<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Fast numerical algorithm for the linear canonical transform</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Hennelly, Bryan M.; Sheridan, John T.</td>
</tr>
<tr>
<td><strong>Publication date</strong></td>
<td>2005-05-01</td>
</tr>
<tr>
<td><strong>Publication information</strong></td>
<td>Journal of the Optical Society of America A, 22 (5): 928-937</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>Optical Society of America</td>
</tr>
<tr>
<td><strong>Link to online version</strong></td>
<td><a href="http://dx.doi.org/10.1364/JOSAA.22.000928">http://dx.doi.org/10.1364/JOSAA.22.000928</a></td>
</tr>
<tr>
<td><strong>Item record/more information</strong></td>
<td><a href="http://hdl.handle.net/10197/3368">http://hdl.handle.net/10197/3368</a></td>
</tr>
<tr>
<td><strong>Publisher's statement</strong></td>
<td>This paper was published in Journal of the Optical Society of America A and is made available as an electronic reprint with the permission of OSA. The paper can be found at the following URL on the OSA website: <a href="http://www.opticsinfobase.org/abstract.cfm?URI=JOSAA-22-5-928">http://www.opticsinfobase.org/abstract.cfm?URI=JOSAA-22-5-928</a>. Systematic or multiple reproduction or distribution to multiple locations via electronic or other means is prohibited and is subject to penalties under law.</td>
</tr>
<tr>
<td><strong>Publisher's version (DOI)</strong></td>
<td><a href="http://dx.doi.org/10.1364/JOSAA.22.000928">http://dx.doi.org/10.1364/JOSAA.22.000928</a></td>
</tr>
</tbody>
</table>

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)
**Fast numerical algorithm for the linear canonical transform**

Bryan M. Hennelly and John T. Sheridan

Department of Electronic and Electrical Engineering, Faculty of Architecture and Engineering, University College Dublin, Belfield, Dublin 4, Republic of Ireland

Received August 26, 2004; revised manuscript received November 19, 2004; accepted November 22, 2004

The linear canonical transform (LCT) describes the effect of any quadratic phase system (QPS) on an input optical wave field. Special cases of the LCT include the fractional Fourier transform (FRT), the Fourier transform (FT), and the Fresnel transform (FST) describing free-space propagation. Currently there are numerous efficient algorithms used for purposes of numerical simulation in the area of optical signal processing to calculate the discrete FT, FRT, and FST. All of these algorithms are based on the use of the fast Fourier transform (FFT). In this paper we develop theory for the discrete linear canonical transform (DLCT), which is to the LCT what the discrete Fourier transform (DFT) is to the FT. We then derive the fast linear canonical transform (FLCT), an \( N \log N \) algorithm for its numerical implementation by an approach similar to that used in deriving the FFT from the DFT. Our algorithm is significantly different from the FFT, is based purely on the properties of the LCT, and can be used for FFT, FRT, and FST calculations and, in the most general case, for the rapid calculation of the effect of any QPS. © 2005 Optical Society of America

**OCIS codes:** 080.2730, 100.2000, 070.4560, 200.2610, 200.3050, 200.4560, 200.4740.

1. INTRODUCTION

The one-dimensional linear canonical transform\(^1\text{-}^3\) (LCT) is a three-parameter (\(\alpha, \beta, \gamma\)) class of linear integral transform. This can be further generalized to a five-parameter transform, the special affine Fourier transform,\(^4\text{-}^6\) in which the additional two parameters are shifts in the spatial and spatial-frequency domains that have no effect on the numerical discussions in this paper. The LCT is a unitary transform and includes as special cases the Fourier transform\(^6\) (FT), the Fresnel transform\(^7\) (FST), the fractional Fourier transform\(^2\) (FRT), and the operations of scaling (magnification) and chirp multiplication (thin lenses).

Optical systems implemented with an arbitrary number of thin lenses and propagation through free space in the Fresnel approximation or through sections of graded-index (GRIN) media belong to the class of systems known as quadratic phase systems\(^7\) (QPSs). The effect of all QPSs can be described by use of the three-parameter LCT. The kernel of the LCT can be shown to be equivalent to chirp multiplication\(^2\) (the amount of chirp being quantified by the first parameter \(\alpha\)) followed by a scaled FT (the amount of scale being quantified by the second parameter \(\beta\)) followed by chirp multiplication (the amount of chirp being quantified by the third parameter \(\gamma\)).

The chirp signals can cause rapid oscillation of the kernel. They are the primary cause of problems in the numerical simulation of the LCT and several of its special cases, and can increase dramatically the number of samples required to fully represent the information\(^8\text{-}^{12}\) beyond that used to represent the original signal. This is true in the general cases of the FRT and FST. In the special case of the FT this chirping is not present (\(\alpha = \beta = 0\)), and the same number of samples are used to fully represent the signal before and after the FT. Numerical implementation of the FT is carried out by using the fast Fourier transform (FFT) with a number of calculations of the order \(N \log N\). Several more questions arise when attempting to digitally compute the LCT: Will the discrete transform (i) be unitary and therefore invertible (this property is particularly important when simulating encryption schemes\(^13\) that utilize the LCT and its special cases), (ii) be additive (i.e., will application of two successive discrete LCT algorithm be equivalent to an easily defined single discrete LCT algorithm in the same way as we expect additively for the continuous LCT), and finally, (iii) will it well approximate the continuous LCT.

