

Efficient computation of quadratic-phase integrals in optics

Haldun M. Ozaktas, Aykut Koç,* and Ilkay Sari†

Department of Electrical Engineering, Bilkent University, TR-06800, Bilkent, Ankara, Turkey

M. Alper Kutay

TÜBİTAK-Scientific and Technological Research Council of Turkey, TR-06100, Kavaklıdere, Ankara, Turkey

Received June 29, 2005; revised manuscript received August 22, 2005; accepted September 12, 2005

We present a fast $N \log N$ time algorithm for computing quadratic-phase integrals. This three-parameter class of integrals models propagation in free space in the Fresnel approximation, passage through thin lenses, and propagation in quadratic graded-index media as well as any combination of any number of these and is therefore of importance in optics. By carefully managing the sampling rate, one need not choose N much larger than the space–bandwidth product of the signals, despite the highly oscillatory integral kernel. The only deviation from exactness arises from the approximation of a continuous Fourier transform with the discrete Fourier transform. Thus the algorithm computes quadratic-phase integrals with a performance similar to that of the fast-Fourier-transform algorithm in computing the Fourier transform, in terms of both speed and accuracy. © 2006 Optical Society of America
OCIS codes: 070.2580, 350.6980, 070.2590.

A quadratic-phase system is a unitary system whose output $g(u)$ is related to its input $f(u)$ through a quadratic-phase integral:

$$g(u) = \sqrt{\beta} \exp(-i\pi/4) \times \int_{-\infty}^{\infty} \exp[i\pi(\alpha u^2 - 2\beta u u' + \gamma u'^2)] f(u') du', \quad (1)$$

where α , β , and γ are real parameters. This relationship is also known by other names, including “linear canonical transforms.”^{1–4}

This input–output relation can model a broad class of optical systems, including thin lenses, sections of free space in the Fresnel approximation, sections of quadratic graded-index media, and arbitrary concatenations of any number of these.^{2,3}

Computation of the Fresnel diffraction integral, which is a special case of Eq. (1) with $\alpha=\beta=\gamma$, has received the greatest attention because it describes the propagation of light in free space (see Refs. 5 and 6 and the references therein). As the input–output relationship represented by the Fresnel integral is space invariant and takes the form of a convolution, one can compute it by taking the Fourier transform (FT) of the input, multiplying by the transfer function of free space, and inverse transforming, with overall time $\sim N \log N$. Here we present an algorithm with which to compute Eq. (1) in $\sim N \log N$ time, despite the fact that the relationship represented by the more general Eq. (1) is not space invariant and is not a convolution. It is important to underline that here N is chosen close to the space–bandwidth product of the set of input signals, which is usually the smallest possible value of N that can be chosen in terms of information-theoretic considerations. Therefore the algorithm is highly efficient and, although we do not present a formal proof of this, is quite possibly nearly

the most efficient that is possible. Indeed, the distinguishing feature of the present approach is the care with which sampling and space–bandwidth product issues are handled. Straightforward use of conventional numerical methods can result in inefficiencies either because their complexity is larger than $\sim N \log N$ or because the highly oscillatory quadratic-phase kernel in Eq. (1) forces N to be chosen much larger than the space–bandwidth product of the signals.

In the method presented here, the samples of the output are computed in terms of the samples of the input, in the same sense that the fast-Fourier-transform (FFT) implementation of the discrete Fourier transform (DFT) computes the samples of the continuous FT of a function. Special care is taken to ensure that the output samples represent the continuous transform in the Nyquist–Shannon sense, so the continuous transform can be fully recovered from the samples. The only source of deviation from exactness arises from the fundamental fact that a signal and its transform cannot both be of finite extent. This is the same source of deviation encountered when the DFT–FFT is used to compute the continuous FT. Indeed, this deviation enters the present algorithm only in the step involving a DFT; all the other steps are fully exact. Therefore the algorithm provides a degree of approximation similar to that of the DFT–FFT and is, in all aspects, to quadratic-phase transforms what the DFT–FFT is to the ordinary FT.

