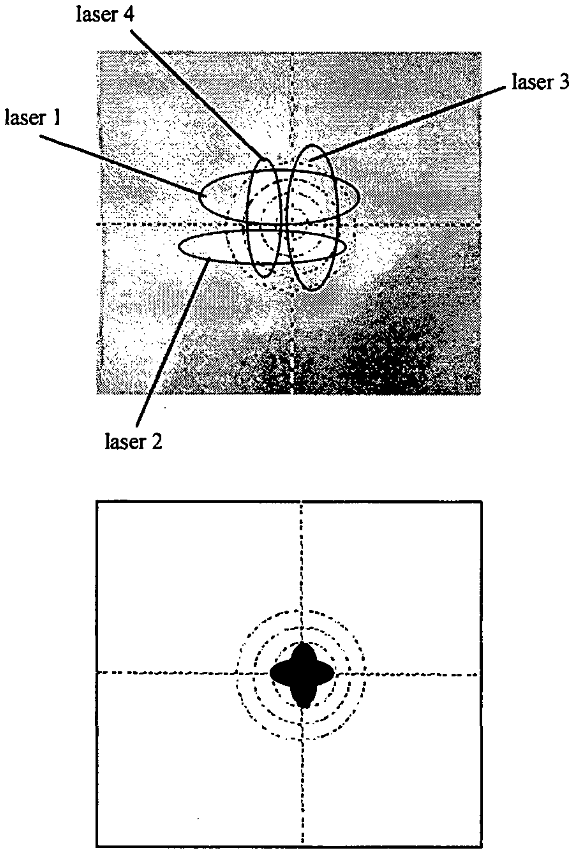


Figure 8.4-2: Top: Imaged laser sources at the input of the fiber. The number of the laser corresponds to the numbers in the layout of Fig. 8.3-14. The dashed lines are the cross hair in the focal plane of the focusing lens. The circles have diameters of 50 µm, 100 µm, 150 µm, and 200 µm. Bold lines are the borders of the images of the emitting junction of the laser diodes. Bottom: Expected intensity pattern according to the simulations.

- About 5% for the losses at the polarizing beam splitter.

These losses result in a transmission of about $(0.96)^{11} = 0.64$ (11 optical surfaces with an average loss of 4% each: Doric lens 2x, collimation lens 2x, mirror combiner 2x, focusing lens 2x, fiber 2x, and polarization beam splitter 1x).

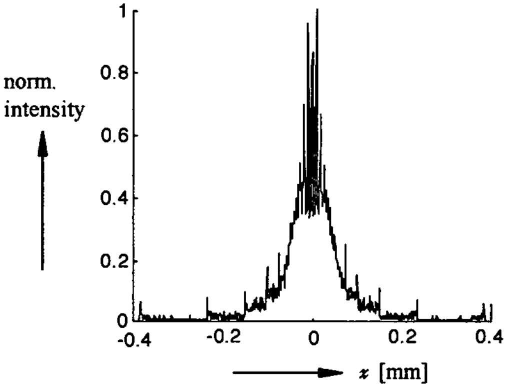


Figure 8.4-3: Plot ofthe normalized intensity distribution of the imaged emitting length l_x of the laser diodes. The intensity distribution was evaluated by ray tracing.

The losses due to the aberration introduced by the spherical lenses in the source module can be explained by Fig. 8.4-3. For this figure, the intensity distribution of the image of the emitting length l_x at the input of the fiber is evaluated by ray tracing. It can be seen that the image of the emitting length is larger than the 0.1 mm of the core diameter of the fiber. The evaluation of the diagram shows that only 66% of the energy flow is within these 0.1 mm. This bad image quality is in fact not surprising, since the numerical aperture of the focusing lens in the direction of l_x is about $NA \approx 0.4$. Multiplying the obtained values for the losses and aberration with the relative energy flow evaluated by the simulation gives
$0.64 \times 0.66 \times 0.41 \approx 0.17$, which is close to the measured value of the 0.14 for the laser diode No.2.

Hence, we are able to explain the low energy flow measured at the output pupil of the fiber for the realized source module. We have shown that because of Fresnel losses and aberration the maximum energy flow at the output of the fiber of the realized prototype does not exceed 17% of energy flow of the lasers.

8.4.3 Measurement of the cross-sectional intensity distribution of the propagating pointing beam

It is important to know how the emitted pointing beam propagates in space, in order to estimate the energy flow which is reflected back by an object to the receiver of the range finder. For long propagation distances, these measurements are usually carried out outdoor. Such measurements are costly and changing atmospheric conditions make repeatability difficult. However, by placing a Fourier lens in front of the exit pupil of the pointing system, the distances between exit pupil and target can be scaled down according to the law of geometrical optics of Chap. 2, so that measurements can be carried out in the laboratory [8.7].