Numerous algorithms have been defined in the literature for efficient, fast numerical implementation of the FT,\(^{14}\) FRT,\(^{15\text{-}18}\) FST,\(^{19\text{-}24}\) and LCT.\(^9,^{10}\) The first algorithm\(^25\) used to digitally calculate the FRT decomposed the signal to be transformed into a summation of the eigenfunctions of the FRT, i.e., the Hermite–Gaussian functions, and then weighted them with appropriate eigenvalues. This method, however, requires \(N^2\) calculations. Various methods have emerged\(^ {15\text{-}18}\) that use the FFT, enabling a more rapid calculation of the FRT. However, each of these algorithms is accurate only for certain limited ranges of fractional order, and none is guaranteed to be either additive or unitary. Recently it has been shown\(^ {10\text{-}12}\) that each algorithm provides identical results if appropriate interpolation and decimation are applied at well-defined critical stages in each algorithm, the only significant difference being the amount of interpolation and decimation necessary. It has also been shown\(^ {10,11}\) that the algorithms can retain the continuous integral transform’s unitary and additive properties if certain conditions are applied. We note that all of these algorithms involve the application of the FFT.

An exactly unitary, index-additive, discrete FRT matrix
matrix has been derived\textsuperscript{26} based on the discrete counterparts of the Hermite–Gaussian functions. However, no closed-form definition has been given, the transform requires \(N^2\) calculations, and only specific sampling intervals can be used in the input and output domains. We believe that the derivation used in this paper may also be applicable to this discrete-matrix transform, since our algorithm is derived by using only the periodic and shift properties of the continuous FRT and LCT.

Similarly for implementation of the FST, various FFT-based algorithms have been presented.\textsuperscript{19–24} Each of these algorithms is accurate only for certain limited ranges of propagation distance. Once again, none of these algorithms is guaranteed to be either additive or unitary. It has also recently been shown\textsuperscript{10,11} that all of these algorithms generate identical results by using appropriate interpolation and decimation, and that they all can be made both unitary and additive.

In this paper we first develop theory for the discrete linear canonical transform (DLCT)—which is to the LCT what the discrete Fourier transform (DFT) is to the FT—based on direct discretization of the LCT kernel for any input sampling interval. We have defined the output sampling interval to be such that the transform is unitary. We then derive the fast linear canonical transform (FLCT), an \(N \log N\) algorithm for its implementation, using an approach analogous to that used to derive the FFT for implementation of the DFT. This is achieved by using the periodicity and shifting properties of the DLCT to exploit symmetries in the DLCT matrix, breaking down the original matrix transform into identical transforms of smaller sizes in the same way that the FFT breaks down the DFT. To our knowledge the approach and result have never been presented before. The resulting algorithm is independent of the FFT and is based purely on the properties of the LCT. In this way we derive a single \(N \log N\) algorithm that can be used to calculate the FT, the FRT, the FST, and, in the most general case, the effect of any QPS.

In overview the paper is organized as follows. In Section 2 we briefly discuss sampling and present two equivalent expressions representing a sampled function. In Section 3 we discuss the continuous LCT and its shifting properties to derive the discrete space LCT (DSLCT), which is the continuous LCT of a sampled input function. In Section 4 we sample the continuous DSLCT to derive the DLCT. In Section 5 we use the periodicity and shifting properties of the DLCT to derive the FLCT algorithm. In Section 6 we briefly discuss how to impose the additive and unitary properties on our algorithm. In Section 7 we present results produced using a single FLCT algorithm to simulate the FT, FRT, FST, and a two-lens QPS. In Section 8 we offer a conclusion.

\section*{2. SAMPLING}

We define \(\delta(x)\) as Dirac’s impulse having a value of 1 at \(x=0\) and a value of 0 at all other values of \(x\). To obtain a sampled version of some continuous signal, we multiply it by a train of such impulse responses:

\[\delta_T(x) = \sum_{n=-\infty}^{\infty} \delta(x - nT),\]  

(1)

where \(n\) is defined as the sampling index, \(T\) is defined as the sampling interval in \(x\), and its inverse is the sampling frequency \(f_s\). Here \(x\) denotes the space domain, although we could allow it to denote the time domain. The sampled version of a continuous signal \(f(x)\) is given by

\[f^T(x) = f(x)\delta_T(x) = \sum_{n=-\infty}^{\infty} f(nT) \delta(x - nT),\]  

(2a)

which can also be written as a Fourier series,

\[f^T(x) = \frac{1}{T} \sum_{k=-\infty}^{\infty} \exp(j2\pi kf_Tx).\]  

(2b)

\section*{3. SHIFTING THEOREM OF THE LINEAR CANONICAL TRANSFORM AND DERIVATION OF THE DISCRETE SPACE LINEAR CANONICAL TRANSFORM}