Before presenting the algorithm we discuss the Wigner distribution (WD) and the fractional Fourier transform (FRT). The WD, $W_f(u, \mu)$, is a phase-space distribution that gives the distribution of signal energy as a function of space and frequency; its definition in terms of $f(u)$ may be found elsewhere.^{2,7} The WD of $g(u)$ can be simply related to the WD of $f(u)$ by a geometrical transformation: $W_g(u, \mu) = W_f(Du - B\mu, -Cu + A\mu)$, where $A = \gamma/\beta$, $B = 1/\beta$, $C = -\beta + \alpha\gamma/\beta$, and

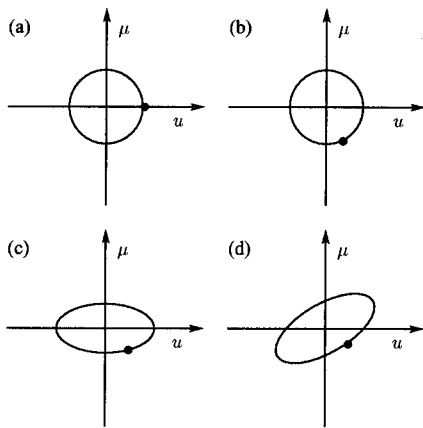


Fig. 1. Decomposition of an arbitrary quadratic-phase system: (a) region to which the signal energy is initially confined, (b) energy distribution after rotation, (c) after scaling, (d) final distribution of signal energy after shearing.

$D = \alpha/\beta$. The FRT $\mathcal{F}^\alpha f(u) = f_a(u)$ is a special case of quadratic-phase transforms² with $\alpha = \gamma = \cot(a\pi/2)$, $\beta = \csc(a\pi/2)$, and an additional factor $\exp(ia\pi/4)$ in front of the integral in Eq. (1), for $-2 \leq a \leq 2$. \mathcal{F}^1 is the ordinary Fourier operator and \mathcal{F}^0 is the identity operator. The FRT operator is additive in index: $\mathcal{F}^{a_1} \mathcal{F}^{a_2} = \mathcal{F}^{a_1+a_2}$. We obtain $W_f(u, \mu)$ by rotating $W_f(u, \mu)$ by $a\pi/2$.

A function and its FT cannot both be of finite extent. In practice, however, we always work with finite spatial and frequency intervals. We assume that a large percentage of the signal energy, as represented by the WD, is confined to an ellipse with diameters ΔS (in meters) in the space dimension and ΔB (in inverse meters) in the frequency dimension. For a given set of functions we can ensure that this is so by choosing ΔS and ΔB sufficiently large. This implies that the space-domain representation of our signal is approximately confined to the interval $[-\Delta S/2, \Delta S/2]$ and that its frequency-domain representation is confined to $[-\Delta B/2, \Delta B/2]$. We then define the space-bandwidth product $\Delta S \Delta B$, which is always ≥ 1 , because of the uncertainty relation. Let us now introduce scaling parameter s (in meters) and scaled coordinates, such that the space- and frequency-domain representations are confined to intervals of length $\Delta S/s$ and $\Delta B s$. Let $s = \sqrt{\Delta S/\Delta B}$ so the lengths of both intervals are equal to the dimensionless quantity $\sqrt{\Delta B \Delta S}$, which we denote by Δu , and the ellipse becomes a circle with diameter Δu . In the new coordinates, our signal can be represented in both domains with Δu^2 samples spaced Δu^{-1} apart. We assume that this dimensional normalization has been performed and that the coordinates are dimensionless.

Our method is based on the decomposition of Eq. (1) into elementary operations. Numerous such decompositions are possible,² but they are not equally suited for numerical purposes. For instance, the direct application of the naive decomposition of chirp multiplication, Fourier transformation, scaling (magnification), and again chirp multiplication that suggests itself on inspection of Eq. (1) will in general

lead to high sampling rates. Our method relies on decomposition into a FRT, magnification, and chirp multiplication. These operations geometrically correspond to rotation, scaling, and shearing in the space-frequency plane (Fig. 1). This decomposition was inspired by the optical interpretation in Ref. 8 and is also a special case of the Iwasawa decomposition. Mathematically,

