Determining spatial modes of lasers with spatial coherence measurements

Carolyn M. Warnky, Betty Lise Anderson, and Charles A. Klein

We explain a technique that extracts both the structure and the modal weights of spatial modes of lasers by analyzing the spatial coherence of the beam. This is the first time, to our knowledge, that an experimental method is being used to measure arbitrary forms of the spatial modes. We applied this method to an edge-emitting Fabry–Perot semiconductor laser with a stripe width of 5 μm and extracted fundamental and first-order lateral modes with relative power weights of 96.2% and 3.8%. There was a single transverse mode. © 2000 Optical Society of America

OCIS codes: 030.1640, 030.4070, 120.3180, 140.5960.

1. Introduction

The quality of a laser beam is a topic of interest to both designers and users of optical systems. The designer wants to specify the optimum characteristics of the laser, and the consumer needs a standard for comparison of different lasers. In addition, a quality factor can be useful in the manufacturing process to determine the acceptance criteria.

Different aspects of the beam quality may have higher priority in various optical systems. Therefore it is difficult to define a quality factor by a single number. Some of the methods to measure beam quality are the M^2 second-moment method, knife-edge scanning, slit scanning, and energy measurement in a variable aperture.[1]

An alternative way to specify the quality of a laser beam is to define the form of the spatial modes and the distribution of power over the modes. This information can be of value to the designer in examining variations in the refractive-index profile, gain distribution, or structure of the laser. Similarly, specifying the modes and weights defines a standard of comparison of beam characteristics for the user.

A common standard for an ideal laser is one with a single Gaussian spatial mode. However, a laser beam could appear to be Gaussian and still be multimode. It is also possible for a laser to be single mode and not be Gaussian. For example, thin-junction semiconductor lasers have an approximately Lorentzian-shaped spatial mode in the plane perpendicular to the junction that may be truncated by packaging because of the large divergence.[2–4] Even parallel to the junction plane there may be asymmetries in fabrication and gain guiding that would cause the fundamental mode to differ from the Gaussian model.

Multimode lasers can be characterized by a basis set of spatial modes and the weights of each mode. Other techniques of determining modal weights rely on assumed expressions for the spatial modes, usually Hermite–Gaussian modes for rectangular symmetry lasers. Although Hermite–Gaussian modes have the advantage of mathematical tractability, they may not adequately represent the physical characteristics of the laser. If the Hermite–Gaussian model is physically inconsistent with the laser, erroneous results will be found for the number of modes and their weights.

The technique outlined in this paper shows how it is possible to find the weights and forms of the spatial modes without any *a priori* assumptions of the shape of the modes. The spatial modes found in this experimental method may show greater insight into the physical characteristics of the laser. This insight is especially useful in laser analysis and design.

Various methods have been proposed for determining the weights of a laser's spatial modes including curve fitting,[5] matrix inversion,[6,7] M^2 analysis,[8] frequency mixing,[9] and coherence measurements.[10–12] All these methods assumed the Hermite–Gaussian

The authors are with the Department of Electrical Engineering, The Ohio State University, 205 Dreese Laboratory, 2015 Neil Avenue, Columbus, Ohio 43210. The e-mail address for B. L. Anderson is anderson@ee.eng.ohio-state.edu.

Received 24 February 2000; revised manuscript received 8 August 2000.

0003-6935/00/336109-09$15.00/0

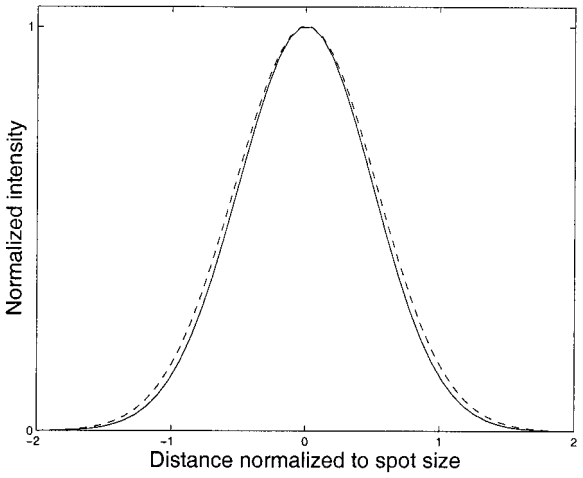

Fig. 1. Intensity profiles of a Gaussian beam (solid curve) and a beam with 95% power in the fundamental mode and 5% in the first-order Hermite–Gaussian mode (dashed curve).

form of the spatial modes. In addition, the first three methods, which use intensity measurements in a plane, require some additional information to uniquely reconstruct the weights,[13] such as intensity in an additional plane[5] or parameters, particularly the spot size, taken from knowledge of the laser cavity.[6–8] The frequency mixing method is not suitable for measuring quantitative mode weights but has been used for real-time adjustments to reduce modes in a slow-flow CO_2 laser.

We propose a technique for taking spatial coherence measurements in an arbitrary axial plane of the laser beam. These measurements comprise a mutual intensity matrix that, when decomposed with our algorithm into eigenvectors and eigenvalues, leads to the spatial modes and mode weights. This technique is particularly suited to semiconductor lasers or other lasers that have a few low-order modes that would be difficult to distinguish by intensity measurements. For example, Fig. 1 demonstrates how a Gaussian beam, shown by a solid curve, is similar to a beam, represented by a dashed curve, with 95% of the total power in the Gaussian mode and 5% in the first-order Hermite–Gaussian mode.

This paper has a total of seven sections. In Section 2 we present the theoretical background of the spatial coherence of modes. In Section 3 we explain the experimental setup that we used to extract information about spatial modes and modal weights from an arbitrary laser beam. In Section 4 we summarize the numerical analysis used to process the experimental data. Experimental results are given in Section 5. In Section 6 we look at the robustness of our method by performing computer simulations with noisy data. A short summary concludes the paper.