The LCT with parameters \(\alpha, \beta, \text{ and } \gamma\) of a function \(f(x)\) is defined as\textsuperscript{2}

\[L_{\alpha\beta\gamma}(y) = \Theta_{\alpha\beta\gamma}[f(x)](y) = \int_{-\infty}^{\infty} K(x,y)f(x)dx,\]  

(3a)

where the LCT kernel is

\[K(x,y) = A \exp[j\pi(\alpha x^2 - 2\beta xy + \gamma y^2)].\]  

(3b)

\(\Theta_{\alpha\beta\gamma}\) is the LCT operator, and \(A\) is a complex constant that is omitted in the following analysis. The LCT obeys the following shift theorems:

\[\Theta_{\alpha\beta\gamma} \left[ \exp(j2\pi \xi x)f(x) \right](y) = \exp(-j\pi \gamma \xi^2/\beta^2) \exp(j2\pi \xi y/\beta) \times \Theta_{\alpha\beta\gamma}[f(x)](y - \xi/\beta).\]  

(4a)

\[\Theta_{\alpha\beta\gamma} \left[ f(x - \xi) \right](y) = \exp(-j\pi \xi^2(\alpha - \gamma \alpha^2/\beta^2)) \exp[j2\pi \xi y/\gamma \alpha/\beta] \times \Theta_{\alpha\beta\gamma}[f(x)](y - \xi \alpha/\beta).\]  

(4b)

Equation (4a) indicates that if we apply a linear change in phase to a signal in the space domain, the LCT of this signal is identical to the LCT of the original signal, except that it has been shifted in the LCT domain by an amount dependent on the change in phase in the space domain, and there is the addition of a constant phase and a linear phase, both dependent on the phase change in the space domain. Equation (4b) shows the analogous effect of a shift in the space domain.

We now apply the LCT to the sampled function \(f^T(x)\). Since we have two equivalent expressions for \(f^T(x)\) in Eq. (2), we can deduce two expressions for its LCT. First we return to Eq. (2b) and apply the LCT to both sides:
\[ \Theta_{\alpha\beta}\{f^T(x)\}(y) = \frac{1}{T} \sum_{k=-\infty}^{\infty} \{ \exp[-j\pi(k/T)^2\gamma\beta] \times \exp[j2\pi(k/T)^\gamma\beta]\Theta_{\alpha\beta}\{f(x)\}(y-kT\beta) \}. \]

(5)

This provides insight into the periodicity of the LCT of the sampled signal. Provided that the sampling is such that aliasing does not occur, the magnitude of \( \Theta_{\alpha\beta}\{f^T(x)\}(y) \) is identical to the magnitude of \( \Theta_{\alpha\beta}\{f(x)\}(y) \) repeated periodically with period \( 1/T\beta \). The phase of \( \Theta_{\alpha\beta}\{f^T(x)\}(y) \) is equivalent to the phase of \( \Theta_{\alpha\beta}\{f(x)\}(y) \) repeated periodically with period \( 1/T\beta \), but also includes phase factors (constant and linear) that are different for every period.

Using Eq. (2a), we can also write that

\[ \Theta_{\alpha\beta}\{f^T(x)\}(y) = \int_{-\infty}^{\infty} \left[ \sum_{n=-\infty}^{\infty} f(n)\delta(x-nT) \right] \exp(-j2\pi\beta xy) \]

\[ \times \exp[-j\pi\alpha^2x^2 + \gamma\beta]dx \]

\[ = \exp[j\pi\gamma\beta^2] \sum_{n=-\infty}^{\infty} f(nT)\exp[j\pi\alpha(nT)^2] \]

(6)

\[ \times \exp(-j2\pi\beta y nT). \]

We will refer to Eq. (6) as the DSLCT, and we shall denote it using the operator \( \Delta_{\alpha\beta} \). If we set \( \alpha = \gamma = 0 \) and \( \beta = 1 \), Eq. (6) reduces to the discrete time (space) Fourier transform (DTFT), and Eq. (5) indicates a periodicity in both magnitude and phase. We note the importance of this periodicity in the DTFT, since it allows the reduction in complexity of certain computations that allow the development of fast algorithms and rapid convolution summations. We also note that certain types of periodicity that have been referred to as chirp periodicity exist in Eq. (5). In this paper we use the existence of this chirp periodicity to create our FLCT algorithm.

### 4. Shift Theorem of the Discrete Space Linear Canonical Transform and Derivation of the Discrete Linear Canonical Transform

Equation (6) shows how to obtain the DSLCT of a discrete function \( f^T(x) \), which we can also write as \( f(nT) \). We now obtain the DSLCT of this function if it has been shifted by \( l \) samples as

\[ \Delta_{\alpha\beta}\{f[(n-l)T]\}(y) = \exp(j\pi\gamma\beta^2) \sum_{n=-\infty}^{\infty} f(n-lT) \]

\[ \times \exp[j\pi\alpha(nT)^2]\exp(-j2\pi\beta y nT). \]

(7a)

Setting \( i=n-l \) gives

\[ \Delta_{\alpha\beta}\{f[(n-l)T]\}(y) = \exp(j\pi\gamma\beta^2) \sum_{i=-\infty}^{\infty} f(iT) \]

\[ \times \exp[j\pi\alpha(i+1)T^2] \]

\[ \times \exp(-j2\pi\beta y(i+1)T) \]