$$g(u) = \exp(-ia\pi/4) \exp(-i\pi q u^2) f_{sc}(u), \quad (2)$$

$$f_{sc}(u) = \sqrt{1/M} f_a(u/M), \quad (3)$$

where $f_a(u)$ is the FRT of $f(u)$ and

$$\alpha = (2/\pi) \operatorname{arccot} \gamma, \quad (4)$$

$$M = \begin{cases} \sqrt{1 + \gamma^2}/\beta & \gamma \geq 0 \\ -\sqrt{1 + \gamma^2}/\beta & \gamma < 0 \end{cases}, \quad (5)$$

$$q = \gamma\beta^2/(1 + \gamma^2) - \alpha. \quad (6)$$

The ranges of the square root and the arccotangent both lie in $(-\pi/2, \pi/2]$. The first operation in this decomposition is the FRT, whose fast computation in $\sim N \log N$ time is presented in Ref. 9. In the research reported in Ref. 9, the FRT was efficiently computed by being decomposed into a chirp multiplication followed by a chirp convolution followed by a final chirp multiplication. The algorithm computed the samples of the continuous FRT in terms of the samples of the original signal. Just as in the present research, care was taken to ensure that the output samples represented the continuous FRT in the Nyquist–Shannon sense. Our algorithm calls the algorithm in Ref. 9 as a subroutine. The only approximation in this subroutine comes from the step involving chirp convolution in which a DFT–FFT is used to approximate the samples of the continuous FT. No other approximation is made, either in this subroutine or in any of the other operations that we employ in our method. Thus the only source of approximation can be traced to the evaluation of a continuous FT by use of a DFT (implemented with a FFT), which is a consequence of the fundamental fact that the signal energy cannot be confined to finite intervals in both domains. The sec-

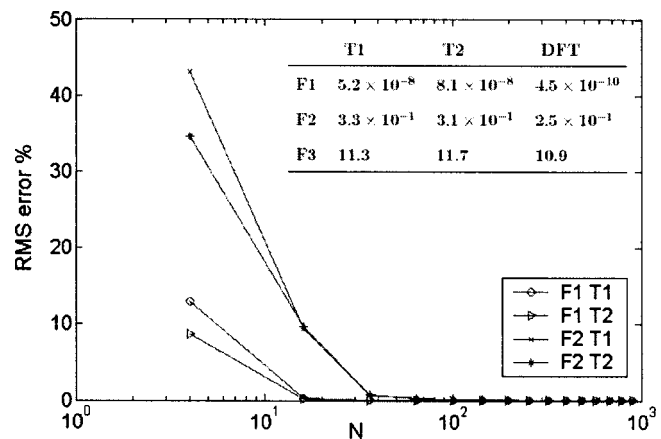


Fig. 2. rms errors versus N . Inset, rms error percentages.

ond operation in our decomposition is scaling, which is actually not a real operation but involves only a re-interpretation of the same samples with a scaled sampling interval. The final operation is chirp multiplication, which takes $\sim N$ time, leading to an overall complexity of $\sim N \log N$. However, there is one final subtlety that must be taken into consideration. Because quadratic-phase systems distort the original region, as shown in Fig. 1, both the space and the frequency extent of the signal as well as its space-bandwidth product may increase, despite the fact that the area of the region remains the same. Therefore, unless some sophisticated basis is used, a greater number of samples than Δu^2 may be needed to represent $g(u)$.

Delaying confrontation of the necessity to deal with this greater number of samples until the very last step is another advantage of our method. Because the FRT corresponds to rotation and scaling only to re-interpretation of the samples, these steps do not require us to increase the number of samples. At the last chirp multiplication step, if we multiply the samples of the intermediate result by the samples of the chirp, the samples obtained will be good approximations of the true samples of $g(u)$ at that sampling interval. If these samples are sufficient for our purposes, nothing further need be done. However, in general these samples will be below the Nyquist rate for $g(u)$ and will not be sufficient for full recovery of the continuous function. To obtain a sufficient number of samples that will permit full recovery, we must interpolate the intermediate result by at least a factor of k , corresponding to the increase in space-bandwidth product:

$$k = 1 + |\gamma - \alpha(1 + \gamma^2)/\beta^2|. \quad (7)$$

Although this formula is sufficient for the approach here, where by careful design the necessity for increasing the number of samples has been postponed to the final step, we note that a systematic approach to determining the increase in the number of samples for arbitrary operations has been presented in Ref. 10, which also brings a number of different perspectives to the same problem.