2. Theory

In this section we define the terms necessary to explain our experimental technique. We discuss spatial modes and spatial coherence and how the two are related.

A laser cavity can have multiple longitudinal modes and multiple spatial modes, depending on the geometry of the cavity. In this paper we are concerned with the spatial modes, sometimes called transverse modes to distinguish them from the longitudinal modes. Multiple spatial modes increase the overall spectral linewidth of the laser, and the higher-order spatial modes also increase the spatial width of the beam.

The spatial modes define a basis set for the spatial variation of the optical field of the laser beam. Using z as the distance along the optical axis, we define the x and y axes suitable to the cavity geometry. If the spatial modes are separable, the wave function can be written as

$$\psi_{mn}(x, y, z) = \phi_m(x, z)\phi_n(y, z)\exp[j\alpha_{mn}(x, y, z)], \quad (1)$$

where m is the mode index in the x direction, n is the mode index in the y direction, $\phi_m(x, z)$ and $\phi_n(y, z)$ are the amplitude functions of the spatial modes, and α_{mn} is a phase term. Amplitude functions need not be analytically known for a physical laser.

In the general case, if the spatial modes, $\phi_m(x, z)$, and $\phi_n(y, z)$ are known in any arbitrary plane, then standard diffraction theory can be used to transform the modes to the form at the exit face of the laser. In the specific case of Hermite–Gaussian modes, the modes are shape invariant during propagation so the same form is found at any point along the z axis. In contrast, the shape of the transverse mode of a thin-junction semiconductor laser is propagation dependent. This mode, perpendicular to the junction, can be modeled as $E_0 \exp(-p|x|)$ at the exit face but the far-field distribution is Lorentzian,[2] which is the Fourier transform of the exit face mode.

Laser cavities with rectangular geometries are usually described by separable Hermite–Gaussian spatial modes based on the work by Fox and Li.[14] This model is often useful because the modes are shape invariant and can be defined functionally. In addition, any arbitrary distribution can be expanded in terms of these modes. They can sometimes describe the physical characteristics of the laser well, especially when the laser is symmetric, has low diffraction losses from finite-aperture mirrors, and the dimensions of the resonator are large compared to the wavelength.[15]

Conversely, certain types of lasers have spatial modes that are not Hermite–Gaussian. These include lasers that are asymmetric either in structure or gain and lasers with small cavity dimensions.

Optical coherence, measured by the mutual coherence function, is both spatial and temporal. The mutual coherence function Γ is defined as the cross correlation between the optical fields at two points, P_1 and P_2, with a time difference of τ and is given as[16]

$$\Gamma_{12}(\tau) = \langle E(P_1, t + \tau)E^*(P_2, t)\rangle, \quad (2)$$

where the angle brackets refer to time averaging and the asterisk refers to complex conjugation. We assume cross-spectral purity which allows the mutual coherence function to be factored into spatial and temporal functions.[17] The time dependence of the mutual coherence function is sinusoidal with the frequency of the light and also has a slowly varying dependence on the time difference τ between the two sample points. If the beam is narrow band, as is the case for lasers, and τ is much less than the coherence time, the beam is temporally coherent and the sinusoidal variation is the only time dependence measured. Light that satisfies these conditions is called quasi-monochromatic[17] and is assumed throughout this paper. Under quasi-monochromatic conditions the mutual coherence function simplifies to[17]

$$\Gamma_{12}(\tau) \cong J_{12} \exp(-j2\pi\nu_0\tau), \tag{3}$$

where $J_{12} = \Gamma_{12}(0)$ is the mutual intensity and ν_0 is the center frequency of the light. The subscripts for J or Γ do not refer to the mode number but to the two points P_1 and P_2. The mutual intensity is possibly complex because of phase differences between P_1 and P_2 and can be thought of as a phasor amplitude of a spatial sinusoidal fringe.[17]

The spatial coherence of a laser beam is determined by the composition of the spatial modes. The modes are mutually uncorrelated, and each mode is completely spatially coherent at the laser mirror surface if there are no frequency degeneracies.[18] (Frequency degeneracies can arise when different combinations of m and n give a constant $m + n$. For thin-junction semiconductor lasers this is not a factor because m is constrained to be 0.)

Because the basis set of modes is complete and orthogonal, the mutual intensity in one dimension can be written as[19]

$$J_{12} = \sum_m \lambda_m \psi_m^*(x_1)\psi_m(x_2), \tag{4}$$

where λ_m is the power weight of the mth mode and ψ_m is the wave function of the mth mode. Wolf originally developed this idea in the frequency domain,[19] but under the quasi-monochromatic conditions the functions in the frequency domain are approximately delta functions and are therefore constant in the time domain.

Mutual intensity can be separated into a phase term, $\exp(j\beta_{12})$, multiplied by a real term $J_{12}{}^r$ or

$$J_{12} = J_{12}{}^r \exp(j\beta_{12}). \tag{5}$$

The real part of Eq. (4) is

$$J_{12}{}^r = \sum_m \lambda_m \phi_m(x_1)\phi_m(x_2), \tag{6}$$

where the ϕ's refer to the amplitude functions of Eq. (1). Using the homogeneous Fredholm integral equation, we obtain the weights of the modes, λ_m, from

$$\int_{-\infty}^{\infty} J_{12}{}^r \phi_m(x_2)\mathrm{d}x = \lambda_m \phi_m(x_1) \tag{7}$$

if the forms of the modes are known. This is the underlying principle that the other spatial coherence methods use for finding modal weights.

For our method we work directly with Eq. (6) to find not only the eigenvalues, or modal weights, but also the eigenvectors, which are the mode shapes. We explain our experimental procedure for measuring $J_{12}{}^r$ in Section 3.

