Fractional Fourier transforms and their optical implementation.   II

Haldun M. Ozaktas

Department of Electrical Engineering, Bilkent University, 06533 Bilkent, Ankara, Turkey

David Mendlovic

Faculty of Engineering, Tel Aviv University, 69978 Tel Aviv, Israel

The linear transform kernel for fractional Fourier transforms is derived. The spatial resolution and the space–bandwidth product for propagation in graded-index media are discussed in direct relation to fractional Fourier transforms, and numerical examples are presented. It is shown how fractional Fourier transforms can be made the basis of generalized spatial filtering systems: Several filters are interleaved between several fractional transform stages, thereby increasing the number of degrees of freedom available in filter synthesis.
    Key words: optical information processing, Fourier optics, spatial filtering, correlation, graded-index media

1.   FRACTIONAL FOURIER TRANSFORMS

Th Fourier transform is one of the most important mathematical tools used in physical optics, optical information processing, linear systems theory, and some other areas. Previously, we defined fractional Fourier transforms[1-3] in a similar spirit in which other fractional operations, such as fractional differentiation, are defined, and we showed how this operation can be optically performed with quadratic graded-index (GRIN) media.   Here we only briefly review the notation and the basic ideas, referring the reader to the above references for details.

The ath Fourier transform of a function $f(x, y)$ is denoted as $\mathcal{F}^a[f(x,y)]$.   We require that our definition satisfy two basic postulates.   First, $\mathcal{F}^1 f$ should be identical to the common Fourier transform, defined as

$$(\mathcal{F}^1 f)(x', y') = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \exp\left[-\frac{i2\pi(x'x + y'y)}{s^2}\right] \mathrm{d}x \mathrm{d}y,$$

$$(1)$$

where s, x, x', y, and y' all have the dimensions of length. In a conventional $2f$ optical Fourier-transforming configuration,[4] x and y would denote the coordinates of the input plane, x' and y' would denote the coordinates of the Fourier plane, and $s^2 = \lambda f$ (λ = wavelength of light, f = focal length of lens).   The first pair of parentheses on the left-hand side of the above equation emphasizes that the variables (x', y') belong to the function $\mathcal{F}^1 f$ and not to f.

Our second postulate requires that $\mathcal{F}^a(\mathcal{F}^b f) = \mathcal{F}^{a+b} f$. Consistent with our two postulates, it is possible to define the fractional Fourier-transform operator physically as the operator that describes propagation along a quadratic GRIN medium.[1,2]   Such a medium has a refractive index profile given by[5]

$$n^2(r) = n_1^2[1 - (n_2/n_1)r^2],$$

$$(2)$$

where $r = (x^2 + y^2)^{1/2}$ is the radial distance from the optical axis (measured in meters) and n_1 and n_2 are the GRIN

medium parameters.   The self-modes of quadratic GRIN media are the two-dimensional Hermite–Gaussian (HG) functions, which form an orthogonal and complete basis set.   The (l, m)th member of this set is expressed as

$$\Psi_{lm}(x, y) = H_l\left(\frac{\sqrt{2}x}{\omega}\right) H_m\left(\frac{\sqrt{2}y}{\omega}\right) \exp\left(-\frac{x^2 + y^2}{\omega^2}\right),   (3)$$

where H_l and H_m are Hermite polynomials of orders l and m, respectively, and ω is a constant that is connected with the GRIN medium parameters.   For a wavelength λ, each HG mode propagates through the GRIN medium with a different propagation constant

$$\beta_{lm} = k\left[1 - \frac{2}{k}\left(\frac{n_2}{n_1}\right)^{1/2}(l + m + 1)\right]^{1/2}$$

$$\approx k - \left(\frac{n_2}{n_1}\right)^{1/2}(l + m + 1),   (4)$$

where $k = 2\pi n_1/\lambda$.   We also note for future use that $\omega = (2/k)^{1/2}(n_1/n_2)^{1/4}$.   Any function $f(x, y)$ can be expressed in terms of the HG basis set as

$$f(x, y) = \sum_l \sum_m A_{lm} \Psi_{lm}(x, y),   (5)$$

where

$$A_{lm} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{f(x, y)\Psi_{lm}(x, y)}{h_{lm}} \mathrm{d}x \mathrm{d}y,   (6)$$

with $h_{lm} = 2^{l+m}l!\, m!\, \pi\omega^2/2$.

It is worth remembering that the reason why we decompose the input function into HG functions is that they are the normal modes of quadratic GRIN media (just as plane waves are the normal modes of free space).

Now, consistent with the physical definition above, the fractional Fourier transform of $f(x, y)$ of order a is defined

mathematically as

$$\mathscr{F}^a[f(x,y)] = \sum_l \sum_m A_{lm} \Psi_{lm}(x,y) \exp(i\beta_{lm} aL). \qquad (7)$$

$L \equiv (\pi/2)(n_1/n_2)^{1/2}$ is the distance of propagation in the GRIN medium that provides the first-order (conventional) Fourier transform. $\Delta z = aL$ is the distance of propagation that gives the ath-order Fourier transform. It was shown[1] that the two postulates mentioned earlier in this section are satisfied by this definition. In the same reference we also discuss and prove some of the properties of fractional Fourier transforms, such as linearity, self-imaging parameters, and intensity shift invariance.

In this paper we further investigate the fractional Fourier-transform operation. The linear transform kernel of the fractional Fourier transform is presented in Section 2. Issues of resolution and space–bandwidth product are investigated in Section 3. In Section 4 we present some computer simulations of the fractional Fourier-transform operation. Fractional Fourier transforms can be made the basis of generalized spatial filtering operations, extending the range of operations possible with optical information processing systems. This is discussed in Section 5. Although the fractional Fourier transform was defined in relation to GRIN media, in certain cases it may be more desirable to employ bulk systems to realize them. In Section 6 we discuss how to realize a bulk system that results in the same light distribution as that of GRIN media at a finite number of planes, which will be sufficient for most applications. In the same section we also discuss other extensions and aspects that cannot be treated fully in this paper.

2. LINEAR TRANSFORM KERNEL FOR THE FRACTIONAL FOURIER-TRANSFORM OPERATION

In order to obtain a direct expression for the ath Fourier transform of a function, we combine Eqs. (6) and (7). Let us first present the one-dimensional results. The analogs of Eqs. (6) and (7) in one dimension are

$$A_l = \int_{-\infty}^{\infty} \frac{f(x)\Psi_l(x)}{h_l} dx, \qquad (8)$$

$$\mathscr{F}^a[f(x)] = \sum_l A_l \Psi_l(x) \exp(i\beta_l aL), \qquad (9)$$

respectively, where $\Psi_l(x) = H_l(\sqrt{2}x/\omega)\exp(-x^2/\omega^2)$, $h_l = 2^l l! \sqrt{\pi}\,\omega/\sqrt{2}$, and

$$\beta_l = k\left[1 - \frac{2}{k}\left(\frac{n_2}{n_1}\right)^{1/2}\left(l + \frac{1}{2}\right)\right]^{1/2} \qquad (10)$$

$$\approx k - \left(\frac{n_2}{n_1}\right)^{1/2}\left(l + \frac{1}{2}\right) \qquad (11)$$

$$= k - \frac{\pi}{2}\frac{1}{L}\left(l + \frac{1}{2}\right), \qquad (12)$$

where $L \equiv (\pi/2)(n_1/n_2)^{1/2}$ as in Section 1.

