Digital Computation of the Fractional Fourier Transform

Haldun M. Ozaktas, Orhan Arikan, Member, IEEE, M. Alper Kutay, Student Member, IEEE,
and Gözde Boztepe, Member, IEEE

Abstract—An algorithm for efficient and accurate computation of the fractional Fourier transform is given. For signals with time-bandwidth product $N$, the presented algorithm computes the fractional transform in $O(N \log N)$ time. A definition for the discrete fractional Fourier transform that emerges from our analysis is also discussed.

I. INTRODUCTION

The fractional Fourier transform [1]-[4] has found many applications in the solution of differential equations [2], [3], quantum mechanics and quantum optics [5]-[11], optical diffraction theory and optical beam propagation (including lasers), and optical systems and optical signal processing [1], [13]-[25], swept-frequency filters [4], time-variant filtering and multiplexing [1], pattern recognition [26], and study of time-frequency distributions [27]. The recently studied Radon transformation of the Wigner spectrum [28]-[30] is also known to be the magnitude square of the fractional Fourier transform [1], [31]. The fractional Fourier transform has been related to wavelet transforms [1], [32], neural networks [32], and is also related to various chirp-related operations [1], [33]-[35]. It can be optically realized much like the usual Fourier transform [1], [13]-[17], [20] and, as we will show in this paper, can be simulated with a fast digital algorithm. Other applications that are currently under study or have been suggested include phase retrieval, signal detection, radar, tomography, and data compression.

In this paper, we will be concerned with the digital computation of the fractional Fourier transform. We are not only interested in a numerical method to compute the continuous transform but also in defining the discrete fractional Fourier transform and show how it can be used to approximate the continuous transform. More precisely, we will show that the samples of the continuous time fractional Fourier transform of a function can be approximately evaluated in terms of the samples of the original function in $O(N \log N)$ time, where $N$ is the time-bandwidth product of the signal.

In many of the above-mentioned applications, it is possible to improve performance by use of the fractional Fourier transform instead of the ordinary Fourier transform. Since in this paper we show that the fractional transform can be computed in about the same time as the ordinary transform, these performance improvements come at no additional cost. To give one concrete example, in some cases, filtering in a fractional Fourier domain, rather than the ordinary Fourier domain, allows one to decrease the mean square error in estimating a distorted and noisy signal [36].

In Section II, preliminaries about the fractional Fourier transform are given, including its relation to the Wigner distribution. Section III reviews some straightforward yet inefficient methods of computing the fractional Fourier transform. Fast computational algorithms are presented in Section IV, and simulation examples are given in Section V. Some alternate methods are discussed in Section VI to better situate the suggested algorithm among other possibilities. Section VII deals with the issue of defining the discrete fractional Fourier transform in some detail. The remainder of the paper constitutes concluding sections.

II. PRELIMINARIES

A. The Fractional Fourier Transform

Let $\{F^s f(x)\}$ denote the Fourier transform of $f(x)$. Integral powers $F^j$ of the operator $F \equiv F^1$ may be defined as its successive applications. Then, we have $\{F^2 f(x)\} = -f(-x)$ and $\{F^3 f(x)\} = f(x)$. The $m$th-order fractional Fourier transform $\{F^m f(x)\}$ of the function $f(x)$ may be defined for $0 < |a| < 2$ as

$$F^a [f(x)] \equiv \{F^m f(x)\} \equiv \int_{-\infty}^{\infty} B_a (x, x') f(x') \, dx',$$

$$B_a (x, x') \equiv A_\phi \exp \left[ i \pi (x^a \cot \phi - 2xx' \csc \phi + x'^2 \cot \phi) \right] ,$$

$$A_\phi \equiv \frac{\exp \left(- i \pi \sgn(\sin \phi)/4 + i\phi/2 \right)}{\left|\sin \phi \right|^{1/2}}$$

(1)

where

$$\phi \equiv \frac{a \pi}{2}$$

(2)

and $i$ is the imaginary unit. The kernel approaches $B_0 (x, x') \equiv \delta(x - x')$ and $B_{\pm 2} (x, x') \equiv \delta(x + x')$ for $a = 0$ and $a = \pm 2$, respectively. The definition is easily extended outside the interval $[-2, 2]$ by remembering that $F^{2j}$ is the identity operator for any integer $j$ and that the fractional Fourier transform operator is additive in index, that is, $F^{a_1} F^{a_2} = F^{a_1 + a_2}$. A complete set of eigenfunctions of the fractional...
Fourier transform are the Hermite–Gaussian functions
\[ F^a \{ \psi_n(x) \} = e^{-ian\pi/2} \psi_n(x), \] (3)
\[ \psi_n(x) = \frac{2^{1/4}}{\sqrt{2^n n!}} H_n(\sqrt{2\pi x}) \exp(-\pi x^2) \] (4)
where \( H_n(x) \) is the \( n \)th-order Hermite polynomial. The spectral expansion of the linear transform kernel is
\[ B_n(x,x') = \sum_{n=0}^{\infty} e^{-ian\pi/2} \psi_n(x) \psi_n(x'). \] (5)
Two and higher dimensional transforms \[ 141 - 181 \] have separable kernels so that most results easily generalize to higher dimensions. Proofs of these and other properties may be found in \[ 1 \]-\[ 4 \], \[ 14 \]-\[ 18 \], and \[ 31 \].

The \( a \)th fractional Fourier transform \( \{ F^a f(x) \} \) of the function \( f(x) \) will be abbreviatedly denoted by \( f_a(x) \).