3. Experimental Setup

To acquire data for spatial coherence, some type of optical interference is necessary. This can be implemented in a variety of ways including two pinholes,[11] a reverse phase-front interferometer,[10] a lateral shear interferometer,[20] or a twin-fiber interferometer.[12] The reverse phase-front interferometer is only capable of interfering a point x with its mirror image $-x$ so is not general enough for our purposes. Use of two pinholes introduces significant diffraction losses, and the setup is not as flexible as the twin-fiber interferometer.

The trade-offs between the lateral shear interferometer and the twin-fiber interferometer are speed of data acquisition versus resolution, data range, and fewer surfaces. The shearing interferometer is inherently faster because split portions of the beam interfere together and are recorded in parallel by a CCD camera. In contrast the twin-fiber interferometer steps through many locations in the beam with the two arms of a fiber coupler sampling the beam. Because of the slow speed of the interface with the optical powermeter, our current setup requires 2–4 h to acquire data using the twin-fiber interferometer. This could be improved by 2 orders of magnitude if a high-speed optical powermeter were used. However, even a slow powermeter has resolution and range advantages over most cameras, especially 8-bit cameras. This advantage is important to distinguish low-amplitude modes from noise. In addition, the lateral shear interferometer has surfaces from beam splitters and neutral-density filters that can introduce unwanted interference effects. For the experimental setup described in this section we used the twin-fiber interferometer.

The twin-fiber interferometer is shown in simplified form in Fig. 2. The figure is drawn in the x–z plane but we could equivalently picture it in the y–z plane by replacing x with y. The interferometer includes a single-mode fiber coupler, indicated by a thick line, with the two input arms sampling the beam on the left and one output arm emitting onto a detector, shown at the top of the figure. The input arms must be cleaved to the same length to maintain temporal coherence. To improve the visibility of the interference, a polarization controller, labeled PC in

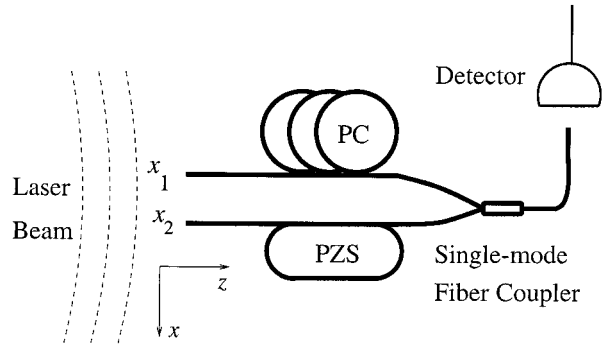

Fig. 2. Twin-fiber interferometer sampling a beam on the left, with the detector at the top right. A polarization controller (PC) is on one arm of the interferometer and a piezoelectric stretcher (PZS) is on the other arm.

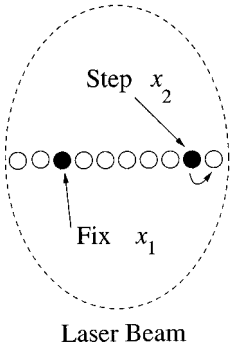

Fig. 4. For each row of data in the correlation matrix, one input is kept at a constant x_1, and the other input steps through the range of positions for x_2.

the figure, is used to match the polarization states of the two arms. One fiber arm is wrapped around a piezoelectric stretcher, labeled PZS in the figure, to vary the phase difference between the two arms. Alternatively, one fiber could be translated in z, varying the phase between points x_1 and x_2.

In our experimental setup, the input arms of the twin-fiber interferometer are on computer-controlled x–y–z stages, and the detected output is recorded by a computer. The bare fibers of the input arms are attached to grooved plates that can be mounted so that the plates lie either in the x–z or the y–z plane. The mounting of the fibers is shown in Fig. 3.

To measure the mutual intensity across the full diameter of the beam we take measurements with x_1 and x_2 at all possible combinations in the beam. This is shown graphically for one row in Fig. 4. For one row one fiber is fixed at a point x_1 in the beam, and the other fiber is then stepped through the possible range of positions for x_2. To take measurements for another row, the first fiber is moved to a different x_1, and the second fiber steps through the new range of x_2. These measurements result in a matrix of data with the steps of x_1 indicating rows and the steps of x_2 indicating columns. As the position of x_1 increases, the row index increases; and as the position of x_2 increases, the column index increases. Because of the physical limitations in the positioning of the fibers, $x_2 > x_1 + d$, where d is the diameter of the fiber.

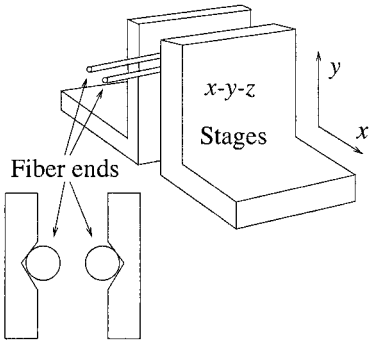

Fig. 3. Input ends of the fibers are attached to metal plates that are mounted on computer-controlled x–y–z translation stages.

The sampled fields are added by the coupler, and the total sum is squared by the detector. To produce interference fringes, we introduce a time difference τ either by stretching one of the fibers with a piezoelectric stretcher or by translating one fiber perpendicular to the wave front. Mathematically the detector intensity is described by

$$I_D = \langle [K_1 E(x_1, t+\tau) + K_2 E(x_2, t)] \times [K_1 E(x_1, t+\tau) + K_2 E(x_2, t)]^* \rangle, \quad (8)$$

where I_D is the detector intensity, the K terms quantify the loss of each fiber arm, and E refers to the field sampled at the two points x_1 and x_2. This simplifies to

$$I_D = I_1 + I_2 + 2|K_1||K_2||J_{12}{}^r|\cos(2\pi\nu_0\tau - \gamma_{12}), \quad (9)$$

where γ_{12} includes the effects of β_{12} from Eq. (5), any phase difference introduced by K_1 and K_2, and 0 or π depending on the sign of $J_{12}{}^r$. Also, I_1 refers to the intensity at point x_1 measured at the detector without interference from any field at x_2. Similarly, I_2 is the intensity at point x_2.