Substituting Eq. (8) into Eq. (9), one can express the ath-order fractional Fourier transform $(\mathscr{F}^a f)(x')$ as

$$\mathscr{F}^a[f(x)] = \int_{-\infty}^{\infty} f(x) B_a(x,x') dx, \qquad (13)$$

with the transform kernel being given by

$$B_a(x,x') = \sum_{l=0}^{\infty} \frac{\exp(i\beta_l aL)}{h_l} \Psi_l(x)\Psi_l(x'). \qquad (14)$$

Using Eq. (12), we obtain

$$\begin{aligned} B_a(x,x') = &\frac{\sqrt{2}}{\omega}\frac{1}{\sqrt{\pi}}\exp\left[ikaL - i\left(\frac{\pi}{4}\right)a\right] \\ &\times \exp\left[-\frac{(x^2 + x'^2)}{\omega^2}\right] \times \sum_l \frac{\exp[-ial(\pi/2)]}{2^l l!} \\ &\times H_l\left(\frac{\sqrt{2}x}{\omega}\right)H_l\left(\frac{\sqrt{2}x'}{\omega}\right). \end{aligned} \qquad (15)$$

We know that when $a = 0$ Eq. (13) must be equal to $f(x')$. In other words, we must have $B_0(x,x') = \delta(x - x')$, the Dirac delta function. Examining Eq. (15) reveals that, apart from certain phase factors up front (which do not have any effect on the transverse profile), choosing a equal to any integer multiple of 4 results in the same kernel. {This holds because $\exp[-ial(\pi/2)] = 1$ for a an integer multiple of 4.}

It is also easy to see that the kernel $B_a(x,x')$ is not space invariant; i.e., it cannot be expressed in the form $B_a(x - x')$, so that Eq. (13) cannot be reduced to a convolution integral.

The fractional Fourier-transform kernel is, of course, at the same time the kernel describing propagation in quadratic GRIN media. As given in Section 1, the distance of propagation Δz is related to a by the relation $\Delta z = aL$.

Following a similar analysis, we can express the two-dimensional kernel as

$$\begin{aligned} B_a(x,y;x',y') = &\frac{2}{\omega^2}\frac{1}{\pi}\exp\left[-\frac{(x^2 + x'^2)}{\omega^2}\right] \\ &\times \exp\left[-\frac{(y^2 + y'^2)}{\omega^2}\right] \\ &\times \exp\left[ikaL - i\left(\frac{\pi}{2}\right)a\right] \\ &\times \sum_l \sum_m \frac{\exp[-ial(\pi/2)]}{2^l l!}\frac{\exp[-iam(\pi/2)]}{2^m m!} \\ &\times H_l\left(\frac{\sqrt{2}x}{\omega}\right)H_l\left(\frac{\sqrt{2}x'}{\omega}\right)H_m\left(\frac{\sqrt{2}y}{\omega}\right) \\ &\times H_m\left(\frac{\sqrt{2}y'}{\omega}\right). \end{aligned} \qquad (16)$$

3. RESOLUTION AND SPACE–BANDWIDTH PRODUCT

Few approximations are inherent in the propagation relationships for quadratic GRIN media, which have been taken from Ref. 5. First, as is commonly done in Fourier optics, the vector nature of optical radiation has been ignored and a scalar theory has been employed. Second, it has been assumed that the refractive index changes little over distances of the order of a wavelength. This is valid provided that $(n_2/n_1)\lambda^2 \ll 1$, which, as we will see below,

must be always satisfied, anyway. The common slowly varying envelope approximation[5] has not been employed.

The third approximation is to assume that the system is of infinite transverse extent. This assumption seems to contradict the fact that the refractive index must always be greater than unity [see Eq. (2)]. Of course, all real systems must be of finite extent, a fact that has several implications. First, since the Fourier transform of the function must also be represented over a finite extent, a finite number of samples is sufficient to characterize the system in the space domain. In other words, the system has a finite resolution. Another implication is that the HG modes beyond a certain order will not be relevant for our purposes because their energy content will mostly lie outside the diameter of our medium. In this section we clarify these issues, and derive the space–bandwidth product of the system.

Since it is difficult to manufacture large index variations, and more fundamentally since we always have that $n(r) \geq 1$, we must ensure that $(n_2/n_1)r^2 \ll 1$ if such a system is to be realizable. Thus the parameters of our medium must satisfy

$$(n_2/n_1)(D/2)^2 \ll 1, \tag{17}$$

where D is the diameter of the medium.

We will now show that if this condition is satisfied, the following consequences hold:

(1) The one-dimensional space–bandwidth product of the system based on sampling considerations is $\simeq D^2/4\omega^2$.

(2) The number of relevant HG modes (those whose energy lies predominantly within the diameter D of the medium) is $\simeq D^2/4\omega^2$.

(3) The approximation for the propagation constant β_{lm} [relation (4)] is valid for these modes.

To show (1), let us assume that the transverse cross section of our GRIN medium is a square with edge length D. Circular systems would probably be more common in practice, but assuming a square will simplify our discussion and the results will differ from that for a circular system merely by geometrical factors of the order of unity.

Our original input function $f(x, y)$ is necessarily bound within the transverse extent $\Delta x = D$ and $\Delta y = D$. The first Fourier transform of $f(x, y)$ is likewise bound within the spatial frequency intervals $\Delta \nu_x = D/s^2$ and $\Delta \nu_y = D/s^2$. [s is the scale parameter appearing in Eq. (1). For a GRIN medium with parameter ω this scale parameter assumes the value $s = \sqrt{\pi}\omega$.[1]] Boundedness in either domain implies that the meaningful information content of the signal in the other domain can be recovered from its samples at the Nyquist rate.[4] Thus, for instance, boundedness in the Fourier domain implies discrete samples in the space domain at intervals of $1/\Delta \nu_x$ and $1/\Delta \nu_y$. Likewise, boundedness in the spatial domain implies discrete samples in the Fourier domain at intervals of $1/\Delta x$ and $1/\Delta y$. The total number of samples in either domain is given by

$$N = N_x N_y = (\Delta_x \Delta \nu_x)(\Delta_y \Delta \nu_y) = (D^2/s^2)^2. \tag{18}$$

The quantity N can also be interpreted as the number of spatial degrees of freedom of the system, or its space–bandwidth product.[4] Now, using the fact that $s = \sqrt{\pi}\omega$,

we obtain

$$N = \left(\frac{D^2}{\pi\omega^2}\right)^2 \tag{19}$$

as the space–bandwidth product of our system. We might guess that the space–bandwidth product for a circular system would be $N = (D^2/4\omega^2)^2$ instead. Thus the minimum resolvable dimension is approximately given by $D/(D^2/\omega^2) = \omega^2/D$.

