Applied and Computational Harmonic Analysis **9**, 19–53 (2000)
doi:10.1006/acha.2000.0305, available online at http://www.idealibrary.com on IDEAL®

Efficient Computation of Oscillatory Integrals via Adaptive Multiscale Local Fourier Bases [1]

A. Averbuch

School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel

E. Braverman

Computer Science Department, Technion, Haifa 32000, Israel

R. Coifman

Department of Mathematics, Yale University, New Haven, Connecticut 06520-8283

and

M. Israeli and A. Sidi

Computer Science Department, Technion, Haifa 32000, Israel

Communicated by Leslie F. Greengard

Received April 15, 1999

The integral $\int_0^L e^{i\nu\phi(s,t)} f(s)\,ds$ with a highly oscillatory kernel (large ν, ν is up to 2000) is considered. This integral is accurately evaluated with an improved trapezoidal rule and effectively transcribed using local Fourier basis and adaptive multiscale local Fourier basis. The representation of the oscillatory kernel in these bases is sparse. The coefficients after the application of local Fourier transform are smoothed. Sometimes this enables us to obtain further compression with wavelets.

1. INTRODUCTION

Computation of oscillatory integrals occurring in scientific and engineering simulations faces difficult problems. As compared to the nonoscillatory case, accurate evaluation of such integrals requires at least a few grid points per oscillation.

For a one-dimensional integral

$$\int_a^b e^{i\lambda\phi(s)} f(s)\,ds, \tag{1.1}$$

[1] This research is supported by a U.S.–Israel Binational Science Foundation grant for 1996–1999.

1063-5203/00 $35.00

with large λ and smooth (nonoscillating) f, there exist methods for fast evaluation, for example, the method of stationary phase first developed in [21] and justified in [22]. Upon assuming λ large, we find that the rapid oscillation of $\exp\{i\lambda\phi(s)\}$ produces cancellation of the integral everywhere except near the stationary points, where the derivative $\phi'(s)$ vanishes. The method proved to be an effective tool in asymptotic expansions and evaluations. However, it requires analytic work in each case (as it cannot be applied automatically), does not reach high accuracy, and is not easily extended to higher dimensions. There exist fast algorithms for special kinds of kernels without regularity assumptions such as [18].

We consider the problem of the fast and accurate evaluation for the integral

$$(Tf)(t) = \int_a^b K(t,s)\, f(s)\, ds, \tag{1.2}$$

where $K(t,s)$ is an oscillatory kernel and f is a smooth function. The kernel is of the type

$$K(t,s) = e^{i\lambda\phi(t,s)}. \tag{1.3}$$

Such integrals arise in acoustic scattering [6] and other applications. Methods, which are effective for evaluation of (1.1), if employed in the two-dimensional case (1.2), become very expensive. If the oscillatory integral (1.2) is computed for each t, even the choice of a minimal number of grid points results in a large-scale matrix problem. Moreover, the matrices obtained at the discretization step are dense, which leads to heavy computation when matrix–vector multiplications are needed. Computation time can be reduced if a sparse representation of $K(t,s)$ can be achieved in some basis. A sparse representation of the kernels (1.3) with wavelets (which appeared to be an efficient tool for sparsification of matrix–vector multiplications in many cases) could be expected. However, wavelet representation of an oscillating matrix appears to be as dense as the original (probably due to the fact that K is oscillating everywhere); i.e., oscillatory kernels cannot be handled efficiently by representing them in wavelet bases.

In [6] a numerical method for fast computation of the integral (1.2) with the kernel (1.3) was described and it was shown analytically that $K(t,s)$ becomes sparse when represented in appropriate local Fourier basis (LFB) (more precisely, the number of elements above a threshold per row grows as $\log N$, where N is the number of grid points in each direction).

The purpose of the present paper is to address more carefully numerical issues arising from [6] and to evaluate the actual numerical efficiency of these methods. In particular, we consider an adaptive basis selection, optimized bell functions, and accurate quadratures. In [6] the location of large coefficients in the matrix realization was related to the geometry of the curve on which the integral is defined. The matrices obtained represent the oscillatory interactions between regions on that curve. The operation count of the algorithm was evaluated to be $O(N \log N)$.

Here we illustrate this method by numerical examples and develop the algorithm by applying multiresolution local cosine bases which leads to representation of (1.3) with a smaller number of coefficients above a threshold due to their adaptivity (about 3 to 4 times fewer coefficients for accuracies 10^{-4} to 10^{-6} and 1024×1024 grid points).

The following integral operator was chosen in [6] as a model problem:

$$\int_0^L e^{i\nu|z(s)-z(t)|} f(s)\,ds. \tag{1.4}$$

Here $\nu \geq 100$, the number of grid points $N \geq 2\nu L$, and $z = z(s)$, $0 \leq s \leq L$, is the parametric representation of a closed curve. The segment $[0, L]$ is divided into M equal intervals with centers a_i. Let $b_i(s)$ be a bell function supported in $[a_{i-1}, a_{i+2}]$ satisfying

$$\sum_{i=1}^{M} b_i^2(s) = 1, \qquad b_i(s) = b_{i-1}(2a_i - s).$$

The functions

$$C_i^k(s) = b_i(s)\left(\frac{2}{a_{i+1}-a_i}\right)^{1/2} \cos\left(\left(k+\frac{1}{2}\right)\pi\frac{s-a_i}{a_{i+1}-a_i}\right) \tag{1.5}$$

form an orthogonal basis of $\mathbf{L}^2([0, L])$ [10]. (In the discrete version grid points are chosen such that a_i are midpoints between adjacent s.) The matrix of coefficients contains $O(N \log N)$ elements above a threshold [6]. There it was evaluated that the number of elements above a threshold is $O(N)$ provided the curve is smooth with bounded derivatives up to a certain order.

The method presented in [6] provides a numerical algorithm of order $O(N \log N)$ as $N \to \infty$. However, to achieve a sparse representation of the oscillatory kernel a *multiscale adaptive local cosine transform* is used rather than a one-level local cosine transform.

Multiscale adaptive use of the local cosine transform, as explained in Section 4, enables the achievement of an impressive compact coding description of oscillatory data. For example, it can be used either in low bit compression of oscillatory seismic data (image compression as described in [4]) or in sparsification of operator kernels as in the case here. The methodology for handling oscillatory integrals and seismic data is almost identical. Therefore, the notion of a 2-D matrix which describes an operator kernel and the notion of image will be used interchangeably. The development of the algorithm in the present paper is independent of the geometry.

We recall that there are two problems when computing oscillatory integrals:

1. Integral (1.2) has to be accurately computed with a comparatively small number of grid points.
2. The matrix obtained after the discretization of the kernel should be effectively presented (sparse representation of the kernel).