The amplitude of the mutual intensity is proportional to the peak-to-peak value of the detected intensity. This can be written as

$$|J_{12}{}^r| = (I_{\max} - I_{\min})/4|K_1||K_2|, \quad (10)$$

where $I_{\max}$ is the detected intensity when the cosine term is 1 and $I_{\min}$ is the intensity when the term is -1. The value of the constant $4|K_1||K_2|$ is not important because it cancels out when relative weights are determined. The phase term γ_{12} is important only if there are jumps of π rad as x_1 and x_2 vary. These jumps point to a sign change of $J_{12}{}^r$ and can be inferred from the shape of the mutual intensity.

4. Numerical Analysis

The next step after we measure the mutual intensity is to decompose the data to determine the spatial

modes and modal weights. We recognize in Eq. (6) that the mutual intensity is the sum of the outer products of the modes, or a $q \times q$ correlation matrix,

$$\mathbf{R} = \mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^T, \tag{11}$$

where the T refers to the transpose, q is the number of data points for x_1 in a complete matrix, and boldface signifies a matrix. The matrix $\mathbf{R}$ is necessarily real and symmetric with eigenvectors $\mathbf{\Phi}$, the spatial modes, and eigenvalues, which are the diagonal values of $\mathbf{\Lambda}$, that are the modal power weights.

We record mutual intensity data in the correlation matrix by varying the two sample points over one dimension x of the beam. This one-dimensional implementation is valid if the modes are separable. (Thin-junction semiconductor lasers have separable spatial modes when measured along the x or y axes.[4]) A row refers to a constant x_1, and a column refers to a constant x_2. With a complete, noise-free matrix the eigenvectors and eigenvalues can be found through standard eigenvector or singular value decomposition (SVD) routines. With enough data the form and weights of the modes can be found without any assumptions about the modes.

There are two experimental issues that make it more difficult to find the actual modes and weights: noise and insufficient data. Under the category of noise there are laser output fluctuations, detector noise, background light variations, and position errors in the fiber translations. Noise may produce false higher-order modes or may distort the shape of low-order modes, or both. Insufficient data can occur if the number of modes exceeds the rank of the matrix or can be due to experimental limitations, such as in the twin-fiber interferometer. Because the fibers cannot physically be superimposed or pass through each other, the data points are in the upper triangle of the correlation matrix. Thus the diagonal and one or more superdiagonals are missing, particularly if the diameter of the fiber is larger than the step size between data points. The symmetry of the matrix is used to fill in the lower triangle of the matrix but the diagonals are unknown. We can circumvent this experimental issue by displacing the fibers in the y axis if the modes are separable. However, there will be a reduction in the overall mutual intensity amplitude, and the intensity levels may be too low. With enough redundancy in the data, however, the modes can be recovered even without these diagonal elements.

Our method to recover the modes is known as vector Newton's method in numerical analysis or Jacobian control in the field of robotic control.[21] It is an iterative algorithm that starts with an initial guess for the modes and weights and then updates these to fit two independent conditions: namely, orthogonality of modes and a least-squares fit to the measured data. Typically we choose to find a specific small number of modes before running the algorithm. Later we rerun the algorithm for progressively larger number of modes until additional modes are clearly noise. The initial guess for the modes is the column eigenvectors found from a SVD of the incomplete matrix, with 0 in the missing diagonal elements. This initialization assures that the results are not biased by any assumptions about the modes.

The Jacobian algorithm develops from Taylor-series approximations. A first-order Taylor-series approximation of a vector function $\mathbf{f}$ with a vector argument $\mathbf{v}$ that changes by $\mathbf{\Delta v}$ is $\mathbf{f}(\mathbf{v} + \mathbf{\Delta v}) = \mathbf{f}(\mathbf{v}) + \mathbf{J}\mathbf{\Delta v}$, where $\mathbf{J}$ is the Jacobian matrix whose elements are defined by $J(i, j) = \partial f_i/\partial v_j$. Now suppose we want to find a particular $\mathbf{\Delta v}$ that causes a new desired value of $\mathbf{f}(\mathbf{v} + \mathbf{\Delta v})$. The change is found when we solve

$$\mathbf{J}\mathbf{\Delta v} = \mathbf{r} = \mathbf{f}(\mathbf{v} + \mathbf{\Delta v}) - \mathbf{f}(\mathbf{v}), \tag{12}$$

where the difference between the desired and current values of $\mathbf{f}$ is called the residual vector $\mathbf{r}$. Because the Jacobian matrix is not generally square, we must use the pseudoinverse $\mathbf{J}^+$ to solve for $\mathbf{\Delta v}$. The resulting solution is

$$\mathbf{\Delta v} = \mathbf{J}^+\mathbf{r} + (\mathbf{I} - \mathbf{J}^+\mathbf{J})\mathbf{z}, \tag{13}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{z}$ is an arbitrary vector in $\mathbf{\Delta v}$ space. The second term represents homogeneous solutions to Eq. (12) when it is underdetermined.

To find the set of modes, there are two sets of conditions and therefore two Jacobian matrices and two residual vectors. Our unknowns, the modes and weights, can be represented as the single vector $\mathbf{v}$. To do this, we first define the weighted spatial modes $\mathbf{v}_k$, where $k = 1, 2 \ldots n$, to be the same as the eigenvectors of Eq. (11) multiplied by the square root of the appropriate eigenvalues. We then form a single vector $\mathbf{v}$ of length nq by concatenating the weighted spatial modes. To represent the two conditions we define $\mathbf{J}_m$ as the data-matching Jacobian matrix and $\mathbf{J}_o$ as the orthogonality condition Jacobian matrix. The corresponding residuals are $\mathbf{r}_m$ and $\mathbf{r}_o$.