As a word on terminology, we believe that ultimately, the term “Fourier transform” should mean, in general, “fractional Fourier transform” and that the presently standard Fourier transform be referred to as the “first-order Fourier transform.” Likewise, DFT should stand for the discrete (fractional) Fourier transform, etc., and the invention of new acronyms and abbreviations should be discouraged.

To avoid confusion, we note that for \( a = 1 \), the fractional transform reduces to the ordinary transform defined as
\[ f(x') \exp(-i2\pi xx') \] dx'.

B. Relation to the Wigner Distribution and the Concept of Fractional Fourier Domains

The Wigner distribution \( W_f(x,\nu) \) of a signal \( f(x) \) can be defined in terms of the time-domain representation \( f(x) \) of that signal as
\[ W_f(x,\nu) = \int_{-\infty}^{\infty} f(x + x'/2) f^*(x - x'/2)e^{-i2\pi xx'} dx'. \] (6)
Roughly speaking, \( W(x,\nu) \) is a function that gives the distribution of signal energy over time and frequency. Properties of the Wigner distribution may be found in \[ 37 \] and \[ 38 \]. We note the following:
\[ \int_{-\infty}^{\infty} W(x,\nu) \, d\nu = |f(x)|^2, \] (7)
\[ \int_{-\infty}^{\infty} W(x,\nu) \, dx = |f_{\nu}(\nu)|^2, \] (8)
\[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(x,\nu) \, dx \, d\nu = \text{Signal energy}. \] (9)
The Wigner distribution can also be defined in terms of any of the fractional transforms of \( f(x) \) and can be written as a function of other coordinate variables in the \( x-\nu \) plane. Thus, it should be considered to be a geometric entity associated with the signal \( f \) in the abstract and not tied to a particular representation of \( f \) in a given domain.

It is possible to show that the Wigner distribution of \( f_a(x) \) is merely a rotated version of that of \( f(x) \) \[ 1 \], \[ 4 \], \[ 18 \], \[ 31 \]
\[ W_{f_a}(x,\nu) = W_f(x \cos \phi - \nu \sin \phi, x \sin \phi + \nu \cos \phi). \] (10)

The same property can be stated in the alternative form \[ 1 \], \[ 4 \], \[ 31 \]
\[ \{ R_\phi W_f(x,\nu) \}(x_a) = |f_a(x_a)|^2 \] (11)
where \( R_\phi \) is the Radon transform operator. \( R_\phi \) takes the integral projection of the 2-D function \( W_f(x,\nu) \) onto an axis making angle \( \phi = a\pi/2 \) with the \( x \) axis. We will refer to this axis as the \( x_a \) axis or the \( a \)th fractional Fourier domain (Fig. 1). The \( x_0 \) axis is the usual time domain \( x \), and the \( x_1 \) axis is the usual frequency domain \( \nu \). Notice that (7) and (8) are special cases of this equation. In general, the projection of the Wigner distribution on the \( a \)th fractional Fourier domain gives the magnitude squared of the \( a \)th fractional Fourier transform of the original function.

There is actually nothing special about any of the continuum of domains; the privileged status we assign to the time and frequency domains can be interpreted as an arbitrary choice of the origin of the parameter \( a \). All of the fractional transforms, including the 0th transform (the function itself), are different functional representations of an abstract signal in different domains. The unitary transformation between these different representations is the fractional Fourier transform \[ 11 \].

C. Compactness in the Time Domain, the Frequency Domain, and Wigner Space

A function will be referred to as compact if its support is so. The support of a function is the subset of the real axis in which the function is not equal to zero. In other words, a function is compact if and only if its nonzero values are confined to a finite interval. It is well known that a function and its Fourier transform cannot be both compact (unless they are identically zero). In practice, however, it seems that we are always working with a finite time interval and a finite bandwidth. This discrepancy between our mathematical idealizations and
the real world is usually not a problem when we work with signals of large time-bandwidth product. The time-bandwidth product can be crudely defined as the product of the temporal extent of the signal and its (double-sided) bandwidth. It is equal to the number of degrees of freedom and the number of complex numbers required to uniquely characterize the signal among others of the same time-bandwidth product.

We will assume that the time-domain representation of our signal is approximately confined to the interval \([-\Delta t/2, \Delta t/2]\) and that its frequency-domain representation is confined to the interval \([-\Delta f/2, \Delta f/2]\). With this statement, we mean that a sufficiently large percentage of the signal energy is confined to these intervals. For a given class of functions, this can be ensured by choosing \(\Delta t\) and \(\Delta f\) sufficiently large. We then define the time-bandwidth product \(N = \Delta t \Delta f\), which is always greater than unity because of the uncertainty relation.

Let us now introduce the scaling parameter \(s\) with the dimension of time and introduce scaled coordinates \(x = t/s\) and \(\nu = f/s\). With these new coordinates, the time and frequency-domain representations will be confined to intervals of length \(\Delta t/s\) and \(\Delta f/s\). Let us choose \(s = \sqrt{\Delta t/\Delta f}\) so that the lengths of both intervals are now equal to the dimensionless quantity \(\sqrt{\Delta t/\Delta f}\), which we will denote by \(\Delta x\). In the newly defined coordinates, our signal can be represented in both domains with \(N = \Delta x^2\) samples spaced \(\Delta x^{-1} = 1/\sqrt{N}\) apart.

From now on, we will assume that this dimensional normalization has been performed and that the coordinates appearing in the definition of the fractional Fourier transform, Wigner distribution, etc., are all dimensionless quantities.