In order to show (2) with the least amount of derivation, we will exploit the analogy with the quantum-mechanical harmonic-oscillator problem. For simplicity, we restrict ourselves to a one-dimensional discussion. In a classical harmonic oscillator the energy is given by $E = k_{ho}x_{max}^2/2$, where k_{ho} is a constant and x_{max} represents the maximum excursion from the equilibrium point. The coordinate variable x remains between the bounds $-x_{max} \leq x \leq x_{max}$. In the quantum-mechanical harmonic oscillator the energy of the lth eigenstate is given by $E_l = \hbar\omega_{ho}(l + 1/2)$, where $\hbar$ is Planck's constant and ω_{ho} is another constant.[5] For large values of l there is a correspondence between the lth eigenstate and the classical solution with the same energy $E = E_l$. Although there is always a finite probability for the absolute value of the coordinate variable to assume values greater than x_{max}, this probability is small for large l. In other words, most of the probability mass of the lth eigenstate lies within the bounds

$$|x| \leq (2E_l/k_{ho})^{1/2} \simeq (2\hbar\omega_{ho}l/k_{ho})^{1/2}. \tag{20}$$

The eigenstates of the quantum harmonic oscillator correspond to the HG modes of quadratic GRIN media. The probability distributions that we have been mentioning above must now be interpreted as energy distributions. Thus, by analogy with the discussion above, we can find the bounds within which most of the energy of the lth HG mode lies. The following relation relating the parameters of the harmonic oscillator and quadratic GRIN media will easily enable us to modify the above results for our purpose[5]:

$$\sqrt{2}/\omega = (k_{ho}/\omega_{ho}\hbar)^{1/2}. \tag{21}$$

Using this relation in inequality (20), we obtain

$$|x|^2 \leq 2l\omega^2/2 = \omega^2 l. \tag{22}$$

Thus most of the energy of the lth mode lies within the interval $[-\omega\sqrt{l}, \omega\sqrt{l}]$. The modes that are relevant for our purpose are those that lie predominantly within the interval $[-D/2, D/2]$. Thus the largest relevant order is determined by $\omega\sqrt{l} \leq D/2$, leading to

$$l = D^2/4\omega^2. \tag{23}$$

It is satisfying that the number of relevant modes is equal to the number of spatial degrees of freedom. Thus, regardless of whether we prefer to represent a signal $f(x, y)$ in terms of its samples or in terms of the coefficients of its HG components, we need the same number of real numbers.

Given (1) and (2), it is easy to demonstrate (3). Since we have shown that the relevant HG orders are those with
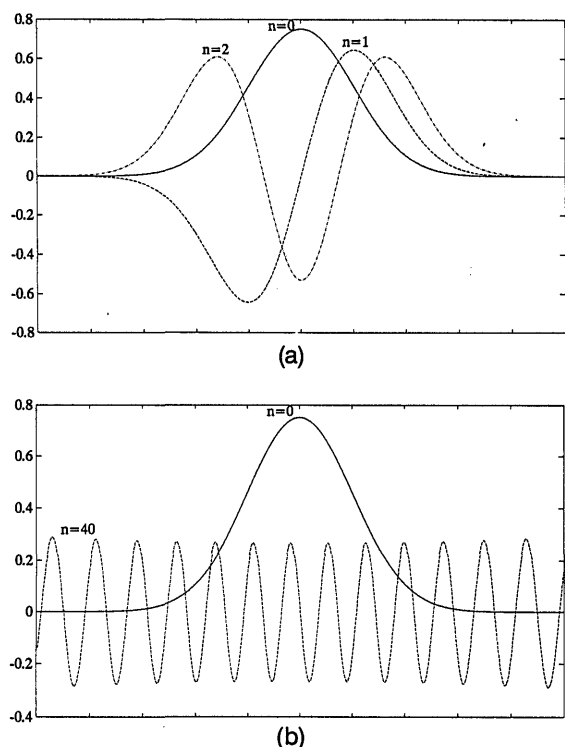
Fig. 1. Hermite–Gaussian (HG) functions of different orders n. (a) Solid curve, $n = 0$; dashed curve, $n = 1$; dashed–dotted curve, $n = 2$; (b) solid curve, $n = 0$; dashed curve, $n = 40$. The one-dimensional version of Eq. (3) has been plotted normalized so as to have unit energy.

$l, m < D^2/4\omega^2$, we can write

$$\frac{2}{k}\sqrt{\frac{n_2}{n_1}}(l + m + 1) \simeq \frac{2}{k}\left(\frac{n_2}{n_1}\right)^{1/2}(l + m) < \frac{2}{k}\left(\frac{n_2}{n_1}\right)^{1/2}\frac{2D^2}{4\omega^2}$$

$$= \frac{n_2}{n_1}\frac{D^2}{2} \ll 1, \quad (24)$$

where the last step follows from the fact that we had assumed that $(n_2/n_1)D^2 \ll 1$ for a realizable GRIN medium. Thus the approximation of β_{lm} [see relation (4)], which is crucial in deriving the exact Fourier-transform relationship,[1] has been justified.

4. COMPUTER-SIMULATED NUMERICAL EXAMPLES OF FRACTIONAL FOURIER TRANSFORMS

In order to demonstrate the fractional Fourier-transform operation, we have carried out some computer simulations for one-dimensional signals.

First, let us illustrate the nature of the HG functions with some examples. Figure 1(a) shows the three lowest orders, and Fig. 1(b) depicts the 40th order. Note that HG functions with high order can be approximated by sinusoidal functions. The horizontal axes in all the plots in this paper, except Figs. 4 and 9, are in arbitrary units.

Now, let us assume that the input function is the common rectangle function. Figure 2 shows this initial signal and the magnitude of its first-order Fourier transform. The phase information of the first Fourier transform is shown in Fig. 3.

The coefficients A_l appearing in Eq. (8) for our specific example are shown in Fig. 4. This figure demonstrates the fact that in this example harmonics of higher orders (>15) contain very little of the total energy.

Figure 5 shows the fractional Fourier transform of order 0.25. Note that this same distribution is also the fractional Fourier transform of orders $1.75, 2.25, \ldots$. Figures 6 and 7 show the fractional Fourier transforms of our input function for the orders $0.5, 1.5, \ldots$ and $0.75, 1.25, \ldots$, respectively. In these figures we can observe the rectangle function evolving into the sinc function as the parameter a is varied from 0 to 1.

Another example that we simulated is the delta function. Figure 8 shows the input signal and its second-order Fourier transform, which is an image of the input signal. This second-order Fourier transform was calculated with Eqs. (8) and (9). One can easily explain the finite width of the output by examining Fig. 9, which is the coefficient
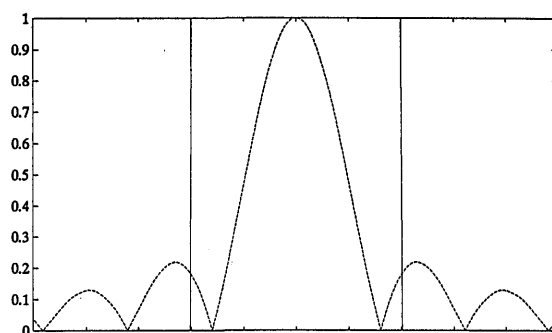

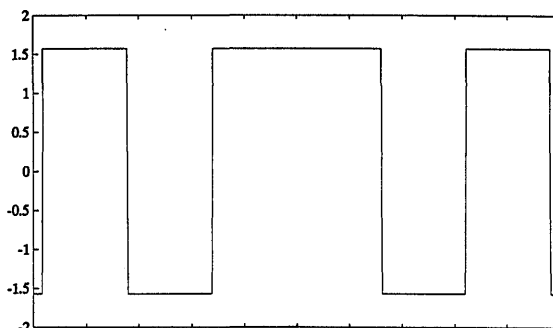Fig. 2. Solid lines, input signal; dashed curve, magnitude of its first-order Fourier transform.