The second problem has been discussed already. It is well known that for convergence of a quadrature formula the number of grid points per oscillation should be greater than two. Thus, our goal is to find the best quadrature which gives an admissible accuracy for 8ν grid points only in each direction. For accurate computation of the integral we apply an *improved trapezoidal rule* which is described in Section 2. This method allows us to achieve accuracy of 10^{-6} to 10^{-7} for the computation of integral (1.4) when $N = 8\nu = 1024$ in each direction, $L = 2\pi$. This result can be improved when using Richardson extrapolation. This method uses the quadratures obtained for coarser grids to

sharpen the result in a finer grid. The problem of computing oscillatory integrals is that computations with two or fewer grid points per oscillation are not relevant. Thus, for each grid point there exists an optimal number of levels employed in Richardson extrapolation. This is discussed in Section 2.4.

When compared to [6] the methods for computing the oscillatory integral (1.4) are extended in the following direction:

1. An improved trapezoidal rule is applied for accurate computations of (1.4) for a comparatively small number of grid points. Richardson extrapolation can be employed for further improvement of convergence of the integral (1.4).
2. The matrix obtained after the discretization is presented using either a usual or a multiscale local cosine transform.

The paper is organized as follows. In Section 2 the problem of accurate integration is solved by choosing an appropriate quadrature formula. In Section 3 the local Fourier (cosine) transform is applied to the discretization developed in Section 2. Numerical examples are presented, with the analysis of sparsity for the matrix of coefficients obtained by the transform. Section 4 describes a multiscale decomposition with local Fourier basis providing an automatic choice of the best basis for a given matrix. The comparison of the sparsities for matrices of coefficients obtained by the ordinary and multiscale transforms are presented. In the conclusion further development of this method is discussed.

2. IMPROVED QUADRATURES FOR OSCILLATING INTEGRALS

2.1. Statement of the Problem

In this section we discuss the numerical approximation of integrals of the form

$$\int_0^L K(t,s)\,f(s)\,ds. \tag{2.1}$$

We assume that $f(s) \in \mathbf{C}^M[0,L]$, where M is as large as needed, and that $K(t,s)$ is continuous on $[0,L]\times[0,L]$ and differentiable as many times as needed in the closed triangular domains

$$S_+ = \{(t,s): 0 \le t \le s \le L\} \qquad \text{and} \qquad S_- \{(t,s): 0 \le s \le t \le L\},$$

but $(\partial^k/\partial s^k)K(t,s)$ is discontinuous across the diagonal $s=t$, where it has a finite jump discontinuity, $k=1,2,\ldots$, and we set

$$\delta_i(t) = \frac{\partial^i}{\partial s^i}K(t,t+) - \frac{\partial^i}{\partial s^i}K(t,t-), \qquad i=1,2,\ldots. \tag{2.2}$$

As an example, consider

$$K(t,s) = e^{i\nu|z(s)-z(t)|}. \tag{2.3}$$

When the closed curve represented by $z(s)$ is an ellipse with foci at $\pm\sqrt{a^2-b^2}$ and semi-major and semi-minor axes a and b, we have with $L=2\pi$

$$\begin{aligned} &z(t)=a\cos(t)+ib\sin(t),\\ &|z(t)-z(s)|=\sqrt{a^2(\cos(t)-\cos(s))^2+b^2(\sin(t)-\sin(s))^2}. \end{aligned} \tag{2.4}$$

$K(t,s)$ is continuous for all $t,s\in[0,2\pi]$, as is obvious from (2.3). In particular, $K(t,t)=1$ for all t. Next,

$$\begin{aligned} |z(t)-z(s)| &= \sqrt{a^2\left(-2\sin\frac{t+s}{2}\sin\frac{t-s}{2}\right)^2+b^2\left(2\cos\frac{t+s}{2}\sin\frac{t-s}{2}\right)^2}\\ &= 2\sin\frac{|t-s|}{2}\sqrt{a^2\left(\sin\frac{t+s}{2}\right)^2+b^2\left(\cos\frac{t+s}{2}\right)^2}. \end{aligned} \tag{2.5}$$

Now $\sin(|t-s|/2)$ is differentiable for all t and s both in S_+ and in S_- but not across the diagonal $s=t$, where its partial derivatives have finite jump discontinuities. The term

$$\Psi(t,s)=\sqrt{a^2\left(\sin\frac{t+s}{2}\right)^2+b^2\left(\cos\frac{t+s}{2}\right)^2} \tag{2.6}$$

is well defined and differentiable an infinite number of times for all t and s including the line $s=t$, since $[\Psi(t,s)]^2$ in (2.6) never vanishes. This implies that $(\partial^i/\partial s^i)K(t,s)$, $i=1,2,\ldots,$ have finite jump discontinuities across $s=t$. The computation of the jump turns out to be particularly simple. From

$$\frac{\partial}{\partial s}|z(t)-z(s)|=\begin{cases} -\cos\dfrac{t-s}{2}\Psi(t,s)+2\sin\dfrac{t-s}{2}\dfrac{\partial}{\partial s}\Psi(t,s), & s<t,\\ \cos\dfrac{s-t}{2}\Psi(t,s)+2\sin\dfrac{s-t}{2}\dfrac{\partial}{\partial s}\Psi(t,s), & t<s, \end{cases} \tag{2.7}$$

and from the fact that

$$\frac{\partial K}{\partial s}(t,s)=i\nu\left(\frac{\partial}{\partial s}|z(t)-z(s)|\right)K(t,s),$$

we have

$$\begin{aligned} \delta_1(t) &= i\nu\left[\left(\frac{\partial}{\partial s}|z(t)-z(s)|\right)\Big|_{s=t+}-\left(\frac{\partial}{\partial s}|z(t)-z(s)|\right)\Big|_{s=t-}\right]\\ &= 2i\nu\sqrt{a^2\cos^2 t+b^2\sin^2 t}. \end{aligned}$$

In the general case, when $z(s)$ is an arbitrary smooth closed curve without self-intersections, $K(t,s)$ in (2.3) is also continuous and $K(t,t)=1$. Then

$$|z(t)-z(s)|=|t-s|\left|\frac{z(t)-z(s)}{t-s}\right|$$

differs from zero and is differentiable an infinite number of times everywhere except along the diagonal. At the diagonal $(z(t)-z(s))/(t-s)$ tends to the derivative, which is a finite nonzero value, so the first derivative of (2.3) in s has a finite jump across the diagonal that originates from the factor $|t-s|$.