In Eq. (9) \( y \) is still continuous. However, only \( N \) parts can be independent in the \( y \) domain since it is composed of \( N \) sample values. This follows from the number of degrees of freedom. In the case of the DTFT we have periodicity in \( y \) with a period of \( 1/T \), so we can take our \( N \) samples in the \( y \) domain over any range \( 1/T \) we wish. In the case of the DSLCT, we must take into account that the DSLCT is not periodic in \( y \) because of the addition of different phase factors in each period. For the moment, we continue and choose to take our \( N \) samples in the range that is unaffected by the chirp periodic factors,

\[ \frac{1}{2T\beta} \leq y \leq \frac{1}{2T\beta} - \frac{1}{NT\beta} \]

(10)

In steps of \( T_y = 1/NT\beta \). We know that this range describes \( \Theta_{\alpha\beta}\{f(x)\}(y) \), and outside this range we find the same function repeated but with additional phase factors introduced by the discretization of the LCT kernel as seen in Eq. (5). Then...
\[
\exp\left[j \pi \gamma (mT_y)^2 \right] \sum_{n=-N/2}^{N/2-1} f(nT) \exp\left[j \pi \alpha (nT)^2 \right] \\
\times \exp\left[-j 2 \pi \beta \left(mT_y\right) \right] \\
= \exp\left[ j \pi \gamma \left( \frac{m}{NT \beta} \right)^2 \right] \sum_{n=-N/2}^{N/2-1} f(nT) \exp\left[j \pi \alpha (nT)^2 \right] \\
\times \exp\left( -j \frac{2 \pi n m}{N} \right), \tag{11a}
\]

where \( m \) takes the values over the range

\[-N/2 \leq m \leq N/2 - 1. \tag{11b}\]

We call relations (11) the DLCT.

We now investigate the effect of sampling in the \( y \) domain on the input function in the \( x \) domain. We recall that the DSLCT is simply the LCT of a discrete function in the \( x \) domain. We take the LCT as defined in Eq. (3) and sample it at intervals of \( T_y = 1/NT \beta \). Using the definition in Eq. (2b), we obtain

\[
L_{\alpha \beta y}^x(y) = \frac{1}{T_y} L_{\alpha \beta y}^x(y) \sum_{k=-\infty}^{\infty} \exp\left[j 2 \pi k (1/T_y) y \right]. \tag{12}
\]

We now apply the inverse LCT to this to return to the \( x \) domain:

\[
\Theta_{-\gamma \beta \alpha} \left(L_{\alpha \beta y}^x(y)\right)(x) \\
= \Theta_{-\gamma \beta \alpha} \left\{ \sum_{k=-\infty}^{\infty} \exp\left[j 2 \pi k (1/T_y) y \right] \right\}(x). \tag{13}
\]

Using the shifting property of the inverse LCT, which is identical to that of the LCT, we find that Eq. (13) is equal to

Property 1

\[
\exp\left[ j \pi \gamma \left( \frac{m}{NT \beta} \right)^2 \right] \sum_{n=-\infty}^{\infty} \left[f(nT) \exp\left(j 2 \pi \xi n T \right) \right] \exp\left[j \pi \alpha \left(m \pi T \right)^2 \right] \exp\left(-j \frac{2 \pi n m}{N} \right) \\
= \exp\left[ -j \pi \gamma \left( \frac{\xi}{\beta} \right)^2 \right] \exp\left[ j 2 \pi \gamma m \xi \right] \exp\left[j \pi \gamma \left(m - \xi NT \right)^2 \right] \sum_{n=-\infty}^{\infty} f(nT) \exp\left(j \pi \alpha \left(nT\right)^2 \right] \exp\left( -j \frac{2 \pi n \left(m - \xi NT \right)}{N} \right). \tag{15}
\]

Property 2

\[
\exp\left[ j 2 \pi \gamma m \right] \exp\left[ j \pi \gamma \left( \frac{m}{NT \beta} \right)^2 \right] \sum_{n=-\infty}^{\infty} f(nT) \exp\left( -j \frac{2 \pi n m}{N} \right) \\
= \exp\left[ j \pi \gamma \left( \frac{m}{NT \beta} \right)^2 \right] \sum_{n=-\infty}^{\infty} f(nT) \exp\left( j 2 \pi \gamma n T \right) \exp\left(-j \frac{2 \pi n \xi}{\beta^2} \right) \exp\left( j \pi \alpha \left(n - \frac{\xi}{T \beta} \right)^2 \right) \exp\left(-j \frac{2 \pi n \left(n - \xi \beta \right)}{N} \right). \tag{16}
\]

These two properties will be used in conjunction with the chirp periodicity of the DLCT to derive the FLCT algorithm.