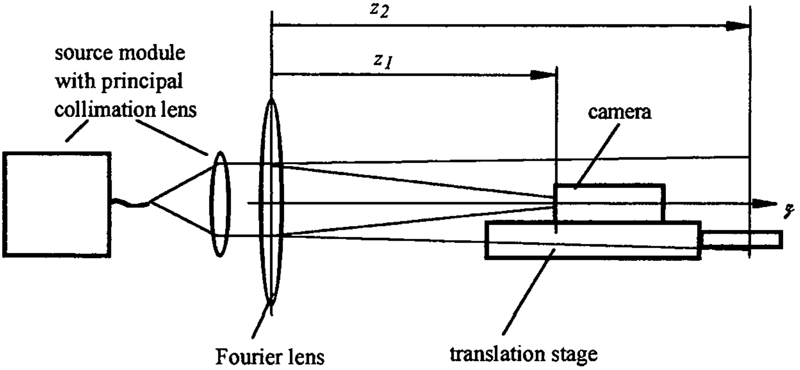


Figure 8.4-4: Set-up for the measurement of the cross-sectional intensity distribution of the output beam of the pointing system. The free propagating beam (dashed line) is imaged by the Fourier lens into a converging beam (bold line) [8.7].

Figure 8.4-4 shows the set-up. The principle plane-convex collimation lens of the pointing system had a focal length of 200 mm and a pupil diameter of 50 mm. The measured cross-sectional intensity distribution at distances equivalent to 10 m, 25 m, 100 m, and 4000 m are shown in Fig. 8.4-5. It is seen that at a distance of 10 m the beam diameter is about 20 mm.

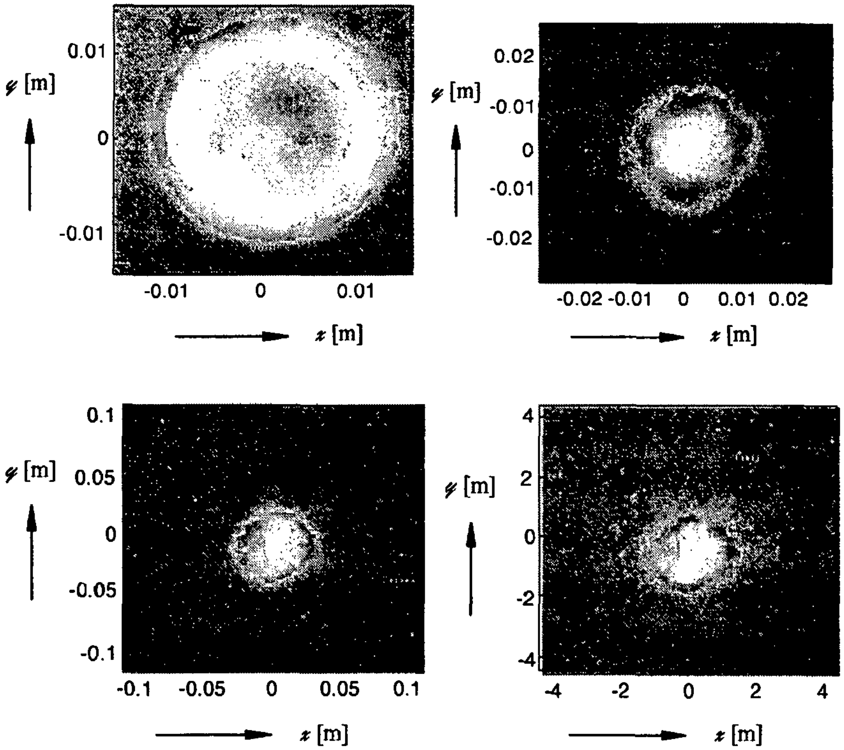


Figure. 8.4-5: Measurements of cross-sectional intensity distribution of the propagating beam of the realized pointing system. The measurements have been caried out by the set-up in Fig. 8.4.3-1 [8-7]. The equivalent propagation distances in free space (z_2) are: at 10 m (top left), 25 m (top right), 100 m (bottom left), and 4 km (bottom right).