If the representation of the signal in the \(\text{ath}\) domain is confined to a certain interval around the origin, the Wigner distribution will be confined to an infinite strip perpendicular to the \(x_a\) axis defined by that interval. Thus, assuming that the representation of the signal in all domains is confined to an interval of length \(\Delta x\) around the origin, is equivalent to assuming that the Wigner distribution is confined within a circle of diameter \(\Delta x\). With this, we mean that a sufficiently large percentage of the energy of the signal is contained in that circle. For any signal, this assumption can be justified by choosing \(\Delta x\) sufficiently large. (Of course, it is in our interest to choose it as small as possible to reduce computational complexity.) For convenience, we will require \(\Delta x\) to be an integer.

Signals whose energy are not concentrated around the origin of Wigner space might be treated more efficiently than by simply choosing \(\Delta x\) large enough to include them, but this extension to the “bandpass” case is not treated in this paper.

D. The Discrete Fourier Transform

The discrete Fourier transform (DFT) is a mapping \(R^N \rightarrow R^N\). The matrix elements of this transformation are defined as

\[
F_{jk} = \frac{1}{\sqrt{N}} e^{-i2\pi jk/N}.
\]  

(12)

The DFT is related to the continuous Fourier transform as follows [39]:

Assume that a function \(f(x)\) and its Fourier transform \(F_1(\nu)\) are both confined to the interval \([-\Delta x/2, \Delta x/2]\). Then, the \(N = \Delta x^2\) samples of the Fourier transform may be found by taking the DFT of the \(N\) samples of the original function, where the sample spacing in both domains is \(1/\sqrt{N} = 1/\Delta x\).

A more precise statement of the above belongs to a class of relations known as Poisson formulas [39]:

\[
\tilde{f}_1(n) = \text{DFT} \left[ f(n) \right]\]

(13)

where \(\tilde{f}(x) = \sum_{n=-\infty}^{\infty} f(x - n\Delta x)\), and \(\tilde{f}_1(x_1) = \Sigma_{n=-\infty}^{\infty} f_1(x_1 - n\Delta x)\).

III. METHODS OF COMPUTING THE CONTINUOUS FRACTIONAL FOURIER TRANSFORM

The defining integral (1) can be rarely evaluated analytically; therefore, numerical integration is called for. Numerical integration of quadratic exponentials, which often appear in diffraction theory, require a very large number of samples if conventional methods are to be employed, due to the rapid oscillations of the kernel. The problem is particularly pronounced when \(a\) is close to 0 or \(\pm 2\). If we assume both the function and its Fourier transform to be confined to a finite interval, we can walk around this difficulty as follows: If \(a \in [0.5, 1.5]\) or \(a \in [2.5, 3.5]\), we evaluate the integral directly. If \(a \in (-0.5, 0.5)\) or \(a \in (1.5, 2.5)\), we use the property \(\mathcal{F}^a = \mathcal{F}^1 \mathcal{F}^{a-1}\), noting that in this case, the \(\mathcal{F}^a\)th transform can be evaluated directly. (Essentially similar issues are discussed in [28]-[30].)

Another method of evaluating (1) would be to use the spectral decomposition of the kernel (see (5)) [1], [14]-[16]. This is equivalent to first expanding the function \(f(x)\) as \(\sum_{n=0}^{\infty} c_n \psi_n(x)\), multiplying the expansion coefficients \(c_n\), respectively, with \(e^{-in\pi z/2}\), and summing the components.

Although both ways of evaluating the fractional Fourier transform may be expected to give accurate results, we do not consider them further since they take \(O(N^2)\) time.

IV. FAST COMPUTATION OF THE FRACTIONAL FOURIER TRANSFORM

The fractional Fourier transform is a member of a more general class of transformations that are sometimes called linear canonical transformations or quadratic-phase transforms [20]. Members of this class of transformations can be broken down into a succession of simpler operations, such as chirp multiplication, chirp convolution, scaling, and ordinary Fourier transformation. Here, we will concentrate on two particular decompositions that lead to two distinct algorithms.

A. Method 1

First, we choose to break down the fractional transform into a chirp multiplication followed by a chirp convolution followed by another chirp multiplication [17], [39].

In this approach, we assume \(a \in [-1, 1]\). Manipulating (1), we can write

\[
f_a(x) = \exp[-i\pi x^2 \tan(\phi/2)]g(x),
\]

(14)
\[ g'(x) = A_\phi \int_{-\infty}^{\infty} \exp \{i \beta (x - x')^2\} g(x') \, dx', \quad (15) \]
\[ g(x) = \exp \{-i \pi x^2 \tan (\phi/2)\} f(x) \quad (16) \]

where \( g(x) \) and \( g'(x) \) represent intermediate results, and \( \beta = \csc \phi \).

In the first step (see (16)), we multiply the function \( f(x) \) by a chirp function. As we discuss in the Appendix, the bandwidth and time-bandwidth product of \( g(x) \) can be as large as twice that of \( f(x) \). Thus, we require the samples of \( g(x) \) at intervals of \( 1/2\Delta x \). If the samples of \( f(x) \) spaced at \( 1/\Delta x \) are given to begin with, we can interpolate these and then multiply by the samples of the chirp function to obtain the desired samples of \( g(x) \). There are efficient ways of performing the required interpolation [40].