2.2. Corrected Trapezoidal Rules

Let us proceed to the problem of the accurate evaluation of the integral with kernel (2.3). The problem of the numerical approximation of integrals of the form described above was recently considered in [20]. The approach of [20], in turn, was inspired by the earlier work [19], and both approaches are based on the trapezoidal rule.

Pick a number of grid points N and let $h = L/N$ and $s_k = kh$, $k = 0, 1, \ldots, N$ (then $s_0 = 0$ and $s_N = 1$). Fix $t = s_i$ for any i, $i \in \{0, 1, \ldots, N\}$. By the fact that the curve represented by $z(t)$ is closed, $f(s)$ is periodic in s with period L and $K(t, s)$ is periodic both in t and in s with period L. Consequently, $K(t, s) f(s)$ is periodic in s with period L, and the trapezoidal rule approximation $T(h)$ for $\int_0^L K(t, s) f(s)\, ds$ can be written as

$$T(h) = h \sum_{j=1}^{N} K(s_i, s_j) f(s_j). \tag{2.8}$$

As shown in [20], $T(h)$ has the asymptotic expansion

$$T(h, s_i) = (Tf)(s_i) + \sum_{\mu=1}^{p-1} \frac{B_{2\mu}}{(2\mu)!} \left[\frac{\partial^{2\mu-1}}{\partial s^{2\mu-1}} G(s_i, 1) - \frac{\partial^{2\mu-1}}{\partial s^{2\mu-1}} G(s_i, 0) \right] h^{2\mu}$$
$$- \sum_{\mu=1}^{p-1} \frac{B_{2\mu}}{(2\mu)!} \left[\frac{\partial^{2\mu-1}}{\partial s^{2\mu-1}} G(s_i, s_i+) - \frac{\partial^{2\mu-1}}{\partial s^{2\mu-1}} G(s_i, s_i-) \right] h^{2\mu} + M_i^{(p)} h^{2p}, \tag{2.9}$$

where $G(t, s) \equiv K(t, s) f(s)$ and $|M_i^{(p)}| \le M^{(p)}$, $i = 0, 1, \ldots, N$, for some $M^{(p)} > 0$. Here B_i are Bernoulli numbers. By the fact that $G(t, s)$ is periodic, the first summation in (2.9) vanishes, and we have

$$T(h, s_i) = (Tf)(s_i) - \sum_{\mu=1}^{p-1} \frac{B_{2\mu}}{(2\mu)!} \left[\frac{\partial^{2\mu-1}}{\partial s^{2\mu-1}} G(s_i, s_i+) - \frac{\partial^{2\mu-1}}{\partial s^{2\mu-1}} G(s_i, s_i-) \right] h^{2\mu} + M_i^{(p)} h^{2p}. \tag{2.10}$$

Note that $T(h, s_i) - (Tf)(s_i) = O(h^2)$ as $h \to 0$.

We can improve the accuracy of $T(h, s_i)$ by adding to it first of all the term with $\mu = 1$ in (2.10). By the fact that

$$\frac{\partial}{\partial s} G(t, s) = \left[\frac{\partial K}{\partial s}(t, s) \right] f(s) + K(t, s) f'(s),$$

and using the value $B_2 = -\frac{1}{6}$, we thus obtain the corrected trapezoidal rule $\hat{T}(h, s_i)$ given by

$$\hat{T}(h, s_i) = T(h, s_i) + \frac{1}{12} \delta_1(s_i) f(s_i), \tag{2.11}$$

with $\delta_i(t)$ as defined in (2.2). Obviously, $\hat{T}(h, s_i) - (Tf)(s_i) = O(h^4)$ as $h \to 0$, and

$$\hat{T}(h, s_i) = (Tf)(s_i) - \sum_{\mu=1}^{p-1} \frac{B_{2\mu}}{(2\mu)!} \left[\frac{\partial^{2\mu-1}}{\partial s^{2\mu-1}} G(s_i, s_i+) - \frac{\partial^{2\mu-1}}{\partial s^{2\mu-1}} G(s_i, s_i-) \right] h^{2\mu} + M_i^{(p)} h^{2p}. \tag{2.12}$$

The accuracy of the method is $O(h^4)$ compared to $O(h^2)$ with the usual trapezoidal rule. Further improvement of the convergence is discussed in Section 2.4.

2.3. *Numerical Results for the Corrected Trapezoidal Rule*

The computational results for the usual and the corrected trapezoidal rules are presented below for two cases: (i) the kernel oscillates very slowly ($\nu = 1$) and (ii) the kernel is oscillatory (ν is up to 256). In the following numerical examples we compute

$$\int_0^{2\pi} \sin\Big(\nu\sqrt{(\cos(t)-\cos(s))^2+(\sin(t)-\sin(s))^2}\Big) f(s)\,ds, \tag{2.13}$$

which is a particular case of (2.3), (2.4) for the circle ($a = b = 1$), with $f(t) = \cos t$.

We will use the following measures to estimate the errors:

$$\varepsilon_{\mathrm{MAX}} = \max_{1\le i\le N} |T(h_i, s) - (Tf)(s_i)|,$$

$$\varepsilon_{\mathrm{MSQ}} = \frac{1}{N}\sqrt{\sum_{i=1}^{N}\big(T(h_i, s) - (Tf)(s_i)\big)^2},$$

$$\varepsilon_{\mathcal{L}^2} = \sqrt{\frac{\sum_{i=1}^{N}(T(h_i, s) - (Tf)(s_i))^2}{\sum_{i=1}^{N}(Tf)(s_i)}}.$$

Table 1 illustrates the order of the regular and the corrected trapezoidal rules for integral (2.13) and $f(s) = \cos s$ in the nonoscillatory case $\nu = 1$.

In Table 1 we observe a good fit between the numerical results and the theoretical ones: $O(h^4)$ accuracy for the corrected trapezoidal rule compared to $O(h^2)$ for the usual trapezoidal rule.

Now we consider the dependence of the accuracy on the number of oscillations of the kernel. Table 2 presents the dependence of the error $\varepsilon_{\mathrm{MAX}}$ on the number of oscillations.