Since the entire exit pupil of the principal collimation lens is illuminated one would expect a beam diameter of about 50 mm. This great difference is due to the aberration of the spherical collimation lens. The phase at the exit pupil of the principal collimation lens has been evaluated by ray tracing. The result is shown at the left hand side of Fig. 8.4-6. It can be seen that the collimation lens introduces spherical aberrations. The resulting intensity distribution at 10 m has also been evaluated by ray tracing and the result is shown at the right hand side of the Fig. 8.4-6. The intensity distribution corresponds well to the measured one: the beam

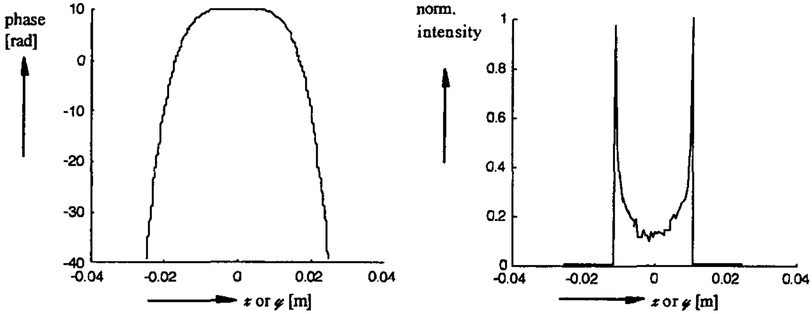


Fig. 8.4-6: The two diagrams have been evaluated by ray tracing. On the left the phase of the beam at the exit pupil of the principal collimation lens. The beam is clearly aberrated. In the optimal case the phase should be a constant. On the right, the resulting intensity distribution of the aberrated beam at 10 m distance. The intensity distribution agrees well with the measured in picture on the top left of Fig. 8.4-5.

diameter is about 20 mm and the highest intensity is found at the border of the intensity pattern. It can be expected that due to these aberrations only a small amount of the energy flow will hit the target at 4 km.

8.5 Conclusion

In this chapter we have carried out the design and realization of an optical pointing system. The design was based on the theory of the preceding chapters and the layout of the proposed design has been made suitable for the automated OSMD assembling. We have seen that the performance of the simulated and realized pointing system did not agree well. We have found that the disagreement was due to the losses and aberrations, which were not taken into account by the simulation program, and due to the badly aligned optical elements. However, as the simulation shows, eliminating all alignment problems, coating all surfaces with an anti-reflection layer, and using aberration free lenses (which have aspherical surfaces), the energy flow hitting the target at 4 km would not exceed 24% of the emitted energy flow of the laser diodes. This is mainly due to the badly matched phase space volumes of the laser diodes, the fiber, and the principal collimation lens with target. Introducing a source transformer, which matches the phase space volume of the source to that of the fiber, the performance of the pointing system could be increased to about 70%. Using a principal collimation lens with larger pupil diameter, the performance of the system would further increase. However, the increase of pupil diameter would result in an increase of the required size for the optical instrument incorporating the pointing system.

In general, we could show that the steps taken in this design were correct. The simulation program based on Chap. 7 (wave propagation) has been proven to be versatile tool to check the performance of the design. However, before using this simulation program, the optical system should be optimized for minimum aberration, using a standard ray tracing program. The simulation program based on Chap. 7 will afterwards reveal effects caused by diffraction, which might decrease the performance of the optical system.

References

[8.1] W. Andreasch, *Konzeption und Entwicklung einer Technologie zur automatisierten Oberflächenmontage optischer Elemente (Optical SMD)* (Swiss Federal Institute of Technology (EPFL), thesis No. 1591, 1996).

[8.2] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, New York, 1995), pp. 271.

[8.3] *Doric Lenses Inc.* (1439, St-Clément, Ancienne-Lorette, PQ, Canada, G2E 3L7).

[8.4] *Team Technologies Inc.*, (1216 High St., Unit A, Auburn, CA 95603, USA).

[8.5] *Blue Sky Research*, (4030 Moonpark Avenue, Suite 123, San Jose, CA 95117, USA).

[8.6] *LIMO – Lissotschenko* (Mikrooptik GmbH, Hauert 7, 44227 Dortmund, Germany).

[8.7] Urs Vokinger, Hans Peter Herzig, "Intensity distribution measurement at intermediate and far distance of nearly collimated beams," EOS, Topical Meetings Digests Series **14** (1987).

Chapter 9

Final conclusion and outlook

We have used the concept of phase space volume and entropy and the program with algorithms for wave propagation to realize a pointing system for rangefinders. The calculations with the program showed that in the proposed design not more than 41% of the total emitted energy flow of the lasers will be coupled into the fiber of the source module and that only 24% of the total energy flow will reach the target in the ideal case (paraxial approximation, no spherical aberration and no Fresnel losses). This rather low performance was principally due to the low coupling efficiency of the beam into the fiber. However, the performance of the coupling can be theoretically improved to a 100%, if the phase space volume of the combined beam is properly matched to the fiber by a source transformation. The energy flow on the target would be increased to 68%. Using a principal collimation lens with larger pupil diameter, the performance of the system could be further increased. However, the increase of pupil diameter increases the dimension of the optical instrument in which the pointing system will be incorporated.