The next step is to convolve \( g(x) \) with a chirp function, as given in (15). To perform this convolution, we note that since \( g(x) \) is bandlimited, the chirp function can also be replaced with its bandlimited version without any effect, that is
\[ g'(x) = A_\phi \int_{-\infty}^{\infty} h(x - x') g(x') \, dx'. \quad (17) \]
where
\[ h(x) = \int_{-\Delta x}^{\Delta x} H(\nu) \exp (i2\pi \nu x) \, d\nu \quad (18) \]
and
\[ H(\nu) = \frac{1}{\sqrt{\beta}} e^{i\pi/4} \exp (-i\pi \nu^2 / \beta) \quad (19) \]
is the Fourier transform of \( \exp (i\pi \beta x^2) \). It is possible to express \( h(x) \) explicitly in terms of the Fresnel integral defined as
\[ F(z) = \int_0^z \exp (\pi z^2 / 2) \, dz. \quad (20) \]

Now, (15) can be sampled, giving
\[ g'(\frac{m}{2\Delta x}) = \sum_{n=-N}^{N} h\left( \frac{m-n}{2\Delta x} \right) g\left( \frac{n}{2\Delta x} \right). \quad (21) \]

This convolution can be evaluated using a fast Fourier transform.

Then, after performing the last step (see (14)), we obtain the samples of \( f_a(x) \) spaced at \( 1/\Delta x \). Since we have assumed that all transforms of \( f(x) \) are bandlimited to the interval \([-\Delta x/2, \Delta x/2]\), we finally decimate these samples by a factor of 2 to obtain samples of \( f_a(x) \) spaced at \( 1/\Delta x \).

Overall, the procedure starts with \( N \) samples spaced at \( 1/\Delta x \), which uniquely characterize the function \( f(x) \), and returns the same for \( f_a(x) \). If we let \( f \) and \( f_a \) denote column vectors with \( N \) elements containing the samples of \( f(x) \) and \( f_a(x) \), the overall procedure can be represented as
\[ f_a = F^H \Phi f, \quad (22) \]
\[ F^H = DAH_\phi AJ. \quad (23) \]

Here, \( D \) and \( J \) are matrices representing the decimation and interpolation operations [40]. \( A \) is a diagonal matrix that corresponds to chirp multiplication, and \( H_\phi \) corresponds to the convolution operation. We notice that \( F^H \) allows us to obtain the samples of the \( a \)th transform in terms of the samples of the original function, which is the basic requirement for a definition of the discrete fractional Fourier transform matrix.

If we are merely interested in computing and plotting the fractional Fourier transform of a given continuous \( f(x) \), then the decimation and interpolation steps can be eliminated.

Note that the described algorithm works for \(-1 \leq a \leq 1 \).

If \( a \) lies outside this interval, we can use the properties \( \{ F^g f \}(x) = f(x) \) and \( \{ F^2 f \}(x) = f(-x) \) to easily obtain the desired transform.

**B. Method II**

We now turn our attention to an alternative method that does not require Fresnel integrals. The defining equation for the fractional Fourier transform can be put in the form
\[ \{ F^a f \}(x) = A_\phi e^{i\alpha ax^2} \int_{-\infty}^{\infty} e^{-i2\pi \beta x' x} [e^{i\pi ax^2} f(x')] \, dx' \quad (24) \]
where \( \alpha = \cot \phi \) and \( \beta = \csc \phi \). We are again assuming that the Wigner distribution of \( f(\cdot) \) is zero outside a circle of diameter \( \Delta x \) centered around the origin. (This was discussed in detail in Section II-C.) Under this assumption, and by limiting the order \( a \) to the interval \( 0.5 \leq |a| \leq 1.5 \), the amount of vertical shear in Wigner space resulting from the chirp modulation is bounded by \( \Delta x/2 \). Then, the modulated function \( e^{i\pi ax^2} f(x') \) is bandlimited to \( \Delta x \) in the frequency domain. Thus, \( e^{i\pi ax^2} f(x') \) can be represented by Shannon’s interpolation formula
\[ e^{i\pi ax^2} f(x') = \sum_{n=-N}^{N} e^{i\alpha (n/2\Delta x)^2} f\left( \frac{n}{2\Delta x} \right) \cdot \text{sinc} \left( 2\Delta x \left( x' - \frac{n}{2\Delta x} \right) \right) \quad (25) \]

where \( N = (\Delta x)^2 \). The summation goes from \(-N \) to \( N \) since \( f(x') \) is assumed to be zero outside \([-\Delta x/2, \Delta x/2]\). By using (25) and (24) and changing the order of integration and summation, we obtain
\[ \{ F^a f \}(x) = A_\phi e^{i\alpha ax^2} \sum_{n=-N}^{N} e^{i\pi (n/2\Delta x)^2} f\left( \frac{n}{2\Delta x} \right) \]
\[ \cdot \int_{-\infty}^{\infty} e^{-i2\pi \beta x' x} \cdot \text{sinc} \left( 2\Delta x \left( x' - \frac{n}{2\Delta x} \right) \right) \, dx'. \quad (26) \]