TABLE 1
Comparison between the Accuracies of the Regular and the Accuracies of the Corrected Trapezoidal Rules for Integral (2.13) with $\nu = 1$ and with $f(s) = \cos s$

N	Corrected trapezoidal rule			Ratio	Usual trapezoidal rule			Ratio
	$\varepsilon_{\mathrm{MAX}}$	$\varepsilon_{\mathrm{MSQ}}$	$\varepsilon_{\mathcal{L}^2}$	$\varepsilon_{N/2}/\varepsilon_N$	$\varepsilon_{\mathrm{MAX}}$	$\varepsilon_{\mathrm{MSQ}}$	$\varepsilon_{\mathcal{L}^2}$	$\varepsilon_{N/2}/\varepsilon_N$
32	2.8e−6	2.0e−6	2.0e−5		1.0e−3	7.3e−4	7.1e−3	
64	1.7e−7	1.2e−7	1.2e−6	16.5	2.6e−4	1.8e−4	1.8e−3	3.9
128	1.1e−8	7.7e−9	7.6e−8	15.5	6.4e−5	4.5e−5	4.4e−4	4.1
256	6.8e−10	4.8e−10	4.7e−9	16.2	1.6e−5	1.1e−5	1.1e−4	4.0
512	4.3e−11	3.0e−11	3.0e−10	15.8	4.0e−6	2.8e−6	2.8e−5	3.9
1024	2.7e−12	1.9e−12	1.8e−11	15.9	1.0e−6	7.1e−7	6.9e−6	4.1
2048	1.7e−13	1.2e−13	1.2e−12	15.9	2.5e−7	1.8e−7	1.7e−6	4.1
4096	1.2e−14	6.9e−15	6.8e−14	14.2	6.2e−8	4.4e−8	4.3e−7	4.0

TABLE 2
Dependence of the Error $\varepsilon_{\mathrm{MAX}}$ on the Number of Oscillations for the Corrected Trapezoidal Rule, where $N = 64, 128, 256$

N	$\nu = 2$	$\nu = 4$		$\nu = 8$		$\nu = 16$		$\nu = 32$		$\nu = 64$	
	$\varepsilon_{\mathrm{MAX}}$	$\varepsilon_{\mathrm{MAX}}$	$\varepsilon_4/\varepsilon_2$	$\varepsilon_{\mathrm{MAX}}$	$\varepsilon_8/\varepsilon_4$	$\varepsilon_{\mathrm{MAX}}$	$\varepsilon_{16}/\varepsilon_8$	$\varepsilon_{\mathrm{MAX}}$	$\varepsilon_{32}/\varepsilon_{16}$	$\varepsilon_{\mathrm{MAX}}$	$\varepsilon_{64}/\varepsilon_{32}$
64	6.0e−7	3.2e−6	5.3	2.2e−5	6.9	1.8e−4	8.2	1.8e−3	10.0	–	–
128	3.7e−8	2.0e−7	5.4	1.4e−6	7.0	1.1e−5	7.9	9.0e−5	8.2	8.8e−4	9.8
256	2.3e−9	1.2e−8	5.2	8.6e−8	7.2	6.7e−7	7.8	5.4e−6	8.1	4.5e−5	8.3

From (2.12) we see that the error in $\hat{T}(h, s_i)$ behaves like

$$\left[\frac{\partial^3}{\partial s^3}G(s_i, s_i+) - \frac{\partial^3}{\partial s^3}G(s_i, s_i-)\right]h^4,$$

which for increasing ν behaves like $O(\nu^3)$. Consequently, we expect $\varepsilon_\nu/\varepsilon_{\nu/2}$ to approach $2^3 = 8$. The results of Table 2 provide the justification of this prediction, and we can see that as ν gets closer to $N/2$, the faster the loss of accuracy is. In [6] the number of grid points necessary to compute the oscillatory integral (2.1) with (2.3) to a sufficient accuracy is estimated by

$$N \geq 2L. \tag{2.14}$$

Table 3 presents the accuracy of the computation for $N = 8\nu, 16\nu$, $L = 2\pi$.

We observe that the error slowly decays as ν and N grow.

2.4. Improvement of the Convergence by Richardson Extrapolation (Romberg Integration)

The accuracy of the corrected trapezoidal rule can be improved by applying the Richardson extrapolation (Romberg integration). The description of this method is contained, for example, in [17]. The idea of this method is the following. The convergence of the quadrature can be essentially improved if together with a finer grid some coarser

TABLE 3
The Accuracy for the Corrected Trapezoidal Rule and the Oscillatory Kernels with $N = 8\nu, 16\nu$

N	$\nu = N/16$	$\varepsilon_{\mathrm{MAX}}$	$\varepsilon_{\mathrm{MSQ}}$	$\varepsilon_{\mathcal{L}^2}$	$\nu = N/8$	$\varepsilon_{\mathrm{MAX}}$	$\varepsilon_{\mathrm{MSQ}}$	$\varepsilon_{\mathcal{L}^2}$
16	1	4.6e−5	3.2e−5	3.2e−4	2	1.6e−4	1.1e−4	4.0e−4
32	2	9.6e−6	6.8e−6	2.4e−5	4	5.2e−5	3.7e−5	2.9e−4
64	4	3.2e−6	2.2e−6	1.8e−5	8	2.2e−5	1.6e−5	6.7e−4
128	8	1.4e−6	9.8e−7	4.2e−5	16	1.1e−5	7.6e−6	2.7e−4
256	16	6.7e−7	4.7e−7	1.7e−5	32	5.4e−6	3.8e−6	1.8e−4
512	32	3.3e−7	2.3e−7	1.1e−5	64	2.7e−6	1.9e−6	4.1e−5
1024	64	1.6e−7	1.2e−7	2.5e−6	128	1.3e−6	9.4e−7	3.6e−5
2048	128	8.2e−8	5.8e−8	2.2e−6	256	6.6e−7	4.7e−7	2.6e−5
4096	256	3.9e−8	2.7e−8	1.5e−6	512	3.1e−7	2.2e−7	1.5e−5

TABLE 4
Comparison among the Accuracies of Integration for the Kernel (2.13), with $\nu = 1$, Obtained by the Regular Trapezoidal Quadrature, the Corrected Trapezoidal Quadrature, and the Algorithm Using Two Grids Simultaneously

N	Corrected two-step rule			Corrected trapezoidal rule			Usual trapezoidal rule		
	ε_{MAX}	ε_{MSQ}	$\varepsilon_{\mathcal{L}^2}$	ε_{MAX}	ε_{MSQ}	$\varepsilon_{\mathcal{L}^2}$	ε_{MAX}	ε_{MSQ}	$\varepsilon_{\mathcal{L}^2}$
32	4.2e−8	3.0e−8	2.9e−7	2.8e−6	2.0e−6	2.0e−5	1.0e−3	7.3e−4	7.1e−3
64	6.4e−10	4.5e−10	4.4e−9	1.7e−7	1.2e−7	1.2e−6	2.6e−4	1.8e−4	1.8e−3
128	9.9e−12	7.0e−12	6.9e−11	1.1e−8	7.7e−9	7.6e−8	6.4e−5	4.5e−5	4.4e−4
256	1.6e−13	1.1e−13	1.1e−12	6.8e−10	4.8e−10	4.7e−9	1.6e−5	1.1e−5	1.1e−4
512	2.2e−14	2.2e−15	2.2e−14	4.3e−11	3.0e−11	3.0e−10	4.0e−6	2.8e−6	2.8e−5
1024	2.4e−15	7.0e−16	6.9e−15	2.7e−12	1.9e−12	1.8e−11	1.0e−6	7.1e−7	6.9e−6

grids are included in the quadrature. The asymptotic expansion of $T(h, s_i)$ contains the powers h^4, h^6, h^8, etc. The Richardson extrapolation can thus be applied as follows:

1. Pick $h_0 = L$ and define $h_k = h_0/2^k$, $k = 1, 2, \ldots$.
2. Compute $T_0^{(k)} = \hat{T}(h_k, s_i)$, $k = 0, 1, \ldots$, where $\hat{T}(h, s_i)$ is defined by (2.11).
3. Compute $T_m^{((k)}$ for $m = 1, 2, \ldots$ by the recursion

$$T_m^{(k)} = \frac{T_{m-1}^{(k+1)} - c_m T_{m-1}^{(k)}}{1 - c_m}, \qquad c_m = \left(\frac{1}{4}\right)^{m+1}. \tag{2.15}$$

Here $T_m^{(k)}$ are approximations to $(Tf)(s_i)$ and $T_m^{(k)} - (Tf)(s_i) = O(4^{-(m+2)k})$ as $k \to \infty$.

In our computations we have taken $m = 1$ (two-step) and $m = 2$ (three-step algorithm applying two grids simultaneously). This improves the accuracy in the following way: if the original quadrature has $O(h^4)$ convergence (as in our case), then by applying two grids (say, 16 and 32 points per interval) the accuracy of $O(h^6)$ can be achieved; three grids give $O(h^8)$ accuracy. However, for oscillatory integrals we cannot employ too many

TABLE 5
Accuracies of Integration for Oscillatory Kernels (2.13), with $N = 8\nu$, Obtained by the Algorithm Using Two Grids Simultaneously

ν	N	ε_{MAX}	ε_{MSQ}	$\varepsilon_{\mathcal{L}^2}$
1	8	3.6e−4	2.5e−4	2.5e−3
2	16	2.4e−5	1.7e−5	6.1e−5
4	32	4.1e−6	2.9e−6	2.3e−5
8	64	1.3e−6	9.2e−7	3.9e−5
16	128	5.6e−7	4.0e−7	1.4e−5
32	256	2.7e−7	1.9e−7	9.0e−6
64	512	1.3e−7	9.5e−8	2.1e−6
128	1024	6.7e−8	4.8e−8	1.8e−6
256	2048	3.6e−8	2.5e−8	1.4e−6
512	4096	3.7e−8	2.6e−8	1.8e−6

TABLE 6
Accuracy of the Richardson Extrapolation

ν	N	Corrected trapezoidal rule		Two-step rule		Three-step rule	
		ε_{MAX}	$\varepsilon_{N/2}/\varepsilon_N$	ε_{MAX}	$\varepsilon_{N/2}/\varepsilon_N$	ε_{MAX}	$\varepsilon_{N/2}/\varepsilon_N$
8	32	3.8e−4	–	1.2e−4	-	2.9e−4	–
	64	2.2e−5	17.2	1.3e−6	92.3	6.0e−7	415
	128	1.4e−6	15.7	1.9e−8	68.4	1.5e−9	400
	256	8.6e−8	16.3	2.9e−10	65.5	5.4e−12	278
	512	5.4e−9	15.9	4.5e−12	64.4	2.1e−14	257
	1024	3.4e−10	15.9	7.1e−14	63.4	1.0e−15	21
16	32	3.6e−3	–	1.0e−3	–	4.0e−5	–
	64	1.8e−4	20.0	4.9e−5	20.4	6.6e−5	0.6
	128	1.1e−5	14.4	5.6e−7	87.5	2.0e−7	330
	256	6.7e−7	16.4	8.2e−9	68.3	5.6e−10	357
	512	4.2e−8	16.0	1.3e−10	63.1	2.0e−12	280
	1024	2.6e−9	16.2	2.0e−12	65.0	8.1e−15	247
32	64	1.8e−3	–	2.7e−3	–	3.8e−3	–
	128	9.0e−5	20	2.3e−5	117	1.9e−5	200
	256	5.4e−6	16.7	2.7e−7	85.2	8.7e−8	218
	512	3.3e−7	16.4	4.0e−9	67.5	2.5e−10	348
	1024	2.1e−8	15.7	6.1e−11	65.6	9.3e−13	269
	2048	1.3e−9	16.2	9.6e−13	63.5	4.3e−15	216
64	128	8.8e−4	–	3.9e−3	–	5.2e−3	–
	256	4.5e−5	19.6	1.1e−5	354	5.0e−5	104
	512	2.7e−6	16.7	1.3e−7	84.6	4.2e−8	1190
	1024	1.7e−7	15.9	2.0e−9	65.0	1.2e−10	350
	2048	1.0e−8	17.0	3.0e−11	66.7	4.6e−13	261
128	256	4.4e−4	–	3.9e−3	–	5.1e−3	–
	512	2.2e−5	20.0	5.6e−6	696	5.7e−5	89.5
	1024	1.3e−6	16.9	6.7e−8	83.6	2.1e−8	2714
	2048	8.2e−8	15.9	9.8e−10	68.4	6.2e−11	339
256	512	2.2e−4	–	3.6e−3	–	4.6e−3	–
	1024	1.1e−5	20	2.8e−6	1286	5.4e−5	85.2
	2048	6.7e−7	16.4	3.3e−8	84.8	1.0e−8	5400

grids simultaneously since for very coarse grids (less than 2 to 3 points per oscillation) the result of the integration is not relevant and does not improve the one obtained for the finer grid. Tables 4 and 5 demonstrate the accuracy of $\nu = 1$ for various n and the accuracy for $N = 8\nu$. Table 6 illustrates the search for the best integration formula for various ν, $N \geq 2\nu$ (corrected trapezoidal rule and two-step and three-step Richardson extrapolation are applied).

We can see from Table 6 that, for example, for $\nu = 128$ and $N = 2\nu$ both two-step and three-step rules are less accurate than one-step rules; for $N = 4\nu = 512$ the best result (error $= 5.6 \times 10^{-6}$) is obtained by the two-step rule, while the results of the three-step rule are even worse than those obtained by the corrected trapezoidal rule; for $N = 8\nu$ each step adds to the accuracy. Thus, by applying simultaneously three grids the accuracy $O(10^{-8})$ for the integration can be achieved for $\nu = 128$, $N = 8\nu = 1024$. A similar effect can be observed for other ν.