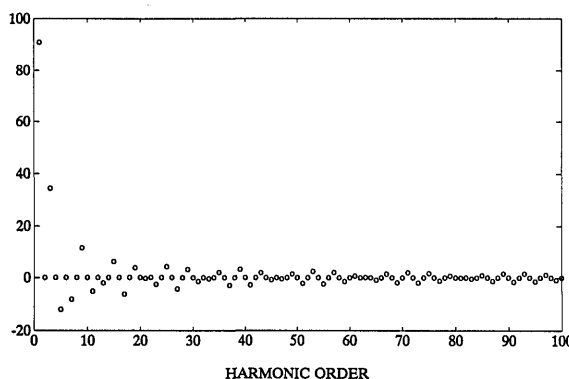Fig. 3. Phase of the first-order Fourier transform of the input signal in Fig. 2.

HARMONIC ORDER

Fig. 4. Energy distribution among the first 100 HG orders.

(a)

(b)
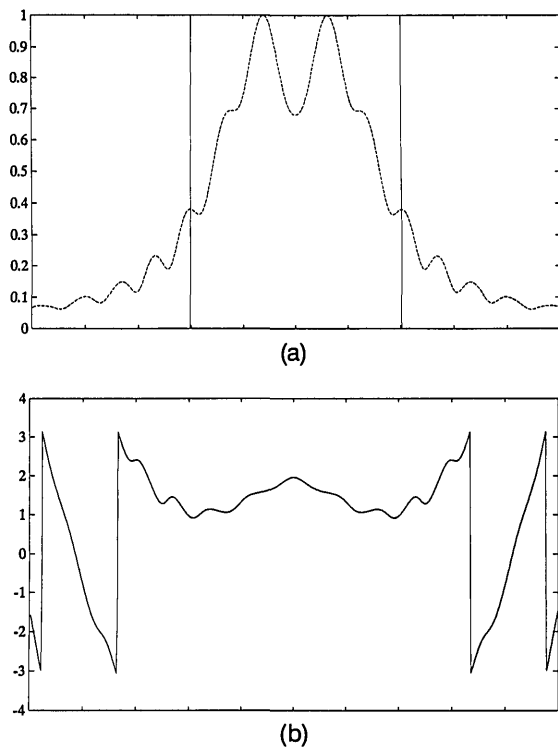
Fig. 5. Fractional Fourier transform of order $a = 0.25$. (a) Magnitude, (b) phase.

series A_l, and noting that there is a relatively large amount of energy contained in the higher orders. This being the case, the first 100 orders that we worked with in these simulations are seen to be insufficient. Remember that in Section 3 we found that the required number of HG orders is equal to the number of resolution elements. In our simulations the number of resolution elements is 1000 pixels, so that one must calculate at least this many orders to recover the delta function at the same resolution in the second Fourier plane. Doing so was less important with the rectangle function because not so much energy was concentrated in the higher orders, so that neglecting them had a much smaller impact.

Figure 10 is the magnitude of the 0.1th fractional Fourier transform of the delta function, Fig. 11 is the 0.4th order, and Fig. 12 is the first Fourier transform (which is seen to be approximately constant as expected). In these figures we can observe the delta function evolving into the constant function as the parameter a is varied from 0 to 1.

5. GENERALIZED SPATIAL FILTERING BASED ON FRACTIONAL FOURIER TRANSFORMATIONS

Fractional Fourier transforms can form the basis of generalized spatial filtering operations, extending the range of operations possible with optical information processing systems. Conventional Fourier plane filtering systems are based on a spatial filter introduced at the Fourier plane, which limits the operations achievable to linear space-invariant ones (i.e., operations that can be expressed as a convolution of the input function with a space-invariant impulse response). By introducing several filters at different fractional Fourier planes, one can im-

plement a wider class of operations. [We should note that full space-variant operations can be implemented with approaches such as (or equivalent to) vector-matrix multiplier architectures[6] or multifacet architectures,[7] but these result in a heavy penalty in terms of space–bandwidth
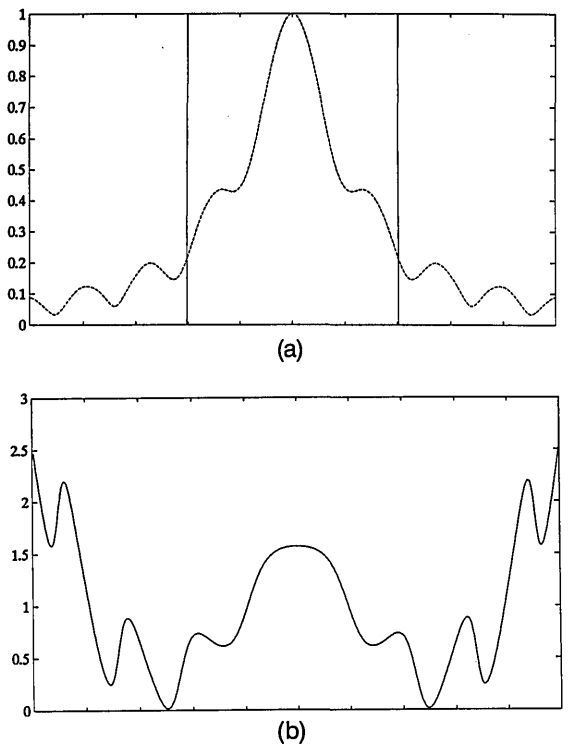

(a)

(b)

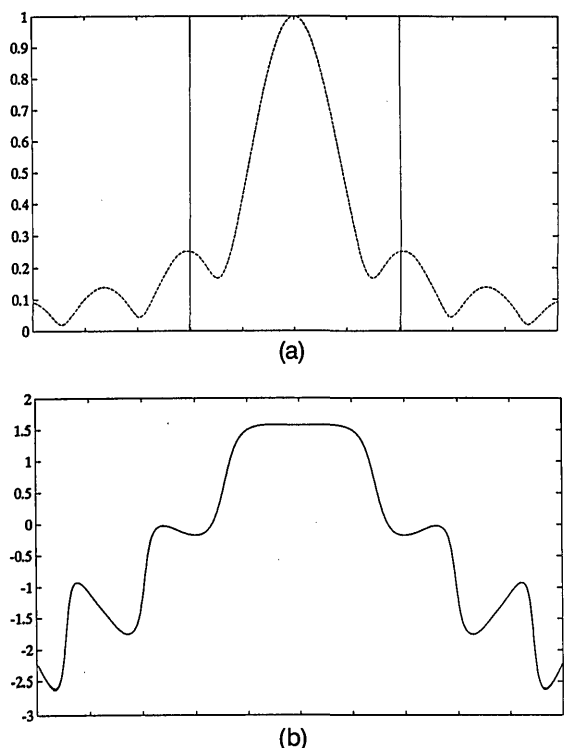Fig. 6. Fractional Fourier transform of order $a = 0.5$. (a) Magnitude, (b) phase.