The integral is equal to \( e^{-i2\pi \beta (n/2\Delta x)} (1/2\Delta x) \ rect \left( \beta x/2\Delta x \right) \). For the range of \( 0.5 \leq |a| \leq 1.5 \), \( \text{rect} \left( \beta x/2\Delta x \right) \) will always be equal to unity on the support \(|x| \leq \Delta x/2 \) of the transformed function. Hence, we can write
\[ \{ F^a f \}(x) = \frac{A_\phi}{2\Delta x} \sum_{n=-N}^{N} e^{i\pi ax^2} e^{-i2\pi \beta x (n/2\Delta x)} e^{i\pi (n/2\Delta x)^2} \]
\[ \cdot f\left( \frac{n}{2\Delta x} \right). \quad (27) \]
Then, the samples of the transformed function are obtained as

\[
\{F^a f\} \left( \frac{m}{2\Delta x} \right) = \frac{A_\phi}{2\Delta x} \sum_{n=-N}^{N} e^{i\pi(\alpha(m/2\Delta x)^2 - 2\beta[mn/(2\Delta x)^2] + \alpha(n/2\Delta x)^2)}
\]

which is a finite summation, allowing us to obtain the samples of the fractional transform in terms of the samples of the original function. Direct computation of this form would require \(O(N^2)\) multiplications. An \(O(N \log N)\) algorithm can be obtained as follows. We put (28) into the following form after some algebraic manipulations:

\[
\{F^a f\} \left( \frac{m}{2\Delta x} \right) = \frac{A_\phi}{2\Delta x} e^{i\pi(\alpha - \beta)(m/2\Delta x)^2} \sum_{n=-N}^{N} e^{i\pi\beta[(m-n)/2\Delta x]^2} \cdot e^{i\pi(\alpha - \beta)(n/2\Delta x)^2} f\left( \frac{n}{2\Delta x} \right).
\]

It can be recognized that the summation is the convolution of \(e^{i\pi\beta(n/2\Delta x)^2}\) and the chirp-modulated function \(f(\cdot)\). The convolution can be computed in \(O(N \log N)\) time by using the FFT. The output samples are then obtained by a final chirp modulation. Hence, the overall complexity is \(O(N \log N)\).

As in method I, by assuming appropriate \(\times 2\) interpolation and decimation, the procedure starts with \(N\) samples spaced at \(1/\Delta x\), which uniquely characterizes the function \(f(x)\) and returns the same for \(f_a(x)\). Again, letting \(f\) and \(f_a\), denoting column vectors with \(N\) elements containing the samples of \(f(x)\) and \(f_a(x)\), the overall procedure can be represented as

\[
f_a = F_{II}^a f, \quad F_{II} = DK_a J
\]

where

\[
K_a(m,n) = \frac{A_\phi}{2\Delta x} e^{i\pi(\alpha(m/2\Delta x)^2 - 2\beta[mn/(2\Delta x)^2] + \alpha(n/2\Delta x)^2)}
\]

for \(|n| \leq N\). Just as \(F_I\), we notice that \(F_{II}^a\) also allows us to obtain the samples of the \(a\)th transform in terms of the samples of the original function.

We have assumed \(0.5 \leq |a| \leq 1.5\) in deriving this algorithm. Using the index additivity property of the fractional Fourier transform, we can extend this range to all values of \(a\) easily. For instance, for the range \(0 \leq a \leq 0.5\), we observe that

\[
F^a = F^{a-1+1} = F^{a-1} F^1.
\]

Since \(0.5 \leq |a - 1| \leq 1\), the algorithm derived earlier, in conjunction with an ordinary Fourier transform, can be used in this case as well. More concretely, since both the signal and its Fourier transform is assumed to be limited to the interval \([-\Delta x/2, \Delta x/2]\), samples of the fractional Fourier transform can be related to the samples of the Fourier transform as

\[
\{F^a f\}(x) \left( \frac{m}{2\Delta x} \right) = \frac{A_\phi}{2\Delta x} \sum_{n=-N}^{N} e^{i\pi(\alpha(m/2\Delta x)^2 - 2\beta[mn/(2\Delta x)^2] + \alpha(n/2\Delta x)^2)} \cdot F[f] \left( \frac{n}{2\Delta x} \right)
\]

where \(\phi' = \pi(a - 1)/2\), \(\alpha' = \cot \phi'\), and \(\beta' = \csc \phi'\). In this case

\[
F_{II}^a = F_{II}^{a-1} F
\]

where \(F\) is the ordinary DFT matrix.

V. EXAMPLES

Both of the above presented fast methods give results that are in perfect agreement for all of the examples we tried. We prefer Method II, which does not involve Fresnel integrals, since it is faster than the first.

We first tested our algorithm by calculating the fractional Fourier transform of the Hermite-Gauss functions. We verified (4) for the first eight orders with excellent precision.

We then evaluated the Fourier transform of the common rectangle function (Fig. 2). It is interesting to see the evolution of the rectangle function continuously to the sinc function as...
I. Introduction

The fractional Fourier transform (FrFT) is a generalization of the standard Fourier transform. It is defined as the Fourier transform of a signal at a particular order, denoted by $a$, where $a$ varies from 0 to 1. This evolution is closely related to the evolution of a diffraction wave [20].

In Fig. 3, we see the fractional transform of triangular, sinusoidal, and chirp functions, respectively. The chirp was deliberately fractional transformed at the order $a$ that concentrates it maximally. Were it not for the finite extent of the chirp, it would have been transformed into a delta function. (Our numerical algorithm cannot in general deal with functions that are not of finite extent.)

Our last example is an elementary application (Fig. 4). Here, we take the fractional transform of a simple Gaussian signal on which a chirp distortion has been superimposed (Fig. 4(a)). By taking the fractional transform of this function at the appropriate order, it is possible to separate the distortion from the Gaussian (Fig. 4(b)). After masking the distorting term (Fig. 4(c)), we are able to transform back to the original domain (Fig. 4(d)), where the chirp distortion has been eliminated [1], [36].