The calculated modes must be orthogonal and have a least-squares fit to the measured data. These requirements dictate that we treat the orthogonality condition as the primary criterion and the data matching as secondary. Applying the orthogonality condition in Eq. (13), we obtain $\mathbf{\Delta v} = \mathbf{J}_o{}^+\mathbf{r}_o + (\mathbf{I} - \mathbf{J}_o{}^+\mathbf{J}_o)\mathbf{z}$. Next $\mathbf{\Delta v}$ is substituted into the Jacobian equation for data matching, and we solve for a $\mathbf{z}$ which finds a least-squares fit to the data while preserving the orthogonality of the modes. When we substitute this $\mathbf{z}$ into Eq. (13) and follow the simplifying procedure of Ref. 21, the result is

$$\mathbf{\Delta v} = \mathbf{J}_o{}^+\mathbf{r}_o + [\mathbf{J}_m(\mathbf{I} - \mathbf{J}_o{}^+\mathbf{J}_o)]^+(\mathbf{r}_m - \mathbf{J}_m\mathbf{J}_o{}^+\mathbf{r}_o). \tag{14}$$

The initial guess for $\mathbf{v}$ results from the SVD of the measured matrix with 0 for the missing diagonal elements. After computing Eq. (14), we update $\mathbf{v}$ by adding $\mathbf{\Delta v}$. This process is repeated until the size of $\mathbf{\Delta v}$ or the residual errors are less than a defined criterion.

To see how we calculate the orthogonality Jacobian

matrix, first note that orthogonality requires that the inner product of two different spatial modes be 0. Mathematically, this is expressed as $\mathbf{v}_k{}^T\mathbf{v}_1 = 0$ for $k \neq l$, where $k, l = 1, 2 \ldots n$ and n is the number of modes. Because the order of k and l does not matter, there are $n(n-1)/2$ unique combinations. The size of $\mathbf{J}_o$ is therefore $n(n-1)/2$ by nq and $\mathbf{r}_o$ is a column vector of length $n(n-1)/2$. We desire that each unique combination be 0, therefore the elements of the residual vector are the negatives of the current products.

For a concrete example, consider a problem in which we are looking for three modes. The updated modes, $\mathbf{v}_1 + \Delta\mathbf{v}_1$ and $\mathbf{v}_2 + \Delta\mathbf{v}_2$, must be orthogonal. Therefore their inner product is 0, or

$$\begin{aligned} 0 &= (\mathbf{v}_1 + \Delta\mathbf{v}_1)^T(\mathbf{v}_2 + \Delta\mathbf{v}_2) \\ &= \mathbf{v}_1{}^T\mathbf{v}_2 + \Delta\mathbf{v}_1{}^T\mathbf{v}_2 + \mathbf{v}_1{}^T\Delta\mathbf{v}_2 + \Delta\mathbf{v}_1{}^T\Delta\mathbf{v}_2. \end{aligned} \quad (15)$$

The first term is considered a known residual error, and the last term is dropped in a first-order expansion. The residual vector in the next iteration will include the effects of our using only a first-order approximation in the current iteration; as the system converges, all errors go to 0. Likewise, we can expand the orthogonality conditions for the other mode combinations: modes $\mathbf{v}_1$ and $\mathbf{v}_3$ and modes $\mathbf{v}_2$ and $\mathbf{v}_3$. Combining these expressions in matrix form, we generate the orthogonality Jacobian and residual vector:

$$\mathbf{J}_o\Delta\mathbf{v} = \begin{bmatrix} \mathbf{v}_2{}^T & \mathbf{v}_1{}^T & 0 \\ \mathbf{v}_3{}^T & 0 & \mathbf{v}_1{}^T \\ 0 & \mathbf{v}_3{}^T & \mathbf{v}_2{}^T \end{bmatrix} \begin{bmatrix} \Delta\mathbf{v}_1 \\ \Delta\mathbf{v}_2 \\ \Delta\mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} -\mathbf{v}_1{}^T\mathbf{v}_2 \\ -\mathbf{v}_1{}^T\mathbf{v}_3 \\ -\mathbf{v}_2{}^T\mathbf{v}_3 \end{bmatrix} = \mathbf{r}_o. \quad (16)$$

The Jacobian system for data matching can be generated in a similar way. We want the elements of the calculated $q \times q$ matrix $\mathbf{R}$ of Eq. (11) to be as close as possible to the measured data elements of matrix $\mathbf{D}$. If $1 + 2p$ diagonals are missing from $\mathbf{D}$, then it can be shown that there are exactly $[(q-p)^2 - (q-p)]/2$ unique measured elements. The factor of 2 comes from the symmetry of the matrix. Suppose $D(i, j)$ is one of the measured elements of $\mathbf{D}$. The corresponding i, j element of $\mathbf{R}$ can be expressed as

$$R(i,j) = \sum_{k=1}^{n} v_k(i)v_k(j). \quad (17)$$

Performing a first-order expansion of Eq. (17), as in Eq. (15), we determine the following pattern for generating $\mathbf{J}_m$. Start $\mathbf{J}_m$ as a zero valued $[(q-p)^2 - (q-p)]/2$ by an nq matrix. Then we obtain the rows, indexed by h, by stepping through the combinations of $i = 1 \ldots q - (p+1)$ and $j = i + (p+1) \ldots q$ and defining, for $k = 1 \ldots n$,

$$\begin{aligned} J_m[h, (k-1)q + i] &= v_k(j), \\ J_m[h, (k-1)q + j] &= v_k(i). \end{aligned} \quad (18)$$

The elements of the residual vector $r_m(h)$ are then $D(i,j) - R(i,j)$, where h determines the choice of i and j.