(a)

(b)

Fig. 7. Fractional Fourier transform of order $a = 0.75$. (a) Magnitude, (b) phase.
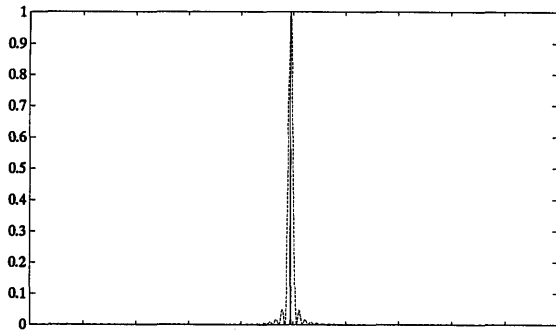
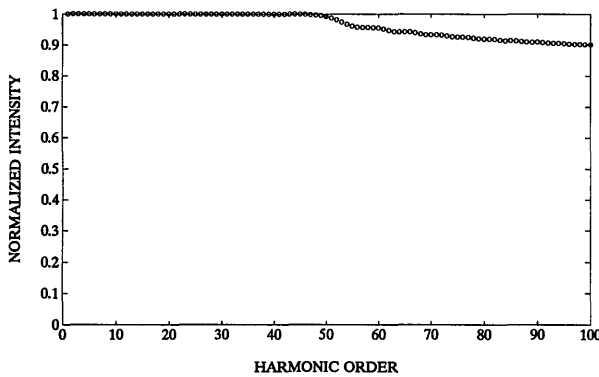Fig. 8. Solid line, input signal (delta function); dashed curve, its second-order Fourier transform.

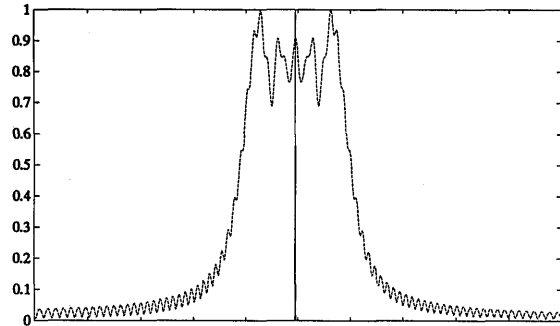Fig. 9. Energy distribution among the first 100 HG orders for the delta function.

Fig. 10. Magnitude of the $a = 0.1$th-order fractional Fourier transform of the delta function.

product utilization.[8] In this context, also see Refs. 9 and 10.]

A system for performing generalized spatial filtering would consist of a certain length of quadratic GRIN medium, followed by the first spatial filter, followed by another length of quadratic GRIN medium, followed by the second spatial filter, and so on (Fig. 13). The effect of propagation in quadratic GRIN media has been analyzed elsewhere[1] and summarized in Section 1 above. The effect of a spatial filter is simply to multiply the functional form of the light distribution incident upon it by its complex amplitude transmittance function.

One can realize such a system by cutting a certain length of quadratic GRIN medium at a finite number of places and inserting filters between the cuts. If GRIN fibers are found inconvenient to work with, it is possible to use a bulk system that produces identical light distribu-

tions at the desired planes. Such systems are described in Section 6.

Conventional 4f spatial filtering configurations are merely a very special case of the above system. Only one filter is used, and this filter is inserted at the $a = 1$st Fourier plane. The common 4f system is, by design, such that if the single filter is removed from the system, the output becomes identical to the input (apart from a coordinate inversion). That is, there is an imaging relationship between the input and output planes. However, in general, we see no *a priori* reason to impose this restriction on our systems, so that the overall length of the system becomes an additional degree of freedom during design.

For notational convenience, we will work with one-dimensional signals throughout this section. Furthermore, we will represent our functions with discrete vectors with a finite number of elements. In Section 3 we showed that the one-dimensional space–bandwidth product of a fractional Fourier transformer based on a GRIN medium of transverse extent D and parameter ω was given by $D^2/4\omega^2$. This was also the number of HG modes with a significant contribution within this diameter. Let us denote this number as N. The fact that there exists only a finite number of degrees of freedom over this finite di-
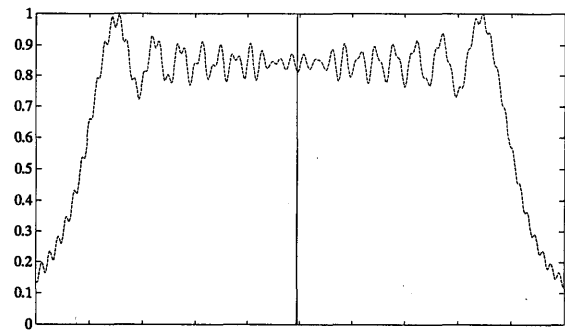

Fig. 11. Magnitude of the $a = 0.4$th-order fractional Fourier transform of the delta function.
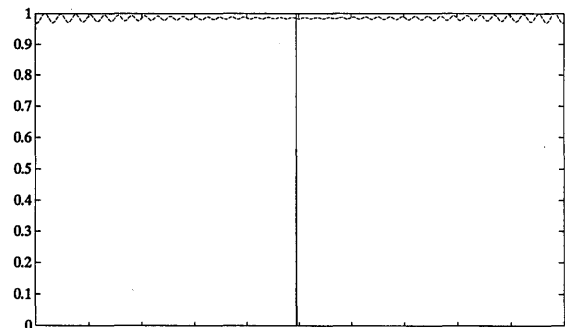

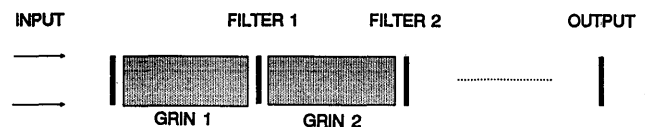Fig. 12. Magnitude of the first-order Fourier transform of the delta function.

Fig. 13. System for performing generalized spatial filtering consists of several fractional Fourier transform stages with spatial filters inserted in between.

ameter justifies the use of the samples of functions at appropriate intervals to represent the functions.

Let the function $f(x)$ be represented by its N samples in vector form as

$$\mathbf{f} = [f_1 f_2 \cdots f_N]^{\mathrm{T}}. \qquad (25)$$

T denotes the transpose operator, i.e., $\mathbf{f}$ is a column vector. $\mathbf{f}$ will be used to denote the transverse amplitude distribution of light in the GRIN medium. The distribution at axial location z will be denoted as $\mathbf{f}(z)$. If the axial location $z = z_1$ is chosen as the input of our system, $\mathbf{f}(z_1)$ will represent the input function. After propagation by a distance $\Delta z_{12} = z_2 - z_1$ to z_2 the light distribution $\mathbf{f}(z_2)$ will be given by the linear relationship

$$\mathbf{f}(z_2) = \mathbf{B}(\Delta z_{12})\mathbf{f}(z_1), \qquad (26)$$

where $\mathbf{B}$ is an $N \times N$ matrix operator that is equal to the identity matrix $\mathbf{1}$ when $\Delta z_{12} = 0$. The fact that $\mathbf{B}$ depends only on Δz_{12} and not on z_1 and z_2 individually is a consequence of the fact that the system is insensitive to a shift of origin in the z direction. Another physically self-evident property is that

$$\mathbf{B}(\Delta z_{23})\mathbf{B}(\Delta z_{12}) = \mathbf{B}(\Delta z_{23} + \Delta z_{12}) = \mathbf{B}(\Delta z_{13}). \qquad (27)$$

This additivity property is, by design, equivalent to the second postulate mentioned in Section 1. More generally,

$$\prod_i \mathbf{B}(\Delta z_{i,i+1}) = \mathbf{B}\left(\sum_i \Delta z_{i,i+1}\right). \qquad (28)$$

Evidently, $\mathbf{B}(\Delta z)$ is the discrete fractional Fourier transform operator with order $a = \Delta z/L$ and will also be denoted as $\mathbf{B}_a$.