VI. DISCUSSION OF ALTERNATE METHODS

The method presented in the previous section is only one of many possible ways to decompose the calculation. Another decomposition is the following:

$$f_x(x) = \exp[-i\pi x^2 \sec(\phi/2)]g'(x) \sec(\phi/2),$$  \hspace{1cm} (36)

The first step corresponds to a chirp convolution that can be performed efficiently by use of the FFT, as explained in the previous section. In the second step, the result of the first step is multiplied by a chirp function after a scaling operation. The overall sequence of steps is of the form convolution-scaling multiplication. Yet another decomposition is the following:

$$f_x(x) = \int_{-\infty}^{\infty} \exp[i\pi \cot(\phi/2)(x-x')] g'(x') dx'. \hspace{1cm} (37)$$

In this case, the sequence of steps are in the form scaling-multiplication convolution. The required convolution can again be computed by using the FFT.

All these methods result in the same time complexity. However, in practice, these alternate methods are not preferable because they require coordinate scaling. Not having to perform scaling is an advantage if the original data has already been sampled. This is because scaling would require additional interpolations, etc., which will require additional computation. Another advantage of Methods I and II is that they require only a moderate amount of oversampling (by a factor of two) since the shearing operations in Wigner space are of limited extent.

The above are not the only possible ways of decomposing this class of integrals [39], [41]. However, among all the

\[ g'(x) = \exp[-i\pi/2x^2 \sin(\phi)]f(x \cos(\phi/2)). \hspace{1cm} (39) \]
alternatives we have examined, the chosen method II has the advantages of resulting in $O(N \log N)$ computation, of requiring a moderate amount of oversampling by distributing the shears in the Wigner domain in the $x$ and $\nu$ directions in a balanced way, and of minimizing the need for overhead operations.

VII. THE DISCRETE FRACTIONAL FOURIER TRANSFORM

In Section V, we obtained matrices that when multiplied with the samples of a function, gave us the samples of the fractional Fourier transform of the function. For instance, take the matrix defined in (30). This property is certainly one that we desire of the discrete fractional Fourier transform matrix. Now, the matrix that maps the samples of any function into the samples of its Fourier transform is unique. Thus, it would seem that we have found the discrete fractional Fourier transform matrix.

However, remember that we have made the assumption that the Wigner distribution of the signal is confined to a circle of diameter $\Delta x$ around the origin. Thus, there might be several discrete fractional Fourier transform matrices, some of which are more elegantly expressible than the others, which gives the same result within the accuracy of this approximation. Of course, all of these matrices will yield results that are in increasingly better agreement as the signal energy contained in the circle is increasingly closer to the total energy; therefore, this definition, together with any other definitions, will approach each other in this limit.

A. Discrete Fractional Fourier Transformation

The relation between the DFT and the continuous ordinary Fourier transform was given in Section II-D. We wish a similar relation to be valid for the definition of the discrete fractional Fourier transform. That is, we wish a definition that maps the sample vector $f$ of the original function into the sample vector $f_J$ of the fractional transform. This mapping may be represented by the matrix $F^a$, which we call the discrete fractional Fourier transform matrix. In the Appendix, we have argued at length that this matrix should not be the simple functional $a$th power of the ordinary DFT matrix if we are to reach a definition of the discrete fractional Fourier transform relation that directly corresponds to the continuous transform. (The continuous fractional transform is also not a simple functional power of the continuous ordinary transform.)

Two candidates for $F^a$ that are consistent with our basic requirement are the matrices $F_J^a$ and $F^a_J$, which were derived in Section IV. In both cases, we are unable to write a simple analytic expression of the matrix elements in terms of elementary functions. However, the fractional transform matrices can be written in piecewise manner for different ranges of $a$ as the product of a small number of relatively simple matrices. For instance, considering method II, we can express $F^a_J$ as

$$F^a_J = \begin{cases} DK_{\hat{a}J}, & \text{for } 0.5 \leq |\hat{a}| \leq 1.5 \\ DK_{a-1J}, & \text{for } |\hat{a}| \leq 0.5 \text{ or } 1.5 \leq |\hat{a}| \leq 2 \end{cases}$$

where $\hat{a}$ is the unique modulo-4 equivalent of $a$ lying in the interval $[-2,2]$.

Whereas these definitions are technically acceptable, it may be possible to come up with simpler looking definitions or definitions with other analytically desirable properties. Thus, we feel that it is best not to freeze the definition until opinion is voiced by other members of the community.

VIII. CONCLUSIONS

The fractional Fourier transform is a subclass of a more general class of integral transformations characterized by quadratic complex exponential kernels. It is often not possible to evaluate these by direct numerical integration because the fast oscillations of the phase of the complex exponential would imply excessively large sampling rates. Many decompositions of these integral transformations into suboperations are possible. However, these methods might also require sampling rates that are significantly higher than the Nyquist rate, depending on the order and particular decomposition employed. This, in turn, results in greater time of computation, larger numerical inaccuracy, and the need for more memory.

The computationally efficient method ($O(N \log N)$ time) of calculating the fractional Fourier transform that we have presented requires oversampling by only a factor of 2, regardless of the order of the transform. Since the computation of the fractional transform does not take much longer than the computation of the ordinary Fourier transform, algorithms that can improve performance by employing the fractional transform instead of the ordinary transform can be implemented at no additional cost. Those working on the many applications listed in the introduction should also benefit from the presented algorithm.