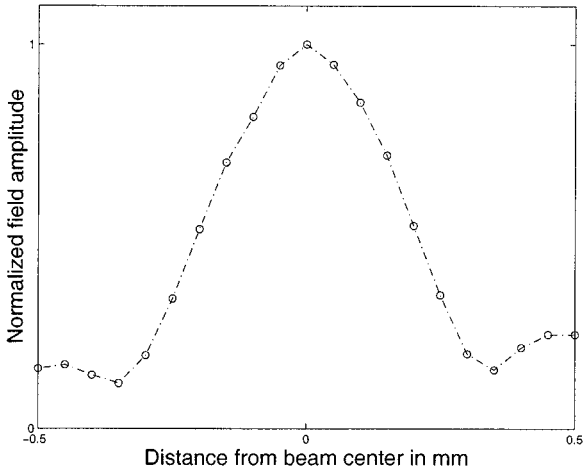

Fig. 5. Single transverse spatial mode extracted by the Jacobian method from measured data.

5. Experimental Results

To test our spatial mode extraction method we performed some experiments on an infrared, ridge-waveguide, multiple-quantum-well laser with a stripe width of 5 μm. The laser was fabricated of GaAs and AlGaAs, had a Fabry–Perot cavity, and was edge emitting. In the lateral direction, in the plane of the junction, one approximately Gaussian-shaped mode was expected with possible higher-order modes. In the transverse direction, perpendicular to the junction, only one mode was expected with a probable Lorentzian shape. Mutual intensity measurements were made separately in the lateral and transverse directions.

We found a single spatial mode in the transverse direction. This mode, shown in Fig. (5), is a good example of when one retrieves physical modes rather than predetermined analytical modes, e.g., Lorentzian. The transverse mode has some sidelobes, similar to a sinc function, that were the diffraction effects of the beam being truncated by the collimating lens. It was still fully spatially coherent because spatial coherence propagates through an optical system just as beams do. If Lorentzian or Hermite–Gaussian modes had been assumed, the sidelobes in the intensity profile would erroneously indicate multiple higher-order modes.

In the lateral direction, the data were also analyzed with the Jacobian algorithm. We extracted two modes that are plotted with +'s in Fig. 6. As a comparison, Hermite–Gaussian modes with the same relative powers are plotted with dashed curves. The relative field weights of the experimental modes are 83.4% and 16.6%, making the power weights 96.2% and 3.8%.

We analyzed the same data again requesting three modes. The strongest mode was virtually identical for either case, whether we specified two or three modes in the algorithm. The next smaller mode in weight had similar shapes for both cases and the same norms. In the presence of noise, our algorithm

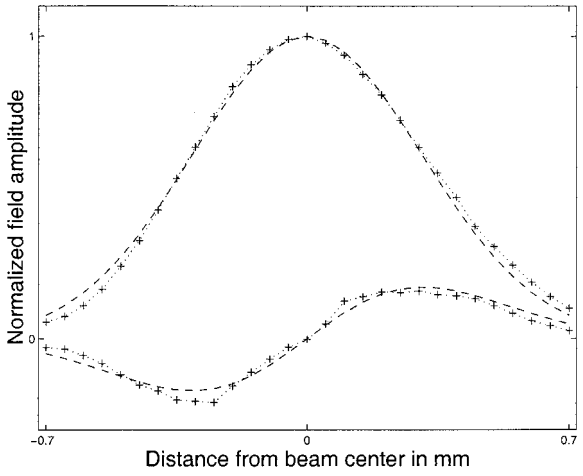

Fig. 6. Lateral spatial modes extracted by the Jacobian method from measured data indicated by +'s. Exact Hermite–Gaussian modes are shown for comparison with a dashed curve. The actual physical modes resemble, but are not exactly, Hermite–Gaussian.

forces some of the noise energy into higher-order modes if too many modes are requested. When we requested three modes from the experimental data, it resulted in a third mode with high spatial frequency variations, which is typical of noise. From the fact that the two strongest modes were consistent and from the spurious nature of the third mode, we concluded that only the two stronger modes were genuine physical modes.

An important point must be emphasized for the results shown in Fig. 6. The algorithm used no assumptions about the shape of the modes because the initial guess in the algorithm came from a SVD of the incomplete correlation matrix. The fact that the two modes are similar to Hermite–Gaussian modes reflects the geometry of the Fabry–Perot cavity and not any presuppositions of the mode shapes.

There are some differences between the physical modes shown in Fig. 6 and similar Hermite–Gaussian modes. The asymmetry in the shape of the stronger mode was probably caused by asymmetry in the laser itself. Some of the variation in the weaker mode near the peaks may be due to noise. As we discuss in Section 6, noise effects on the fundamental mode are confined mainly to the center of the beam. The weaker mode is more affected by the noise.

6. Noise Simulations

As a check on the validity of the mode extraction algorithm in the presence of noise, we performed a number of computer simulations. We generated one-dimensional spatial modes as an example of an almost single-mode Gaussian beam. We used two Hermite–Gaussian modes with relative power weights of 95% in the fundamental mode and 5% in the first-order mode. These weights correspond to 81.3% and 18.7% in terms of the electric field. The modes were normalized so the total energy was 1. The spatial width was $4w_0$, where $2w_0$ is the diameter of the fundamental mode $1/e$ field spot size. The width was normalized to $w_0 = 1$, and the range of positions was $-2 \le x \le 2$.