We now seek to find the explicit form of $\mathbf{B}$ in terms of previously defined quantities. Let the sampled version of the lth member of the HG function set be denoted in vector form as $\boldsymbol{\Psi}_l = [\tilde{\Psi}_{1l}\tilde{\Psi}_{2l}\cdots\tilde{\Psi}_{Nl}]^{\mathrm{T}}$. Here we define $\tilde{\Psi}_{il} = \Psi_{i,l-1}/\sqrt{h_{l-1}}$ so as to ensure that the matrix $\boldsymbol{\Psi}$, with $\boldsymbol{\Psi}_l$ as its columns, is not only orthogonal but also orthonormal, and also so that the order of the HG set l now ranges from 1 to N rather than from 0 to $N-1$. Based on the results of Section 3, $\mathbf{f}$ can be expanded into the first N HG functions to a good approximation (since only for these first N functions is the energy predominantly contained within the extent D):

$$f_i = \sum_{l=1}^{N} \tilde{A}_l \tilde{\Psi}_{il}, \qquad (29)$$

$$\mathbf{f} = \boldsymbol{\Psi}\mathbf{A}, \qquad (30)$$

where $\tilde{A}_l = A_{l-1}\sqrt{h_{l-1}}$, $\mathbf{A} = [\tilde{A}_1\tilde{A}_2\cdots\tilde{A}_N]^{\mathrm{T}}$, and $\boldsymbol{\Psi}$ is the matrix made up of the coefficients Ψ_{il}. The above approximation is best when the coefficients A_l appearing in Eq. (29) are identical to those appearing in the infinite-summation Eq. (5) (see Appendix A). Equation (30) can also be inverted to yield

$$\mathbf{A} = \boldsymbol{\Psi}^{-1}\mathbf{f}. \qquad (31)$$

The columns of the matrix $\boldsymbol{\Psi}$ are eigenmodes of the operator $\mathbf{B}(\Delta z)$, regardless of Δz. This means that application of the operator $\mathbf{B}$ on an input function of the form

$\mathbf{f} = \boldsymbol{\Psi}_l$ produces an output function directly proportional to $\boldsymbol{\Psi}_l$. More precisely,

$$\mathbf{B}(\Delta z)\boldsymbol{\Psi}_l = \exp(i\beta_l \Delta z)\boldsymbol{\Psi}_l, \qquad (32)$$

where β_l is given by Eq. (12). The factor $\exp(i\beta_l \Delta z)$ is the eigenvalue associated with the lth eigenfunction.

Now, let the input function $\mathbf{f}(z_1)$ be expanded according to Eq. (30) with coefficients $\tilde{A}_l(z_1)$; i.e.,

$$\mathbf{f}(z_1) = \boldsymbol{\Psi}\mathbf{A}(z_1), \qquad (33)$$

$$\mathbf{A}(z_1) = \boldsymbol{\Psi}^{-1}\mathbf{f}(z_1). \qquad (34)$$

On propagation from z_1 to z_2 each eigenvalue is multiplied by the eigenvalue given above. This can be interpreted as a change in the expansion coefficient; i.e.,

$$\tilde{A}_l(z_2) = \exp(i\beta_l \Delta z_{12})\tilde{A}_l(z_1), \qquad (35)$$

$$\mathbf{A}(z_2) = \boldsymbol{\beta}(\Delta z_{12})\mathbf{A}(z_1), \qquad (36)$$

where the diagonal matrix $\boldsymbol{\beta}$ has $\exp(i\beta_l \Delta z_{12})$ as its lth diagonal element. Now, using Eq. (30), we can write

$$\mathbf{f}(z_2) = \boldsymbol{\Psi}\mathbf{A}(z_2). \qquad (37)$$

Now, combining Eqs. (34), (36), and (37), we arrive at

$$\mathbf{f}(z_2) = \boldsymbol{\Psi}\boldsymbol{\beta}(\Delta z_{12})\boldsymbol{\Psi}^{-1}\mathbf{f}(z_1), \qquad (38)$$

so that

$$\mathbf{B}(\Delta z_{12}) = \boldsymbol{\Psi}\boldsymbol{\beta}(\Delta z_{12})\boldsymbol{\Psi}^{-1}. \qquad (39)$$

The last equation may be inverted to yield

$$\boldsymbol{\Psi}^{-1}\mathbf{B}(\Delta z_{12})\boldsymbol{\Psi} = \boldsymbol{\beta}(\Delta z_{12}). \qquad (40)$$

In the parlance of matrix theory the unitary transformation effected by the matrix $\boldsymbol{\Psi}$ diagonalizes the operator $\mathbf{B}$. This is to be expected, since $\boldsymbol{\Psi}$ is constructed with the eigenvectors of $\mathbf{B}$ as its columns. The elements of the diagonal matrix $\boldsymbol{\beta}$ are, as mentioned above, the eigenvalues of $\mathbf{B}$.

Let us now discuss some properties of the above matrices. It is well known that the orthogonal HG set satisfies

$$\int \Psi_l(x)\Psi_k(x)\mathrm{d}x = \delta_{lk}h_l, \qquad (41)$$

so that

$$\int \tilde{\Psi}_l(x)\tilde{\Psi}_k(x)\mathrm{d}x = \delta_{lk}. \qquad (42)$$

If N is large, then, by the definition of Riemann integrals, the inner product

$$\boldsymbol{\Psi}_l^{\mathrm{T}}\boldsymbol{\Psi}_k = \sum_{i=1}^{N} \tilde{\Psi}_{il}\tilde{\Psi}_{ik} \qquad (43)$$

is a good approximation of the previous integral, so that $\boldsymbol{\Psi}_l^{\mathrm{T}}\boldsymbol{\Psi}_k = \delta_{lk}$, which verifies the fact that the matrix $\boldsymbol{\Psi}$ is orthonormal, i.e., $\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi} = \boldsymbol{\Psi}\boldsymbol{\Psi}^{\mathrm{T}} = \mathbf{1}$, where $\mathbf{1}$ is the identity matrix.

The formalism of this section can be interpreted in two ways. We may consider the discrete vectors as being

sampled versions of continuous functions. In this case we must bear in mind that the orthonormality of the discrete (sampled) HG function set is approximate, the approximation becoming better and better as $N \to \infty$. In this case it must be assumed that N is large enough that the orthonormality condition holds to a sufficiently good approximation. Alternatively, one may choose to interpret the following as an *a priori* discrete approach. Then, the fractional Fourier-transform operation can be effected by multiplication of a given vector by the ath power of the discrete Fourier-transform matrix. The computation of the ath power of the discrete Fourier-transform matrix, as well as its eigenvalues and eigenvectors, is discussed in Ref. 11. In this reference it is shown that a set of eigenvectors satisfying the orthonormality condition exactly can be found.