3. DECOMPOSITION IN LOCAL FOURIER BASES (LFB) AND SPARSITY ANALYSIS

3.1. Local Fourier Bases

The local Fourier basis can capture well oscillatory patterns (see, for example, [3, 4, 6]) in contrast to wavelets or multiwavelets; therefore we attempt to get effective computation of the oscillatory integral by representing its kernel in LFB; i.e., we computed the following coefficients,

$$A_{k,l}^{i,j} = \int_0^{2\pi} \int_0^{2\pi} K(t,s) C_k^i(t) C_l^j(s)\, ds\, dt, \tag{3.1}$$

$$B_{k,l}^{i,j} = \int_0^{2\pi} \int_0^{2\pi} K(t,s) S_k^i(t) S_l^j(s)\, ds\, dt, \tag{3.2}$$

where

$$C_k^i(s) = b_i(s) \left(\frac{2}{a_{i+1} - a_i} \right)^{1/2} \cos\left(\left(k + \frac{1}{2} \right) \pi \frac{s - a_i}{a_{i+1} - a_i} \right), \tag{3.3}$$

$$S_k^i(s) = b_i(s) \left(\frac{2}{a_{i+1} - a_i} \right)^{1/2} \sin\left(\left(k + \frac{1}{2} \right) \pi \frac{s - a_i}{a_{i+1} - a_i} \right). \tag{3.4}$$

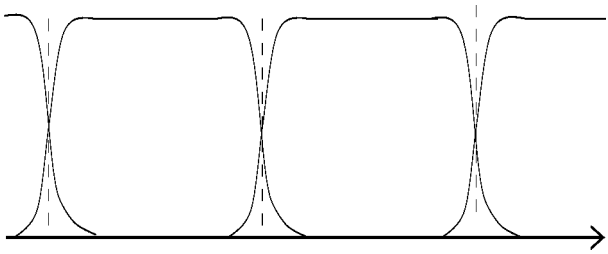
Let $[0, 2\pi]$ be divided into M intervals $I_i = (a_i, a_i + 1)$ and $b_i(s)$ be a collection of smooth window (bell) functions supported in $I_{i-1} \cup I_i \cup I_{i+1}$ such that

$$\sum_{i=1}^{M} b_i^2(s) = 1,$$

$$b_i(s) = b_{i-1}(2a_i - s).$$

A discrete system of basis functions is constructed. Let N be the number of discretization points. Following [6] we choose $N \geq 2\nu L$, i.e., $N \geq 4\pi\nu$. In all the numerical implementations we assumed $N = 8\nu$.

To get better performance we use the optimized bells that were developed by [15] which minimize the number of local Fourier coefficients necessary for their representation (3.5). These bells generalize the Coifman–Meyer (CM) construction [10] by extending the choice of the bell functions. Each bell is defined on three adjacent intervals as shown in Scheme 1,



SCHEME 1. The collection of bell functions. Each bell is defined on three adjacent intervals.

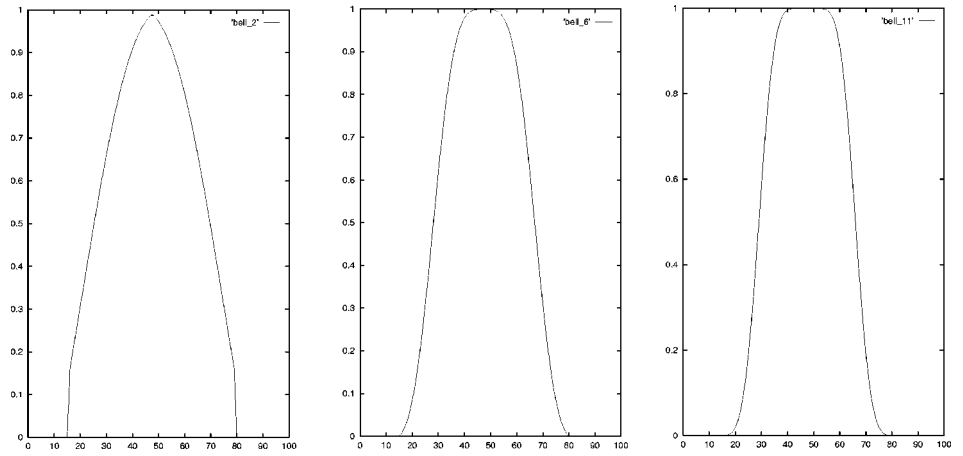


FIG. 1. The form of the bell for the steepness $k = 2, 6, 11$, respectively.

where each one is given by

$$b_k(x) = \begin{cases} \frac{1}{2}\left(1 + \sum_{n=0}^{k-1} g_n \sin(n + \frac{1}{2})\pi x\right) & \text{for all } -\frac{1}{2} \le x \le \frac{1}{2}, \\ \frac{1}{2}\left(1 + \sum_{n=0}^{k-1} (-1)^n g_n \cos(n + \frac{1}{2})\pi x\right) & \text{for all } \frac{1}{2} \le x \le \frac{3}{2}, \\ 0 & \text{otherwise} \end{cases} \tag{3.5}$$

where the coefficients g_n are the solutions of a linear system and its values for $k = 0, \ldots, 11$ are tabulated in [15]. Here k is a parameter that characterizes the smoothness of the bell and its steepness at the sides. Figure 1 illustrates the forms of the bells for $k = 1, 6, 11$.

The basis which was built corresponds to the discretization of the continuous basis functions where N discretization points x_j are taken such that a_i are midpoints between adjacent x_j.

The kernels and their discretizations were chosen so that they correspond to the corrected trapezoidal rule which was described in Section 2.2.

3.2. Numerical Results: Sparsity, Error Estimates, and Comparison with Wavelets

Tables 7 and 8 present the number of coefficients above a threshold ε in the representation of the kernel

$$G(t, s) = \sin\left(\nu\sqrt{(\cos(t) - \cos(s))^2 + (\sin(t) - \sin(s))^2}\right), \qquad t, s \in [0, 2\pi],$$

in local Fourier basis. The discretization which corresponds to the corrected trapezoidal rule was described in Section 2.2, i.e.,

$$G_{ij} = \frac{2\pi}{N} G(t_i, t_j) + \nu \frac{2\pi^2}{3N^2},$$

where the grid points are equispaced $t_i = 2\pi i/N$. Table 7 presents the number of significant coefficients in the LFB kernel representation and the accuracy obtained (the

TABLE 7

Sparsity of the Representation of the Oscillatory Kernel ($N = 8\nu$ is a Number of Points in Each Direction) in LFB and the Accuracy Obtained with the Corrected Trapezoidal Rule (2.12) When the Coefficients below the Threshold Are Neglected