From the simulated modes we generated a 29×29 correlation matrix, the size corresponding to the data matrix used in experimental measurements. The noise-free correlation matrix was made up of the sum of the outer products of the exact weighted modes. We calculated the matrices of maximum and minimum intensities at each point, $\mathbf{I}_{\max}$ and $\mathbf{I}_{\min}$, using Eq. (9), with $K_1 = K_2 = 1$.

The number of missing diagonals was varied to see how well the modes could be retrieved as the available data points were reduced. For an error-free correlation matrix, five superdiagonals, or a total of 11 diagonals out of 57, could be missing, and the two remaining modes were still extracted with virtually no error. Even when three or four modes were requested, two modes were found with the correct weights and the extra modes converged to 0. For the simulated noise tests that follow, the number of missing superdiagonals was kept at two, for a total of five missing diagonals.

To simulate noise, we modeled two different types of noise. The first is background noise, with a constant variance over the whole beam, and the second is intensity-dependent noise, proportional to the detected intensity at each point. The noisy correlation matrix was then $(\tilde{\mathbf{I}}_{\max} - \tilde{\mathbf{I}}_{\min})/4$, where the tilde signifies that noise has been added.

To model the background noise we generated a matrix of noise with a uniform distribution from -0.5 to 0.5 and then multiplied it by a percentage of the peak of the $\mathbf{I}_{\max}$ matrix. Different realizations of this noise were added to $\mathbf{I}_{\max}$ and $\mathbf{I}_{\min}$. We used three levels of background noise: 0, $\pm 0.5\%$, and $\pm 1\%$.

The intensity-dependent noise was generated for the maximum and minimum intensities at each data point. Random noise matrices, uniformly distributed over -0.5 to 0.5, were multiplied element by element by a percentage of $\mathbf{I}_{\max}$ or $\mathbf{I}_{\min}$. The resulting noise was added to $\mathbf{I}_{\max}$ or $\mathbf{I}_{\min}$, as appropriate. We used 21 levels of intensity-dependent noise: 0, $\pm 0.5\%$, $\pm 1\%$, $\pm 1.5\%$, . . . $\pm 10\%$.

The results for the relative weight of the first-order mode are summarized in Table 1. The rows represent the varying amounts of intensity-dependent noise, as listed in the leftmost column. The second, third, and fourth columns show the relative weight for background noise of 0%, $\pm 0.5\%$, and $\pm 1\%$. The relative weight should be 0.05, because 5% of the power is in the first-order mode, but the effect of noise is shown by increased power being attributed to the weaker mode. The values listed are the average for the 50 different noise simulations for each combination.

The extracted modes for six of the 63 combinations of noise from Table 1 are illustrated in Fig. 7. The horizontal axes represent the position in the beam, normalized to spot size. The vertical axes show elec-

Table 1. Relative Weight of the First-Order Mode as Intensity-Dependent Noise and Background Noise are Varied[a]

Intensity Dependent (%)	Background 0%	Background ±0.5%	Background ±1%
0	0.0500	0.0502	0.0507
±0.5	0.0500	0.0501	0.0508
±1.0	0.0501	0.0503	0.0509
±1.5	0.0500	0.0502	0.0508
±2.0	0.0503	0.0505	0.0508
±2.5	0.0502	0.0504	0.0510
±3.0	0.0504	0.0509	0.0506
±3.5	0.0505	0.0507	0.0517
±4.0	0.0512	0.0511	0.0518
±4.5	0.0514	0.0508	0.0515
±5.0	0.0517	0.0513	0.0517
±5.5	0.0518	0.0517	0.0520
±6.0	0.0524	0.0512	0.0521
±6.5	0.0513	0.0530	0.0532
±7.0	0.0531	0.0526	0.0521
±7.5	0.0520	0.0539	0.0539
±8.0	0.0537	0.0545	0.0534
±8.5	0.0535	0.0532	0.0556
±9.0	0.0555	0.0547	0.0566
±9.5	0.0545	0.0551	0.0567
±10.0	0.0552	0.0560	0.0562

[a]The correct relative weight is 0.05.

tric field amplitude, normalized so that the total energy in the noise-free beam is 1. The subplots each represent a different set of conditions. Each subplot superimposes the results from 50 different noise realizations of the same conditions. A comparison of the first column of subplots with the second column shows an increase in background noise. The rows of subplots have increasing intensity-dependent noise as can be seen from top to bottom. In the first row of plots there is no intensity-dependent noise, and the background noise is ±0.5% in Fig. 7(a) and ±1% in Fig. 7(b). The second row has ±2.5% intensity-dependent noise with no background noise in Fig. 7(c) and ±1% background noise in Fig. 7(d). In the third row, the intensity-dependent noise is increased to ±5%, with no background noise in Fig. 7(e) and ±1% background noise in Fig. 7(f).

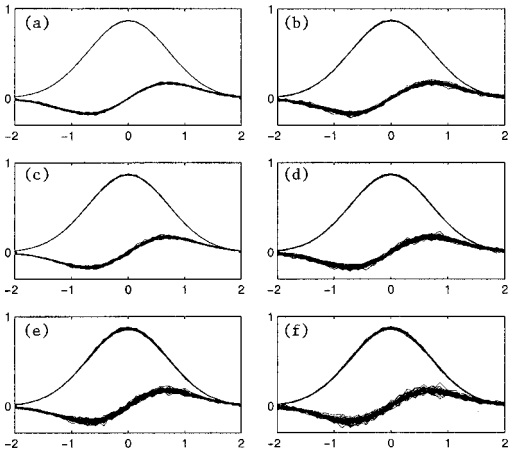

Fig. 7. Noise simulations with 50 realizations superimposed. Background noise is 0% in (c) and (e), ±0.5% in (a), and ±1% in (b), (d), and (f). Intensity-dependent noise is 0% in (a) and (b), ±2.5% in (c) and (d), and ±5% in (e) and (f).