We have also dealt with the definition of the discrete fractional Fourier transform. A definition satisfying the basic requirement of providing a mapping from the samples of the original function to the samples of the fractional transform was obtained. However, as is the case with the ordinary DFT, this requirement does not uniquely determine the definition of the discrete fractional transform. In this paper, we arrived at two distinct definitions, which of course give results that are equal within the intrinsically necessary approximation of assuming the signals to be limited in both time and frequency. (This is the same approximation that is made when we use the DFT to compute a continuous ordinary Fourier transform.) Ultimately, the definition to be agreed upon will assert itself with qualities such as internal consistency, simplicity, and being part of a broader analytical framework.

IX. EXTENSIONS AND FUTURE RESEARCH

We have limited our discussion to real values of $a$. Complex ordered transforms also have potential applications (such as modeling wave propagation in media with loss or gain) so that it might be of interest to generalize the simulation algorithm to this case.

The decompositions employed in this paper are also applicable to other quadratic-phase transform integrals [20]. Thus, by relatively straightforward extension of the method of this paper, $O(N \log N)$ algorithms for these integrals can also be found.
There remains much to be worked out in the theory of the discrete fractional Fourier transform. It would be quite desirable to find an expression for the discrete transform that is related to the discrete Wigner distribution through a relation resembling (11). Furthermore, to tie everything together, it would be of interest to define discrete versions of the Hermite–Gauss functions in such a way that they are the eigenvectors of the discrete fractional Fourier transform matrix. Whatever the final definition of the discrete fractional Fourier transform matrix that emerges from future work, we know that it will give the same result in simulating a continuous transform as the matrix defined in this paper within the accuracy limited by representing a continuous function with a discrete vector.

APPENDIX A

EFFECT OF CHIRP MULTIPLICATION AND CONVOLUTION ON COMPACT SIGNALS

We will refer to a signal as compact if its Wigner distribution function is compact, that is, if it is confined to a circle of some finite diameter \( \Delta x \). As discussed earlier, we are content with having a sufficiently large percentage of the signal energy confined to this circle (Fig. 5(a)).

Now, let us assume that the time domain representation of our signal \( f(x) \) is multiplied by the chirp function \( \exp \left[ i \pi \tan \left( \frac{\phi}{2} \right) x^2 \right] \), where \( -\pi/2 \leq \phi \leq \pi/2 \). It is known that the effect of this is to shear the Wigner distribution in the \( \nu \) direction, as shown in Fig. 5(b) [37]. (Mathematically, this corresponds to convolving the original Wigner distribution with the Wigner distribution of the chirp function \( \delta [\nu - \tan (\phi/2) x] \) in the \( \nu \) variable.) We see that the support of the Wigner distribution remains compact. Whatever percentage of the signal was confined to the bandwidth \( \Delta x \), we are now assured that it is confined to the bandwidth \( [1 + |\tan (\phi/2)|] \Delta x \), as found above.

We can convince ourselves of the same fact by the following intuitive argument. The instantaneous frequency of the chirp function in question is \( \tan (\phi/2) x \). If this chirp is confined to the interval \( [-\Delta x/2, \Delta x/2] \), the largest instantaneous frequency will be \( |\tan (\phi/2)| \Delta x/2 \). If we take this to be approximately equal to the bandwidth of the chirp, then the double-sided bandwidth is \( |\tan (\phi/2)| \Delta x \). Multiplying this chirp with \( f(x) \) results in a convolution of their Fourier transforms, resulting in an overall double-sided bandwidth of \( [1 + |\tan (\phi/2)|] \Delta x \), as found above.

A similar argument holds for convolution of a function with a chirp. This time, shearing of the Wigner distribution in the \( x \) direction is involved [37].

The need to limit the extent of vertical or horizontal shearing of the Wigner distribution has also been discussed in [28]–[30] in a similar context.

APPENDIX B

DEFINITION OF THE DISCRETE FRACTIONAL FOURIER TRANSFORM AS THE FUNCTIONAL POWER OF THE DFT MATRIX

In this Appendix, we discuss a definition of the discrete fractional Fourier transform given by Dickinson and Steiglitz some time ago [44], [45]. According to their definition, the \( \alpha \)-order discrete fractional transform matrix is found by taking the \( \alpha \)-th power of the DFT matrix \( F \). The \( \alpha \)-th power of the DFT matrix \( F^\alpha \) is found by standard procedures, and the ambiguity in taking fractional powers is resolved by choosing the principal roots. Then, one finds that

\[
F^\alpha = \sum_{j=0}^{3} \alpha_j(\alpha) F^j
\]

where explicit expressions for the \( \alpha_j(\alpha) \) are given in [44].

We will now argue that the above definition of the discrete fractional Fourier transform is not the one that corresponds
to the continuous definition given in (1). One of the central properties of the continuous fractional transform is given in (3). The corresponding property for the DFT would look like

$$F^n \frac{\hat{\psi}}{n} = e^{i\pi n/2} \frac{\hat{\psi}}{n}$$  \hspace{1cm} (42)

where \( \hat{\psi} \)'s are the eigenvectors of the DFT matrix. We now prove that the discrete fractional Fourier transformation defined as in (41) cannot have the property given in (42).

**Proposition 1:** Given a set of eigenvectors \( \hat{\psi} \) of the DFT matrix \( F \) such that

$$F \hat{\psi} = e^{-i\pi k/2} \hat{\psi}$$  \hspace{1cm} (43)

if the discrete fractional Fourier transformation is defined as in (41), then (42) will not hold true for noninteger values of \( a \).