Since $\mathbf{\Psi}$ is a real orthonormal matrix, it is also unitary (i.e., its inverse is equal to the complex conjugate of its transpose). The matrix $\boldsymbol{\beta}$ is diagonal and hence symmetric $\boldsymbol{\beta}^{\mathrm{T}} = \boldsymbol{\beta}$. It is also easy to show that $\boldsymbol{\beta}$ is unitary but not Hermitian (i.e., $\boldsymbol{\beta}$ is not equal to the complex conjugate of its transpose). Since $\mathbf{B}$ is a product of three unitary matrices, it is also unitary but is also not Hermitian. Moreover, it is possible to show that $\mathbf{B}$ is also symmetric. Using Eq. (40), the reader can analytically verify the property given in Eq. (27), as well as show that $\mathbf{B}^{-1}(\Delta z) = \mathbf{B}(-\Delta z)$. Another easily verified and useful property is the following: Any diagonal matrix $\mathbf{D}$ that commutes with $\mathbf{B}$ (i.e., satisfies $\mathbf{BD} = \mathbf{DB}$) is necessarily equal to the identity matrix, perhaps multiplied by a scalar constant.

Now, referring to Fig. 13, let us assume that we choose $z = z_0$ as our input plane and $\mathbf{f}(z_0)$ as our input function. If the first filter is located at $z = z_1$, the functional form of the light distribution immediately before the filter (at $z = z_1^-$) will be denoted as $\mathbf{f}(z_1^-)$ and

$$\mathbf{f}(z_1^-) = \mathbf{B}(\Delta z_{01})\mathbf{f}(z_0), \qquad (44)$$

where $\Delta z_{01} = z_1 - z_0$. Let the continuous filter functions be denoted as $h_j(x)$ and their discrete forms as $\mathbf{h}_j = [h_{1j} h_{2j} \cdots h_{Nj}]^{\mathrm{T}}$. Since the effect of the filter is term-by-term multiplication of the components, the functional form of the light distribution immediately after the filter (at $z = z_1^+$) can be expressed as

$$\mathbf{f}(z_1^+) = \mathbf{H}_1\mathbf{f}(z_1^-), \qquad (45)$$

where $\mathbf{H}_j$ is a diagonal matrix with its ith diagonal element given by h_{ij}. Generalizing the last two equations, we obtain

$$\mathbf{f}(z_j^-) = \mathbf{B}(\Delta z_{j-1,j})\mathbf{f}(z_{j-1}^+), \qquad (46)$$

$$\mathbf{f}(z_j^+) = \mathbf{H}_j\mathbf{f}(z_j^-), \qquad (47)$$

where $j = 1, \ldots, M + 1$ if there are M filters for the first equation and $j = 1, \ldots, M$ for the second equation. Thus the overall effect of passage through M filters can be described as

$$\mathbf{f}(z_{\mathrm{out}}) = \mathbf{f}(z_{M+1}) = \mathbf{T}\mathbf{f}(z_0) = \mathbf{T}\mathbf{f}(z_{\mathrm{in}}), \qquad (48)$$

$$\mathbf{T} = \mathbf{B}(\Delta z_{M,M+1})\mathbf{H}_M\mathbf{B}(\Delta z_{M-1,M})\mathbf{H}_{M-1} \cdots \mathbf{B}(\Delta z_{1,2})\mathbf{H}_1\mathbf{B}(\Delta z_{0,1}), \qquad (49)$$

where $\mathbf{T}$ denotes the overall transfer matrix of the system. Given M, $z_1, z_2, \ldots, z_{M+1}$, and $\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_M$, one can calculate $\mathbf{T}$.

More interesting, however, is to start with a desired $\mathbf{T}$, representing an arbitrary linear transformation, and synthesize this $\mathbf{T}$ by choosing the above parameters.

If $\mathbf{T}$ is in Toeplitz form, representing a linear operation expressible as a convolution, only one filter ($M = 1$) will be sufficient. More general matrices will require a larger number of filters. The matrix $\mathbf{T}$ has $N \times N$ degrees of freedom, whereas each filter has only N. Thus a system that can realize any given matrix will need to have at least N stages. Other matrices with certain redundancies or special forms may be implemented with a smaller number of filters. Thus, given a matrix $\mathbf{T}$, the problem is to find the smallest number of filters with which it can be realized and the corresponding filter locations and values.

Since the filter locations do not represent a significant number of degrees of freedom, we may simplify by choosing $z_j = j\Delta z$, where Δz is a constant. Then Eq. (49) becomes

$$\mathbf{T} = \mathbf{B}(\Delta z)\mathbf{H}_M\mathbf{B}(\Delta z)\mathbf{H}_{M-1} \cdots \mathbf{B}(\Delta z)\mathbf{H}_1\mathbf{B}(\Delta z). \qquad (50)$$

The solution of these nonlinear equations is not straightforward and must be left out of the scope of this paper. We satisfy ourselves by identifying two problems for future research:

(1) Solve for the minimum value of M and the components of $\mathbf{H}_j$ starting from a desired matrix $\mathbf{T}$. It would also be useful to relate the minimum value of M to some characteristic number or order of the matrix.

(2) For a fixed number of filters M, give a general characterization of the $\mathbf{T}$ matrices that can be realized.

For completeness, we also provide the dual description of propagation through the sandwiched cascade of filters and GRIN media. Passage through a filter is most simply described in the original function space [Eq. (45)], since $\mathbf{H}_j$ is a diagonal matrix. On the other hand, propagation through a GRIN medium is most simply described in HG coefficient space, since the matrix $\boldsymbol{\beta}$ is diagonal. This is why we transform to HG coefficient space each time we want to propagate our function through a section of a GRIN medium. Alternatively, we may choose to construct our overall transfer function in HG coefficient space. This time we would fall back to original function space each time we want to work our way through a filter. In this case Eqs. (48) and (49) can be written in the form

$$\mathbf{A}(z_{\mathrm{out}}) = \mathbf{\Psi}^{-1}\mathbf{T}\mathbf{\Psi}\mathbf{A}(z_{\mathrm{in}}), \qquad (51)$$

$$\mathbf{T} = \mathbf{\Psi}\boldsymbol{\beta}(\Delta z_{M,M+1})\mathbf{\Psi}^{-1}\mathbf{H}_M \cdots \mathbf{\Psi}\boldsymbol{\beta}(\Delta z_{1,2})\mathbf{\Psi}^{-1}$$
$$\times \mathbf{H}_1\mathbf{\Psi}\boldsymbol{\beta}(\Delta z_{0,1})\mathbf{\Psi}^{-1}, \qquad (52)$$

$$\mathbf{S} = \boldsymbol{\beta}(\Delta z_{M,M+1})\mathbf{G}_M \cdots \boldsymbol{\beta}(\Delta z_{1,2})\mathbf{G}_1\boldsymbol{\beta}(\Delta z_{0,1}), \qquad (53)$$

where $\mathbf{S} = \mathbf{\Psi}^{-1}\mathbf{T}\mathbf{\Psi}$ and $\mathbf{G}_j = \mathbf{\Psi}^{-1}\mathbf{H}_j\mathbf{\Psi}$. In deriving these relations, we made use of Eqs. (30) and (39).