We can see in Fig. 7 that the error in the shape of the extracted modes is most sensitive to background noise. This is reasonable because the areas of low intensity in the beam will have a low signal-to-noise ratio, as opposed to the intensity-dependent noise where the signal-to-noise ratio stays constant. Obviously, an increase in intensity-dependent noise also increases the error in the mode forms, but not as significantly. Note also that for all cases the weaker mode is more corrupted by noise than the fundamental mode.

The mode weights were found within less than a percentage point for the worst case shown in Fig. 7. The relative power weight of the first-order mode in Fig. 7(f) is within the range 0.049–0.055, where 0.05 is the noise-free weight. (On this scale, the weight of the fundamental mode is 1 minus the weight of the first-order mode.) For all the possible noise levels tested, even with ±10%, the relative weight of the weaker mode stayed within the range of 0.048–0.065. The weight of the weaker mode tended to increase as the noise levels increased because most of the noise power was included in the extraction of the weaker mode.

In the experimental results in Section 5, the overall noise was less than ±5%. Background noise was less than ±0.1%. Short-term intensity-dependent fluctuations during 5–10 min were not more than ±3% at any point in the beam. Taking into account laser power drift and possible positional errors, the noise was still less than ±5%. This most closely corresponds to the case of Fig. 7(e). Comparing Fig. 6 with Fig. 7(e) we can see that the irregularities in the weak mode are not surprising for the amount of noise. Any variation in the fundamental mode, however, would basically be confined to the center of the beam. This suggests that the asymmetry observed in the strong mode of Fig. 6 is physical.

7. Summary

We have demonstrated the ability of determining the weights and arbitrary structure of spatial modes in a few-mode laser. No prior knowledge of the laser beam, such as spot size, is necessary. The 5-μm-wide stripe, ridge-waveguide laser that we tested had two modes, with 96.2% of the power in the fundamental mode and 3.8% in a first-order mode. Using exact modes of 95% and 5%, our computer simulations show that this method is robust and extracts the correct modes with weights within one half of a percentage point for a background noise level of ±0% of maximum intensity and an intensity-dependent noise level of ±5%. The experimental measurements were made under similar noise conditions.

We thank Wright-Patterson Air Force Base which provided the laser tested under a Cooperative Research and Development Agreement.

References

1. D. Wright, P. Greve, J. Fleischer, and L. Austin, "Laser beam width, divergence and beam propagation factor–an international standardization approach," Opt. Quantum Electron. **24,** S993–S1000 (1992).
2. W. P. Dumke, "The angular beam divergence in double-heterojunction lasers with very thin active regions," IEEE J. Quantum Electron. **11,** 400–402 (1975).
3. A. Naqwi and F. Durst, "Focusing of diode laser beams: a simple mathematical model," Appl. Opt. **29,** 1780–1785 (1990).
4. X. Zeng and A. Naqwi, "Far-field distribution of double-heterostructure diode laser beams," Appl. Opt. **32,** 4491–4494 (1993).
5. A. Cutolo, T. Isernia, I. Izzo, R. Pierri, and L. Zeni, "Transverse mode analysis of a laser beam by near- and far-field intensity measurements," Appl. Opt. **34,** 7974–7978 (1995).
6. R. Borghi and M. Santarsiero, "Modal decomposition of partially coherent flat-topped beams produced by multimode lasers," Opt. Lett. **23,** 313–315 (1998).
7. M. Santarsiero, F. Gori, R. Borghi, and G. Guattari, "Evaluation of the modal structure of light beams composed of incoherent mixtures of Hermite–Gaussian modes," Appl. Opt. **38,** 5272–5281 (1999).
8. A. E. Siegman and S. W. Townsend, "Output beam propagation and beam quality from a multimode stable-cavity laser," IEEE J. Quantum Electron. **29,** 1212–1217 (1993).
9. A. Liesenhoff and F. Rühl, "An interferometric method of laser beam analysis," Rev. Sci. Instrum. **38,** 4059–4065 (1995).
10. P. Spano, "Connection between spatial coherence and modal structure in optical fibers and semiconductor lasers," Opt. Commun. **33,** 265–270 (1980).
11. E. Tervonen, J. Turunen, and A. T. Friberg, "Transverse laser-mode structure determination from spatial coherence measurements: experimental results," Appl. Phys. B **49,** 409–414 (1989).
12. L. J. Pelz and B. L. Anderson, "Practical use of the spatial coherence function for determining laser transverse mode structure," Opt. Eng. **34,** 3323–3328 (1995).
13. F. Gori, M. Santarsiero, and G. Guattari, "Coherence and the spatial distribution of intensity," J. Opt. Soc. Am. A **10,** 673–678 (1993).
14. A. G. Fox and T. Li, "Resonant modes in a maser interferometer," Bell Syst. Tech. J. **40,** 453–488 (1961).
15. H. Kogelnik and T. Li, "Laser beams and resonators," Proc. IEEE **54,** 1312–1329 (1966).
16. M. Born and E. Wolf, *Principles of Optics* (Pergamon, Oxford, UK, 1980).
17. J. W. Goodman, *Statistical Optics* (Wiley, New York, 1985).
18. E. Wolf and G. S. Agarwal, "Coherence theory of laser resonator modes," J. Opt. Soc. Am. A **1,** 541–546 (1984).
19. E. Wolf, "New theory of partial coherence in the space-frequency domain. Part I: Spectra and cross spectra of steady-state sources," J. Opt. Soc. Am. **72,** 343–351 (1982).
20. C. Iaconis and I. A. Walmsley, "Direct measurement of the two-point field correlation function," Opt. Lett. **21,** 1783–1785 (1996).
21. A. A. Maciejewski and C. A. Klein, "Obstacle avoidance for kinematically redundant manipulators in dynamically varying environments," Int. J. Robotics Res. **4,** 109–117 (1985).