Threshold ε	ν of exp.	N of points	Total coefficients	Coefficients above ε	%	Integration error	Accuracy with thresholding
1e−3	8	64	4096	2067	50%	2.2e−5	2.6e−3
	32	256	65536	7269	11%	5.4e−6	4.8e−3
	128	1024	1048576	17654	1.7%	1.3e−6	2.8e−3
1e−4	8	64	4096	3381	91%	2.2e−5	2.0e−4
	32	256	65536	20304	31%	5.4e−6	4.1e−4
	128	1024	1048576	66307	6.3%	1.3e−6	4.3e−4
1e−5	8	64	4096	3932	96%	2.2e−5	3.5e−5
	32	256	65536	33947	52%	5.4e−6	7.5e−5
	128	1024	1048576	159943	15%	1.3e−6	1.3e−4
1e−6	8	64	4096	4073	99%	2.2e−5	2.2e−5
	32	256	65536	50217	77%	5.4e−6	1.0e−5
	128	1024	1048576	300998	29%	1.3e−6	1.5e−5
1e−7	8	64	4096	4093	99%	2.2e−5	2.2e−5
	32	256	65536	58773	90%	5.4e−6	5.5e−6
	128	1024	1048576	513705	49%	1.3e−6	2.0e−6

Note. "Integration error" describes the accuracy of the quadrature, while the "accuracy" column gives the computation error when all the coefficients below a certain threshold are neglected.

TABLE 8

Comparison between the Sparsity Representation of the Oscillatory Kernel ($N = 1024$) Both in Local Fourier Basis and the Multiwavelet Basis

Threshold ε	ν of exp.	Integration error	Accuracy with thresholding	LFB coefficients $> \varepsilon$ per row	Multiwavelet coefficients per row
1e−3	1	2.7e−12	4.8e−3	4.5	2
	4	4.8e−11	4.0e−3	4.8	5
	16	2.6e−9	3.9e−3	6.9	19
	64	1.6e−7	3.9e−3	12.7	63
	128	1.3e−6	2.8e−3	17.2	97
1e−4	1	2.7e−12	1.2e−3	18.4	6
	4	4.8e−11	1.5e−3	17.4	14
	16	2.6e−9	5.5e−4	21.8	46
	64	1.6e−7	1.2e−3	45.9	184
	128	1.3e−6	4.3e−4	64.8	306
1e−5	1	2.7e−12	1.7e−4	47.6	15
	4	4.8e−11	2.6e−4	48.8	30
	16	2.6e−9	1.4e−4	66.1	87
	64	1.6e−7	1.5e−4	118.0	280
	128	1.3e−6	1.3e−4	156.2	389
1e−6	1	2.7e−12	2.7e−5	112.3	30
	4	4.8e−11	1.9e−5	115.7	56
	16	2.6e−9	1.8e−5	140.1	138
	64	1.6e−7	1.5e−5	218.6	348
	128	1.3e−6	3.6e−5	293.9	425

maximal error) when the coefficients below the threshold are ignored. This accuracy is compared with the accuracy obtained after integration (without thresholding). The sparsity is estimated both by the number of coefficients above a threshold and the percent of significant coefficients.

For accuracies 10^{-5} to 10^{-6} the error obtained after thresholding is close to the error of the integration quadrature. Thus, for $\nu = 128$ and number of grid points $N = 1024$ it is senseless to consider accuracies less than 10^{-6} (see Table 3). One can see that if all the coefficients below a certain threshold ε are omitted, the error obtained is close to ε. Due to the oscillatory nature of the kernel there is nearly no accumulation of error.

Table 8 describes the dependence of the sparsity on ν, where $N = 1024$. In addition to LFB we use multiwavelets [1] with four vanishing moments. The last column in the table gives the number of multiwavelet coefficients per row above the chosen threshold which can be compared with the sixth column that presents the number of elements above the threshold after the LFB.

We observe that for small ν ($\nu < 16$) the number of wavelet coefficients above a threshold is less than the number of LFB coefficients, while for large ν ($\nu > 16$) LFB gives a more compact representation for the kernel. This illustrates the fact that wavelet representation is not efficient for highly oscillatory integral kernels.

3.3. Sparsification of the Kernel by Local Fourier Bases: Theoretical Estimates

We compute the following integral

$$T(f(x)) = \int_0^1 e^{i\nu\phi(x,y)} f(y)\, dy, \tag{3.6}$$

where ϕ is a smooth function satisfying

$$\left| \frac{\partial^{\alpha+\beta}\phi(x,y)}{\partial x^{\alpha}\partial y^{\beta}} \right| \le C. \tag{3.7}$$

If a one-dimensional function is decomposed with local cosine basis, each coefficient of such a decomposition corresponds to a certain location (window) q and a certain frequency p. Let us assume that we have $N = 4\nu$ discretization points (less than two points per oscillation would not work) and $\sqrt{N} = 2\sqrt{\nu}$ bells (windows for local cosine transform). Then, the number of the window q is such that $0 \le q < 2\sqrt{\nu}$ and frequency p is such that $0 \le p < 2\sqrt{\nu}$.

Let us denote by $c_q = q/(2\sqrt{\nu})$ the origin of the qth interval and by $d_q = (q + 1/2)/(2\sqrt{\nu})$ the center of the qth window and bell. Then the basis function corresponding to location q and frequency p is

$$C_{q,p}(x) = 2\sqrt{\nu}\, b(2\sqrt{\nu}(x - d_q)) \cos\left(\left(p + \frac{1}{2} \right) \pi (x - c_q) 2\sqrt{\nu} \right), \tag{3.8}$$

where $b(x)$ is a bell function.

The collection of functions $C_{q,p}$ forms an orthogonal basis. Thus the interaction between two windows q and q' and two frequencies p and p' is given by the expression

$$
\begin{aligned}
G_{(q',p')(q,p)} &= \langle C_{q',p'}, T(C_{q,p})\rangle \\
&= 2\sqrt{\nu}\int_0^1\int_0^1 e^{i\nu\phi(x,y)} b\big((x-d_{q'})2\sqrt{\nu}\big) b\big((y-d_q)2\sqrt{\nu}\big) \\
&\quad \times \cos\left(\left(p'+\frac{1}{2}\right)\pi\left(x-\frac{q'}{2\sqrt{\nu}}\right)2\sqrt{\nu}\right) \\
&\quad \times \cos\left(\left(p+\frac{1}{2}\right)\pi\left(y-\frac{q}{2\sqrt{\nu}}\right)2\sqrt{\nu}\right) dx\,dy. \qquad (3.9)
\end{aligned}
$$