6. DISCUSSION AND EXTENSIONS

There are many aspects of fractional Fourier transforms and their optical implementation that we have been un-

able to cover in this paper, as well as several open issues. Here we briefly mention some of these so as to point to future directions.

The definition of fractional Fourier transforms can be extended to complex orders. Optical realization of complex ordered transforms would involve attenuating or amplifying media.

It would be interesting to explore what kind of fractional transformations are performed by GRIN media that have other than quadratic profiles. One can define a whole family of transformations by considering propagation in GRIN media with the refractive index profile

$$n^2(r) = n_1^2[1 - (n_p/n_1)r^p], \tag{54}$$

where p would act as an additional parameter of the transformation family. $p = 2$ corresponds to quadratic media, which we have analyzed.

As we mentioned in Section 5, a detailed treatment of how many filters are needed and how they should be synthesized to realize a given operation must be postponed for future study. Here we must satisfy ourselves by noting that this process, which in its extreme constitutes a generalization of conventional planar spatial filtering to volume spatial filtering, would ultimately be limited by noise and scatter from the several spatial filters.

It is worth making several additional comments about the generalized spatial filtering process:

(1) The system can be optically implemented with many pieces of a GRIN medium in cascade, with spatial filters sandwiched between them. However, GRIN media tend to have rather low space–bandwidth products. For this reason a bulk optics implementation of fractional Fourier transforms may be desirable. Two options are suggested. The first is to directly implement Eqs. (6) and (7). An ordinary optical correlator can be used to calculate the A_{lm} parameters [Eq. (6)], and then, with a mask and another correlator, we can reconstruct $\mathcal{F}^a[f(x, y)]$ [Eq. (7)]. This is a very complicated configuration. The second option, which is based on a recent alternative definition for the fractional Fourier transform, is more provocative.[3] This definition employs Wigner distributions. The ath fractional Fourier transform is defined as follows:

Treat the Wigner distribution of a one-dimensional function as a two-dimensional function and rotate it by $a(\pi/2)$ rad. The unique function (apart from a constant phase factor) having this Wigner distribution is the ath Fourier transform of the original function.

Based on the fact that a rotation of the Wigner distribution can be achieved simply with a system involving free-space propagation followed by a lens and then another free-space propagation, a bulk optics implementation of this definition of the fractional Fourier transform was suggested.[3] Elsewhere we show that both definitions are equivalent,[12] so that this system can be used in our generalized spatial filtering configuration.

(2) Generalized spatial filtering need not involve fractional Fourier transforms. One can cascade several conventional first-order filtering stages and investigate the range of operations possible under this restriction. Doing

so results in a special case of our formalism with $\Delta z = L$ for all the stages.

(3) Generalized spatial filtering need not utilize the fractional Fourier transform as we have defined it. For instance, one can simply take any conventional imaging system and insert spatial filters at any number of intermediate planes. A similar matrix formulation would be possible in this case as well, although it would not be so elegant, as such systems are not uniform in the longitudinal direction. (For a comparison of conventional $2f$ Fourier-transform configurations with GRIN-based systems, see Ref. 1.)

(4) An interesting question is the following: What is the minimum physical distance between two consecutive filters? In other words, what is the total number of filters that we can use simultaneously in a system with given overall length? The answer to this question should be related to the depth of focus of the system.

7. CONCLUSION

We derived the linear transform kernel for fractional Fourier transforms. We then discussed the spatial resolution and the space–bandwidth product for propagation in GRIN media in direct relation to fractional Fourier transforms and also presented some simple numerical examples. We showed how fractional Fourier transforms can be made the basis of generalized spatial filtering systems. In conventional optical spatial filtering systems the longitudinal dimension is wasted. We have suggested that this idle dimension be used by introducing many consecutive filters, thereby increasing the number of degrees of freedom available in filter synthesis.

APPENDIX A

Let the function $f(x)$ be expanded in terms of the basis set $\Psi_l(x)$ as follows:

$$f(x) = \sum_{l=0}^{\infty} A_l \Psi_l(x), \tag{A1}$$

$$A_l = \int \frac{f(x)\Psi_l(x)}{h_l} \mathrm{d}x. \tag{A2}$$

Now let $f(x)$ be approximated by the finite summation

$$f(x) \simeq \sum_{l=0}^{N-1} C_l \Psi_l(x). \tag{A3}$$

Then the mean-square error

$$\int \left[f(x) - \sum_{l=0}^{N-1} C_l \Psi_l(x) \right]^2 \exp\left(-\frac{x^2}{\omega^2}\right) \mathrm{d}x \tag{A4}$$

is minimized by the choice of $C_l = A_l$.[13]

REFERENCES

1. D. Mendlovic and H. M. Ozaktas, "Fractional Fourier transforms and their optical implementation. I," J. Opt. Soc. Am. A **10**, 1875–1881 (1993).

2. H. M. Ozaktas and D. Mendlovic, "Fourier transforms of fractional order and their optical interpretation," Opt. Commun. **101**, 163–169 (1993).

3. A. W. Lohmann, "Image rotation, Wigner rotation, and the fractional Fourier transform," J. Opt. Soc. Am. A **10**, 2181–2186 (1993).

4. A. W. Lohmann, "The space–bandwidth product, applied to spatial filtering and holography," Research Paper RJ-438 (IBM San Jose Research Laboratory, San Jose, Calif., 1967).

5. A. Yariv, *Quantum Electronics*, 3rd ed. (Wiley, New York, 1989), Chap. 2, pp. 18–26; Chap. 6, pp. 106–127.

6. J. W. Goodman, A. R. Dias, and L. M. Woody, "Fully parallel, high-speed incoherent optical method for performing discrete Fourier transforms," Opt. Lett. **2**, 1–3 (1978).

7. S. K. Case, P. R. Haugen, and O. J. Løberge, "Multifacet holographic optical elements for wave front transformations," Appl. Opt. **20**, 2670–2675 (1981).

8. D. Mendlovic and H. M. Ozaktas, "Optical-coordinate transformations and optical-interconnection architectures," Appl. Opt. **32**, 5119–5124 (1993).

9. H. I. Jeon, M. A. G. Abushagur, A.A. Sawchuk, and B. K. Jenkins, "Digital optical processor based on symbolic substitution using holographic matched filtering," Appl. Opt. **29**, 2113–2125 (1990).

10. M. W. Haney and J. J. Levy, "Optically efficient free-space folded perfect shuffle network," Appl. Opt. **30**, 2833–2840 (1991).

11. B. W. Dickinson and K. Steiglitz, "Eigenvectors and functions of the discrete Fourier transform," IEEE Trans. Acoust. Speech Signal Process. **ASSP-30**, 25–31 (1982).

12. D. Mendlovic, H. M. Ozaktas, and A. W. Lohmann, "The effect of propagation in graded index media on the Wigner distribution function and the equivalence of two definitions of the fractional Fourier transform," submitted to Appl. Opt.

13. G. Arfken, *Mathematical Methods for Physicists*, 3rd ed. (Academic, New York, 1985), pp. 712–760.