It has to be decided for the continuation of the development of the source module if it would not be better to replace the four laser diodes by a single laser diode having a four times longer emitting length which is then folded by one source transformer to match the phase space volume of the fiber. Theoretically this is possible as long the phase space volume of the single laser diode is smaller than the phase space volume of the fiber. The single laser approach has many advantages: a combiner is not necessary and the number of optical elements is reduced at least by a factor of four, resulting in a cheaper assembling and a very compact source module. Therefore, such a combiner will be more competitive. The development of a source transformer folding the source at least twice is in any case necessary. Therefore, why not the current (December 1998) development of a source transformer in micro-optical technology extend to a development of a general type of source transformer, which folds a linear source N times. Of course this will be more expensive but will reduce the development cost of future pointing system, which have to emit an even higher energy flow.

The alternative concept of wave propagation, entropy, and phase space volume has given a better insight into the design of the optical systems. Based on this concept we were able to explain WHY the problem has to be solved in this way or WHY no further optimization is possible. This knowledge is of importance for the development of a competitive product.
The work has shown that the simulation program needs to be further improved so that large aberrations can be taken into account. However, if the design is first optimized by standard ray tracing for low aberrations, then the current version of the simulation program can be used to check the influence of diffraction for the aberration free system.
The problem of assembling is more difficult to solve. The procedure of alignment has to be changed and the OSMD assembling automate has to be further optimized to increase precision and reliability. A new type of assembling automate (TRIMO) is currently developed at the EPFL.

It has been experienced during this work that in general the (automated) assembling of optical system is a problem. Especially in micro-optics with the demanding tolerances of less than 10 μm the possibility of standard mechanics has reached a limit and new assembling techniques have to be found. Furthermore, CCD cameras and other optical systems necessary to carry out the alignment must be adapted to the new assembling technique. If the goal of a low cost optical systems by automated assembling shall be continued, it is important to focus the development on these problems. To find a satisfactory solution, as for the automated assembling of electronic circuit, not only financial investment is needed but also a lot of creative engineering work is necessary.

Acknowledgements

This work would not have been possible without the generosity of Leica Geosystems Ltd. and the Institute of Microtechnology at the University of Neuchâtel. Therefore, I would like to thanks to Dr. B. Gächter for having organized the financial support at Leica and for having given me the time to write this work. I wish to thank to Prof. Dr. R. Dändliker for having welcomed me into his department and for having directed the thesis. I greatly appreciated his lectures and the discussions during the correction of this work. I would also like to thank Dr. H. P. Herzig for giving me the chance to join his group and for his encouragement during the work although the topic was not at all in the domain of his group.

I had the chance to share an office with Dr. P. Blattner. Throughout the three years' period that we passed together he never got tired of helping me to solve the problems and to come up with new ideas. This help was not only restricted to heavy mathematical and physical problems. He also helped to solve quickly any problems with my computer and to find information on the internet. Thank you Peter!

Without the contribution of Mr. Ch. de Graffenried at the Swiss Federal Institute of Technology in Lausanne, Messrs. E. Hugelshofer (Hügi) and H. Fend, both at Leica Geosystems Ltd., the optical pointing system could never have been realized. Mr. E. Hugelshofer fabricated the mechanical parts, Mr. Fend designed and fabricated the necessary electronic devices to drive the lasers, and Mr. de Graffenried planned and carried out the major part of the assembly. I would like to take the opportunity here to express my thanks for their important contributions.

Mr. A. Schilling and Dr. P. Kipfer supported me during the tests of the far field measurement set-up which had to be carried out during the night. The tests had to be repeated since I erased the results of the first test on the computer by mistake. I wish to thank them for their help again. Furthermore, over the three years, Peter Kipfer never gave up trying to convince me of the importance of thermodynamics and of the superior quality of German beer and cars. As the attentive reader may find out, he was successful in the first point, but, for the latter two points, I can only say that I still prefer Italian cars and the Spanish wine 'Don Pascual'.

I would finally like to thank Messrs. J. R. von Allmen and J. L. Kumin for having welcomed and helped me in their mechanical work shop at the university, to Dr. Z. Szalmassy for having calculated the mirror layer of the mirror combiner, to Mrs. Dagorov for having measured the losses of this mirror, and to Messrs. L. Stauffer and R. Appius for measuring the radiation characteristic of the laser diodes and optical fiber.