Consider the chirped pulse function (F1) (Fig. 2) $\exp(-\pi u^2 - i\pi u^2)$ and the trapezoidal function (F2) $1.5 \text{tri}(u/3) - 0.5 \text{tri}(u)$, whose transforms have been computed both by the fast method presented here and by a highly inefficient brute-force numerical approach that is taken here as a reference [$\text{tri}(u) = \text{rect}(u) * \text{rect}(u)$]. Because these functions are well confined to a circle with diameter $\Delta u = 8$ we take $N = 8^2$. We also consider the binary sequence (F3) 01101010 occupying $[-8, 8]$, with each bit two units in length, such that $N = 16^2$. Also consider two transforms, the first (T1) with parameters $(\alpha, \beta, \gamma) = (-3, -2, -1)$ (leading to $k = 1.5$) and the second (T2) with parameters $(-4/5, 1, 2)$ (leading to $k = 7$). The results

are tabulated in the inset of Fig. 2. We note that the errors are in line with the errors in approximating the FT with the DFT. Figure 2 shows the error versus the number of sample points; the error decreases toward the space-bandwidth product but saturates afterward, demonstrating that we can choose the number of points to be comparable to the space-bandwidth product.

In conclusion, we have presented a fast, $\sim N \log N$ algorithm for computing any quadratic-phase integral. Our approach is based on concepts from signal analysis and processing rather than on conventional numerical analysis. With careful consideration of sampling issues, N can be chosen close to the space-bandwidth product of the signals. The only source of error is that which arises from the approximation of a FT with the DFT. Thus the algorithm can compute quadratic-phase integrals with a performance similar to that of the FFT algorithm in computing the FT, in terms of both speed and accuracy. This approach not only handles a much more general family of integrals but also effectively addresses certain difficulties, limitations, or trade-offs that arise in other approaches to computing the Fresnel integral (see Ref. 10 for a systematic comparison of several approaches). The presented algorithm can also be used for efficient realization of filtering in linear canonical transform domains.¹¹

We thank one of the anonymous reviewers for bringing the recent Ref. 10 to our attention. H. M. Ozaktas (haldun@ee.bilkent.edu.tr) acknowledges partial support of the Turkish Academy of Sciences.

*A. Koc is currently with the Department of Electrical Engineering, Stanford University, Stanford, California 94305.

[†]I. Sari is currently with the Department of Electrical Engineering, Texas A&M University, College Station, Texas 77843.

References

1. K. B. Wolf, *Integral Transforms in Science and Engineering* (Plenum, 1979).
2. H. M. Ozaktas, Z. Zalevsky, and M. A. Kutay, *The Fractional Fourier Transform with Applications in Optics and Signal Processing* (Wiley, 2001).
3. M. J. Bastiaans, *J. Opt. Soc. Am.* **69**, 1710 (1979).
4. S. Abe and J. T. Sheridan, *Opt. Lett.* **19**, 1801 (1994).
5. D. Mendlovic, Z. Zalevsky, and N. Konforti, *J. Mod. Opt.* **44**, 407 (1997).
6. D. Mas, J. Garcia, C. Ferreira, L. M. Bernardo, and F. Marinho, *Opt. Commun.* **164**, 233 (1999).
7. L. Cohen, *Time-Frequency Analysis* (Prentice-Hall, 1995).
8. H. M. Ozaktas and M. F. Erden, *Opt. Commun.* **143**, 75 (1997).
9. H. M. Ozaktas, O. Arıkan, M. A. Kutay, and G. Bozdağı, *IEEE Trans. Signal Process.* **42**, 2141 (1996).
10. B. M. Hennelly and J. T. Sheridan, *J. Opt. Soc. Am. A* **22**, 917 (2005).
11. B. Barshan, M. A. Kutay, and H. M. Ozaktas, *Opt. Commun.* **135**, 32 (1997).