Various FFT algorithms for determination of the DFT can be derived by use of time (space) decomposition and frequency (spatial frequency) decomposition. In the fol-
We begin with the equation describing the DLCT,

\[ DL_{n;b}(mT_y) = D\Theta_{n;b}^N[f(nT)][mT_y] \]

\[
= \sum_{n=-N/2}^{N/2-1} f(nT) \exp[j\pi a(nT)^2] \exp\left(-\frac{j2\pi mn}{N}\right) \times \exp\left[j\pi\gamma \left(\frac{m}{N} \right)^2\right]. \tag{17}
\]

where \( D\Theta_{n;b}^N \) is the operator notation for the DLCT and \( m \) has the range defined in relation (11b). Impose the following relationship between \( N \) and \( T \):

\[ T = \frac{h}{\sqrt{N}}, \tag{18} \]

where \( h \) is chosen to satisfy Eq. (18) for both number of samples and the range over which they have been taken. We rewrite the DLCT in Eq. (17) as

\[
D\Theta_{n;b}^N[f(nT)][mT_y] = \sum_{n=-N/2}^{N/2-1} f(nT) \exp\left[\frac{j\pi a(nh)^2}{N}\right] \times \exp\left(-\frac{j2\pi mn}{N}\right) \times \exp\left[j\pi\gamma \left(\frac{m}{h}\right)^2\right]. \tag{19}
\]

Once again \( m \) has the range defined in relation (11b). For ease, we rewrite the kernel in Eq. (19) as

\[
W_{N,h}^{m,m} = \exp\left[\frac{j\pi a(nh)^2}{N}\right] \times \exp\left(-\frac{j2\pi mn}{N}\right) \times \exp\left[j\pi\gamma \left(\frac{m}{h}\right)^2\right]. \tag{20}
\]

We call this the “LCT twiddle factor,” and in the case of the FT parameters \( \alpha = \gamma = 0, \beta = 1 \) it reduces to the twiddle factor found in the derivation of the FT in Refs. 27–29. Using Eqs. (18) and (20) we rewrite Properties 1 and 2 as

**Property 1**

\[
\sum_{a=-\infty}^{+\infty} f(nT) \exp\left(\frac{j2\pi\xi n}{\sqrt{N}}\right) W_{N,h}^{m,m} = \exp\left[-j\pi\gamma \left(\frac{\xi}{\beta}\right)^2\right] \times \exp\left(\frac{j2\pi\gamma n}{h\sqrt{N}\beta^2}\right) \times \sum_{n=-\infty}^{+\infty} f(nT) W_{N,h}^{m-\xi h,\xi h}. \tag{21}
\]

**Property 2**

\[
\sum_{a=-\infty}^{+\infty} f(nT) W_{N,h}^{m,m} = \exp\left[j\pi\alpha \left(\frac{\xi}{\beta}\right)^2\right] \times \sum_{n=-\infty}^{+\infty} f(nT) \exp\left(\frac{j2\pi\alpha n}{h\sqrt{N}}\right) W_{N,h}^{m-\xi h,\xi h}. \tag{22}
\]

We wish to calculate the DLCT of the function

\[ f(nT), \quad -N/2 < n < N/2 - 1, \tag{23} \]

which is split (following from the procedure for FFT frequency decomposition) into

\[ a(n) = f(nT), \]

\[ b(n) = f[(n + N/2)T], \quad N/2 < n < -1. \tag{24} \]

We rewrite Eq. (19) using the functions defined in Eq. (24) as

\[
D\Theta_{n;b}^N[f(nT)][mT_y] = \sum_{n=-N/2}^{N/2-1} f(nT) W_{N,h}^{m,m} = \sum_{n=-N/2}^{-1} [a(n)W_{N,h}^{m,m} + b(n)W_{N,h}^{m,3N/4}], \tag{25}
\]

where \( m \) has the range defined in relation (11b). We can rewrite this using Property 2 as

\[
D\Theta_{n;b}^N[f(nT)][mT_y] = \sum_{n=-N/2}^{-1} [a(n)W_{N,h}^{m,m} + b(n)\exp(j\pi ah^2(n + N/4))] W_{N,h}^{m,m}. \tag{26}
\]

We introduce a new function

\[ c(n) = b(n)\mu_0(n), \quad \mu_0(n) = \exp[j\pi ah^2(n + N/4)], \tag{27} \]

and we rewrite the right-hand side of Eq. (26) as

\[
\sum_{n=-N/2}^{-1} [a(n) + c(n)\exp(-j\pi m)] W_{N,h}^{m,m}. \tag{28}
\]

Taking even and odd \( m \) values in expression (28), we get, respectively,

\[
D\Theta_{n;b}^N[f(nT)][2mT_y] = \sum_{n=-N/2}^{-1} [a(n) + c(n)] W_{N,h}^{2m,m}, \tag{29a}
\]

\[
D\Theta_{n;b}^N[f(nT)][(2m + 1)T_y] = \sum_{n=-N/2}^{-1} [a(n) - c(n)] W_{N,h}^{(2m+1),m}. \tag{29b}
\]

where in Eq. (29), \( m \) is now over the range \(-N/4 < m \leq N/4 - 1\). This concludes the frequency decomposition part of the algorithm derivation. We now apply time decomposition. Defining four new functions by splitting up the input functions in Eqs. (29) into their even and odd samples, we define

\[ p(n) = a(2n) + c(2n), \]

\[ q(n) = a(2n + 1) + c(2n + 1), \]

\[ r(n) = a(2n) - c(2n), \]

\[ s(n) = a(2n + 1) - c(2n + 1), \tag{30} \]

where in each of the cases in Eqs. (30) \( n \) takes values in the range \(-N/4 < n \leq -1\). We rewrite Eqs. (29) as
We then set $D = \text{integer}$ and use of appropriate use of Eq. (5), which defines the chirp periodicity. This shift and then determine the required output by making equally manageable, since we could ignore the integer could have arranged for an integer shift, which would be