**Proof:** For integer powers of the DFT matrix, the following holds true:

$$F^k \hat{\psi} = e^{-i\pi k/2} \hat{\psi}$$  \hspace{1cm} (44)

where \( k \) is an integer. This can be seen by repeated application of the DFT matrix to both sides of (43). Assuming that the discrete fractional Fourier transformation matrix given by (41) satisfies (42), we obtain

$$F^n \hat{\psi} = \hat{\psi} = (\alpha_0(a) + \alpha_1(a) + \alpha_2(a) + \alpha_3(a)) \hat{\psi}$$  \hspace{1cm} (45)

$$F^n \hat{\psi} = e^{-2i\pi a} \hat{\psi} = (\alpha_0(a) + \alpha_1(a) + \alpha_2(a) + \alpha_3(a)) \hat{\psi}$$  \hspace{1cm} (46)

Since \( \hat{\psi} \)'s are nonzero, these equations imply the following contradiction for noninteger values of \( a \):

$$1 = \alpha_0(a) + \alpha_1(a) + \alpha_2(a) + \alpha_3(a)$$  \hspace{1cm} (47)

$$e^{-2i\pi a} = \alpha_0(a) + \alpha_1(a) + \alpha_2(a) + \alpha_3(a).$$  \hspace{1cm} (48)

Thus, (41) and (42) are inconsistent. Hence, the proposition is true.

Furthermore, if the chosen discrete fractional Fourier transform matrix satisfies (42), the resulting matrix is not a function of the common DFT matrix:

**Proposition 2:** Given a set of eigenvectors \( \hat{\psi} \) of the DFT matrix \( F \) such that

$$F \hat{\psi} = e^{-i\pi k/2} \hat{\psi}$$  \hspace{1cm} (49)

if \( F^n \hat{\psi} = e^{i\pi n/2} \hat{\psi} \), then there is no function \( T(\cdot) \) that satisfies

$$F^n = T(F).$$  \hspace{1cm} (50)

**Proof:** For any function \( T(\cdot) \) of the DFT matrix \( F \), there exists an equivalent definition in the spectrum of \( F \) [42], [43]:

$$T(F) = \sum_{k=0}^{3} \beta_k F^k.$$  \hspace{1cm} (51)

In Proposition 1, we have proved that the right-hand side cannot satisfy the property given in (42). Thus, the left-hand side cannot satisfy it either, proving Proposition 2.

In conclusion, whereas this definition of the fractional Fourier transform may be useful for certain applications [44], [45], it is not the discrete version of that defined in (1).

ACKNOWLEDGMENT

The authors would like to thank D. Mendlovic of Tel-Aviv University for many useful discussions.

REFERENCES


1 1


[27] J. R. Fonollosa and C. L. Nikias, "A new positive time-frequency distri-


Haldun M. Ozaktas received the B.S. degree from Middle East Technical University, Ankara, Turkey, in 1982 and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, in 1988 and 1991 respectively, all in electrical engineering. During his graduate studies, he was a Research Assistant to J. W. Goodman. He joined Bilkent University, Ankara, Turkey in 1991, where he is presently Associate Professor of Electrical Engineering. In 1992, he was with A. W. Lohmann at the University of Erlangen-Nürnberg, Germany, as an Alexander von Humboldt Foundation Postdoctoral Fellow. Over the summer of 1994, he worked as a Consultant for D. A. B. Miller at Bell Laboratories, Holmdel, NJ. He is the author or coauthor of over 40 refereed journal articles and numerous conference presentations and papers, seven of which were invited. His academic interests include optical information processing, signal and image processing, and optoelectronic and optically interconnected computing systems.

Orhan Arikan (M'91) was born in 1964 in Manisa, Turkey. In 1986, he received the B.Sc. degree in electrical and electronics engineering from the Middle East Technical University, Ankara, Turkey. He received both the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois, Urbana-Champaign, in 1988 and 1990, respectively.

Following his graduate studies, he worked for three years as a Research Scientist at Schlumberger-Doll Research, Ridgefield, CT. During this time, he was involved in the inverse problems and fusion of measurements with multiple modality. Since 1993, he has been an Assistant Professor of Electrical and Electronics Engineering at Bilkent University, Ankara, Turkey. His current research interests are in adaptive signal processing, time-frequency analysis, inverse problems, and array signal processing.

Dr. Arikan is the acting chairman of Turkey Chapter of the IEEE Signal Processing Society.

M. Alper Kutay (S'95) was born in 1972 in Konya, Turkey. After finishing Ankara Science High School, he received the B.S. and M.S. degrees in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 1993 and 1995, respectively.

During his M.S. study, he was a Teaching and Research Assistant. He is currently working toward the Ph.D. degree under the supervision of Prof. H. M. Ozaktas in the Department of Electrical and Electronics Engineering at Bilkent University. His current research areas include time-frequency representations, fractional Fourier transform, and time-varying filtering. He is a scholar of the BDP program of TUBITAK.

Mr. Kutay was awarded third place in the 1994 IEEE region 8 student paper contest for his work entitled "An Adaptive Speckle Suppression Filter for Medical Ultrasonic Imaging."

Gözde Bozdagi (M'95) was born in Ankara, Turkey, in 1967. She received the B.S. degree from Middle East Technical University, Ankara, Turkey, in 1988 and the M.S. and Ph.D. degrees from Bilkent University, Ankara, Turkey, in 1990 and 1994, respectively, all in electrical and electronics engineering.

From March 1994 to October 1995, she was a visiting research associate at the Electrical Engineering Department of the University of Rochester, Rochester, NY. She is currently a member of the research and technology staff at the Digital Imaging Technology Center, Xerox Corporation, Webster, NY. Her current research interests include image segmentation, video image processing, content-based indexing, motion estimation, and noise filtering.