We note that the LCT twiddle factor has the property

Equation (35) is made up of two quarter-size DLCTs with output ranges over $-N/4 \leq m \leq N/4 - 1$. This is twice as large as is necessary since we can make use of the chirp periodicity, so that we need calculate only half of these samples. Using the discussion in Section 4, we can calculate the samples of the two quarter-size DLCTs $A$ and $B$ in the range $-N/8 \leq m \leq N/8 - 1$ directly from Eqs. (36). Over the range $-N/4 \leq m \leq -N/4 - 1$,

and in the range $N/8 \leq m \leq N/4 - 1$,

and applying Properties 1 and 2, we obtain

Equation (35) is made up of two quarter-size DLCTs with output ranges over $-N/4 \leq m \leq N/4 - 1$. This is twice as large as is necessary since we can make use of the chirp periodicity, so that we need calculate only half of these samples. Using the discussion in Section 4, we can calculate the samples of the two quarter-size DLCTs $A$ and $B$ in the range $-N/8 \leq m \leq N/8 - 1$ directly from Eqs. (36). Over the range $-N/4 \leq m \leq -N/4 - 1$,

Large as is necessary since we can make use of the chirp periodicity, so that we need calculate only half of these samples. Using the discussion in Section 4, we can calculate the samples of the two quarter-size DLCTs $A$ and $B$ in the range $-N/8 \leq m \leq N/8 - 1$ directly from Eqs. (36). Over the range $-N/4 \leq m \leq -N/4 - 1$,

Large as is necessary since we can make use of the chirp periodicity, so that we need calculate only half of these samples. Using the discussion in Section 4, we can calculate the samples of the two quarter-size DLCTs $A$ and $B$ in the range $-N/8 \leq m \leq N/8 - 1$ directly from Eqs. (36). Over the range $-N/4 \leq m \leq -N/4 - 1$,
Once again we could choose a value of $q$ to give a more manageable integer shift in the output domain, but we choose to eliminate the half shift, so we set $q = 1/h\sqrt{N}$; then
\[
D\theta_{ap}^{T,N}[f(nT)][(2m + 1)T_y] = \exp\left(\frac{j\pi y}{\beta h^2 N}\right) \exp\left(\frac{j4\pi m y}{h^2 \beta^2 N}\right) \sum_{n=-N/4}^{-1} r(n) \right. 
\times \exp\left(- \frac{j4\pi m}{N} W_{(N/4),h}^m + \exp\left(\frac{j4\pi m y}{h^2 \beta^2 N}\right) \right. 
\times \left. \sum_{n=-N/4}^{-1} s(n) \exp\left(- \frac{j4\pi m}{N} W_{(N/4),h}^{n+1/2, m}\right) \right. 
\left. + \mu_2(m) \sum_{n=-N/4}^{-1} s(n) \mu_2(n) W_{(N/4),h}^{n+1/2, m}\right],
\]
where
\[
\mu_2(m) = \exp\left(\frac{j\pi y}{\beta h^2 N}\right) \exp\left(\frac{j4\pi m y}{h^2 \beta^2 N}\right)
\]
and
\[
\mu_0(n) = \exp\left(- \frac{j4\pi m}{N}\right).
\]

We are now in a position identical to that in Eq. (35). To rid ourselves of the shift in the input domain of the second of these DLCTs, we carry out the same procedure, giving
\[
D\theta_{ap}^{T,N}[f(nT)][(2m + 1)T_y] = \mu_2(m) C(m) + \mu_2(m) \mu_2(m) D(m),
\]
where $m$ has the range $-N/4 \leq m \leq N/4 - 1$, the constant term $\mu_2$ is a result of using both Properties 1 and 2 together, and
\[
C(m) = \sum_{n=-N/4}^{-1} r(n) \mu_2(n) W_{(N/4),h}^m, \\
D(m) = \sum_{n=-N/4}^{-1} s(n) \mu_2(n) \mu_2(n) W_{(N/4),h}^m, \tag{44}
\]

Once again for $C(m)$ and $D(m)$ we need only calculate half the range of $m$ and find the second half by using the chirp periodicity of the DLCT. We calculate the samples of the two quarter-size DLCTs $C$ and $D$ in the range $-N/8 \leq m \leq N/8 - 1$ directly from Eqs. (44).

Over the range $-N/4 \leq m \leq -N/8 - 1$,
\[
C(m) = C(m + N/4) \mu_2(m), \quad D(m) = D(m + N/4) \mu_2(m), \tag{45}
\]
and in the range $N/8 \leq m \leq N/4 - 1$,
\[
D(m) = D(m - N/4) \mu_2(m), \quad D(m) = D(m - N/4) \mu_2(m). \tag{46}
\]

This completes our derivation of the FLCT algorithm as a means of calculating the DLCT defined in Eqs. (11). We have explicitly presented a method to decompose an N-point DLCT into four $N/4$-point DLCTs. Computation directly by Eqs. (11) requires $N^2$ calculations. We note that this algorithm is radix 4 and that the size of the input data must be $N=2^n$. If $n$ is odd, we use the fast algorithm recursively until the DLCT has been broken down into $2 \times 2$ blocks that are then calculated directly by Eq. (11a) in the same way as the FFT finishes with the FFT butterfly. If $n$ is even, we use the fast algorithm recursively until the DLCT has been broken down into $4 \times 4$ blocks that are then calculated directly by Eq. (11a).

We also note that a similar procedure can be applied for other radix prime numbers squared, e.g., radix 9, where with each step, the algorithm will break down a large $N \times N$ matrix into nine smaller matrices of dimension $(N/9 \times N/9)$. As mentioned earlier in this section, our method permits the derivation of a large number of fast algorithms, each of which uses of the order of $N \log N$ calculations with the same efficiency as the FFT, and we have outlined the procedure to derive only one of these.

Substituting $\alpha = 1/\tan \theta$, $\beta = 1/\sin \theta$, and $\gamma = 1/\tan \theta$, which gives us the normalized FRT, we now have a fast FRT algorithm and, carrying out a similar substitution for the Fresnel transform, we also have a new fast algorithm for its determination. In Section 8 we present results obtained for these cases and others using our Fast LCT Software Package (© Hennelly-Sheridan 2004) programmed in C++.

6. USING THE ALGORITHM: OUTPUT EXTENSION AND INTERPOLATION AND DECIMATION

The algorithm as it is described up to this point is unitary, and an inverse transform can be applied to recover exactly our initial discrete signal. However, we note the following concerning the output extent or function width. The input in $x$ was taken to be $W_n = NT$ and from Eq. (5), we deduced that the output samples would uniformly occupy a $1/T\beta$ extent in $y$. From the theory of the LCT, we know that the output extent will not in general be equal to this value. We can expect both a change in the output spatial extent $W_n$ and in the output spatial-frequency extent (bandwidth) $B_n$, and therefore we can also expect a change in their product [the signal space–bandwidth product, $SBP$ (SBP)]; the SBP determines the number of uniform samples required fully to represent the signal. We can account for these changes by tracking variations in the shape of the Wigner distribution function (WDF) as we apply a LCT.

Since in general $1/T\beta$ will not be equal to $W_n$, the initial signal must be interpolated or decimated so as to impose this equality. For example, in the case of $W_n$ being twice the value of $1/T\beta$ we must interpolate by a factor of 2 such that $T \longrightarrow T/2$. Efficient methods of interpolation and decimation can be found in the literature. We also note that there may not be enough samples to adequately
represent the continuous LCT. The SBP \((W_p b_s)\) may still be larger than \(N\), even after interpolation has been applied to increase or decrease the output extent. In this case we apply zero padding of the input function until we have equality between the two.

In the case of the continuous LCT, additivity implies that

\[
\Theta_{a_3b_3\gamma_3}(f(x))(y) = \Theta_{a_3b_3\gamma_3}(\Theta_{a_1b_1\gamma_1}(f(x))(y))(y),
\]

\[
\left[ \begin{array}{c}
\gamma_2/\beta_2 \\
-\beta_2 + a_3 \gamma_2/\beta_3 \\
\alpha_3/\beta_3
\end{array} \right] = \left[ \begin{array}{c}
\gamma_2/\beta_2 \\
-\beta_2 + a_2 \gamma_2/\beta_3 \\
\alpha_2/\beta_3
\end{array} \right] \times \left[ \begin{array}{c}
\gamma_1/\beta_1 \\
-\beta_1 + a_1 \gamma_1/\beta_3 \\
\alpha_1/\beta_3
\end{array} \right].
\]

(47b)

The relationship is particularly important when simulating the FRT and FST, and the question arises whether our algorithm obeys Eq. (46). If the conditions concerning interpolation and zero padding are met, and if we take account of the WDF shape at the output of our \((a_1, \beta_1, \gamma_1)\) FLCT and use it as the input WDF shape to the \((a_2, \beta_2, \gamma_2)\) FLCT, then the LCT additive property defined in Eq. (46) will be met by the FLCT algorithm.9,10

Regarding sampling, we note that the LCT induces deformations of the finite area subtending the WDF, referred to as the support (where the original limits are given by use of the power criterion), and will thus lead to a distortion of the initial input rectangular sampled area. This distortion means that although the signal may be accurately represented by the exact same samples (following transformation) the samples will no longer, in general, occur inside a subtending rectangle in phase space. However, practical sampling typically involves use of a CCD camera with a regular periodic pattern of pixels, which is usually interpreted as corresponding to a regular (rectangular in the one-dimensional signal case) area in phase space. A skewed rectangle (following for example a Fresnel transformation) in our procedure becomes a subarea inside a larger rectangle (unless parts of the original subtending area are intentionally neglected). Thus the total number of samples necessary to ensure capture of all the information input to the system will usually involve a change in the sampling. Following the standard procedure we assume sufficient regular sampling to ensure that aliasing effects can be assumed negligible. While some “oversampling” may occur compared to a situation in which sufficient knowledge is available to allow a priori preprocessing of the data,12 in our method no preprocessing or postprocessing of the data is necessary.

7. NUMERICAL RESULTS

The algorithm derived in Section 5 was written in C++ code (Fast LCT © Hennelly-Sheridan 2004). In this section we apply the FLCT to various special cases of the LCT as well as a general, arbitrary QPS. We begin by implementing the DFT of an input rectangular function \(a\) that is equal to 1 in the range \(-1\) to \(+1\) and is equal to 0 everywhere else. We take 512 samples in the range \(-5\) to \(+5\), and we set \(a = \gamma = 0\), \(\beta = 1\) in the FLCT algorithm. The resulting magnitude and phase of the DFT are shown in Fig. 1. In this case the FLCT algorithm reduces to a mixture of the well-known time and frequency decompositions, and the algorithm is numerically as efficient as the standard FFT algorithms.9

Second, we apply the FLCT algorithm to calculate the discrete (normalized) FRT of angle \(\theta = \pi/4\) of the same function. Again we take 512 samples over the range \(-5\) to \(+5\) and we set \(a = \gamma = 1/\tan(\pi/4), \beta = 1/\sin(\pi/4)\) in the FLCT. The magnitude and phase of the resulting discrete complex function are shown in Fig. 2. These results have been verified as being identical to those found using algorithms in the literature for the FRT.10–14

Third, the FLCT algorithm is applied to calculate the discrete FST for \(\lambda = 500\) nm and \(z = 10\) mm. The input function used is a rectangular function equal to 1 in the range \(-0.5\) mm to \(+0.5\) mm and to 0 everywhere else. 2048 samples over the range \(-5\) mm to \(+5\) mm are necessary because of the need for interpolating when implementing the FST. We set \(a = \gamma = \beta = 1/\lambda z\) in the FLCT. The magnitude and phase of the resulting discrete complex function are shown in Fig. 3. These results were verified.
as being identical to those found using standard algorithms for implementing the FST.16–21

A general LCT for an arbitrary bulk optical (QPS) system is calculated. The two-lens system shown in Fig. 4 is simulated. We arbitrarily choose \( f_1 = f_2 = 100 \text{ mm}, \ d_1 = 10 \text{ mm}, \ d_2 = 20 \text{ mm}, \ d_3 = 30 \text{ mm}, \) for which we calculate \( \alpha = 0.0756, \ \gamma = 0.00557, \ \beta = 0.021459 \) by matrix LCT theory.2,8 The input function is taken to be the rectangular function used in the previous case, and 2048 samples are taken in the range \(-5 \text{ mm} \) to \(+5 \text{ mm}.\) The magnitude and phase of the resulting discrete complex function are shown in Fig. 5. While the result is a good approximation to the continuous QPS output, we note that increasingly accurate results can be found if the input function is interpolated and zero-padded in accordance with the conditions outlined in Section 7.8

We conclude this section by noting that the availability of this fast LCT algorithm will find applications in the areas of both two-dimensional and three-dimensional holographic data encryption33–35 and compression.36

8. CONCLUSION

We have discussed the existing algorithms for the numerical implementation (for purposes of numerical simulation in the area of optical signal processing) of the DFT, FRT, and FST. All of the reported algorithms in the literature apply the FFT method. We have presented an expression for the discrete space linear canonical transform (DSLCT) and the discrete linear canonical transform (DLCT); the latter is to the LCT what the Discrete Fourier transform (DFT) is to the FT and is based on a direct discretization of the LCT kernel for any input sampling interval. We then presented the fast linear canonical transform (FLCT), an \( N \log N \) algorithm for implementation of the DLCT that uses an approach similar to that used in deriving the FFT from the DFT. This is achieved by using the periodicity and shifting properties of the DLCT to exploit symmetries in the DLCT matrix to break down the original matrix transform into identical transforms of smaller sizes in the same way as the FFT breaks down the DFT.

The original \( N^2 \) DLCT matrix transform has been decomposed into four quarter-size discrete transforms of identical type in the same way in which the FFT decomposes the \( N^2 \) DFT matrix into two half-size DFT transforms. A new group of fast algorithms that includes fast fractional Fourier and Fresnel transforms was presented. These new algorithms are entirely independent of the FFT, are based purely on the properties of the LCT, and can be used to directly analyze any QPS. We have discussed how the algorithm can approximate the continuous LCT most accurately in terms of the LCT additive property and in terms of output sample space and extent.

ACKNOWLEDGMENTS

We acknowledge the support of Enterprise Ireland and Science Foundation Ireland through the Research Innovation Fund and the Basic Research Programme and of the Irish Research Council for Science, Engineering and Technology. We thank A. Fagan without whose undergraduate lectures this work would not have been possible.
Corresponding author John Sheridan may be reached by telephone at +353-(0)1-716-1927, by fax at +353-(0)1-283-0921, by e-mail at John.Sheridan@ucd.ie, and at http://www.ucd.ie/eleceng.

REFERENCES