By the change of variables

$$
s=(x-d_{q'})2\sqrt{\nu}, \qquad t=(y-d_q)2\sqrt{\nu}, \qquad (3.10)
$$

we obtain

$$
\begin{aligned}
2x\sqrt{\nu} = s + d_{q'}2\sqrt{\nu} \Rightarrow x = \frac{s}{2\sqrt{\nu}} + d_{q'} \Rightarrow x = \frac{s}{2\sqrt{\nu}} + \frac{q'+1/2}{2\sqrt{\nu}}, \\
2y\sqrt{N} = t + d_q 2\sqrt{\nu} \Rightarrow y = \frac{t}{2\sqrt{\nu}} + d_q \Rightarrow y = \frac{t}{2\sqrt{\nu}} + \frac{q+1/2}{2\sqrt{\nu}}.
\end{aligned}
$$

The boundaries for t and s are

$$
s_1 = -q' - \frac{1}{2} \le s \le 2\sqrt{N} - q' - \frac{1}{2} = s_2, \qquad t_1 = -q - \frac{1}{2} \le s \le 2\sqrt{N} - q - \frac{1}{2} = t_2, \qquad (3.11)
$$

and

$$
x - \frac{q'}{2\sqrt{\nu}} = x - \frac{q'+1/2}{2\sqrt{\nu}} + \frac{1}{4\sqrt{\nu}} = \frac{s}{2\sqrt{\nu}} + \frac{1}{4\sqrt{\nu}}, \qquad (3.12)
$$

$$
y - \frac{q}{2\sqrt{\nu}} = x - \frac{q+1/2}{2\sqrt{\nu}} + \frac{1}{4\sqrt{\nu}} = \frac{t}{2\sqrt{\nu}} + \frac{1}{4\sqrt{\nu}}. \qquad (3.13)
$$

The change of variables (3.10) turns (3.9) into the following integral:

$$
\begin{aligned}
&\frac{1}{2\sqrt{\nu}}\int_{t_1}^{t_2}\int_{s_1}^{s_2} \exp\left\{ i\nu\phi\left(\frac{q'}{2\sqrt{\nu}} + \frac{s}{2\sqrt{\nu}}, \frac{q}{2\sqrt{\nu}} + \frac{t}{2\sqrt{\nu}}\right)\right\} \\
&\quad \times b(s)b(t)\cos\left(\left(p'+\frac{1}{2}\right)\pi\left(s-\frac{1}{2}\right)\right)\cos\left(\left(p+\frac{1}{2}\right)\pi\left(t-\frac{1}{2}\right)\right) ds\,dt.
\end{aligned}
$$

Let us expand function $\nu\phi$ into Taylor series:

$$
\begin{aligned}
&\nu\phi\left(\frac{q'}{2\sqrt{\nu}} + \frac{s}{2\sqrt{\nu}}, \frac{q}{2\sqrt{\nu}} + \frac{t}{2\sqrt{\nu}}\right) \\
&\quad = \nu\phi\left(\frac{q'}{2\sqrt{\nu}}, \frac{q}{2\sqrt{\nu}}\right) + \frac{\sqrt{\nu}}{2}\left[\frac{\partial\phi}{\partial q'}\left(\frac{q'}{2\sqrt{\nu}}, \frac{q}{2\sqrt{\nu}}\right)s + \frac{\partial\phi}{\partial q}\left(\frac{q'}{2\sqrt{\nu}}, \frac{q}{2\sqrt{\nu}}\right)t\right] + r^{\nu}_{qq'}(s,t).
\end{aligned}
$$

Therefore, the residual has the following form:

$$
\begin{aligned}
r^{\nu}_{qq'}(s,t) &= \nu\phi\left(\frac{q'}{2\sqrt{\nu}} + \frac{s}{2\sqrt{\nu}}, \frac{q}{2\sqrt{\nu}} + \frac{t}{2\sqrt{\nu}}\right) - \nu\phi\left(\frac{q'}{2\sqrt{\nu}}, \frac{q}{2\sqrt{\nu}}\right) \\
&\quad - \frac{\sqrt{\nu}}{2}\left[\frac{\partial\phi}{\partial q'}\left(\frac{q'}{2\sqrt{\nu}}, \frac{q}{2\sqrt{\nu}}\right)s + \frac{\partial\phi}{\partial q}\left(\frac{q'}{2\sqrt{\nu}}, \frac{q}{2\sqrt{\nu}}\right)t\right]. \qquad (3.14)
\end{aligned}
$$

We assume $|s| < 1$, $|t| < 1$, which corresponds to overlapping less than half a window. This provides the convergence of the Taylor series. The original integral, which defines the interaction between two windows, has the following form,

$$\frac{1}{2\sqrt{\nu}} e^{i\nu\phi(q'/(2\sqrt{\nu}),q/(2\sqrt{\nu}))} \int_{t_1}^{t_2}\int_{s_1}^{s_2} \beta(s,t) \times \cos\left(\left(p'+\frac{1}{2}\right)\pi\left(s+\frac{1}{2}\right)\right)\cos\left(\left(p+\frac{1}{2}\right)\pi\left(t+\frac{1}{2}\right)\right) e^{i(As+Bt)}\,ds\,dt, \quad (3.15)$$

where

$$\beta(s,t) = b(s)b(t)e^{ir^{\nu}_{qq'}(s,t)} \quad (3.16)$$

and the coefficients A, B are

$$A = \frac{\sqrt{\nu}}{2}\frac{\partial\phi}{\partial q'}\left(\frac{q'}{2\sqrt{\nu}},\frac{q}{2\sqrt{\nu}}\right), \qquad B = \frac{\sqrt{\nu}}{2}\frac{\partial\phi}{\partial q}\left(\frac{q'}{2\sqrt{\nu}},\frac{q}{2\sqrt{\nu}}\right). \quad (3.17)$$

In addition to local cosine bases, the function can be expanded into local sine bases. We next represent cos as a combination of exponentials. To compute (3.15) we have to evaluate the following integral:

$$\begin{aligned}
&\frac{1}{2\sqrt{N}} e^{i\nu\phi(q'/(2\sqrt{\nu}),q/(2\sqrt{\nu}))} \sum_{\pm}\int_{t_1}^{t_2}\int_{s_1}^{s_2} \beta(s,t)\exp\left\{i\left[\pm\left(p'+\frac{1}{2}\right)\pi\left(s+\frac{1}{2}\right)+As\right]\right\}\\
&\quad\times\exp\left\{i\left[\pm\left(p+\frac{1}{2}\right)\pi\left(t+\frac{1}{2}\right)+Bt\right]\right\}ds\,dt\\
&=\frac{1}{2\sqrt{\nu}}\exp\left\{i\nu\phi\left(\frac{q'}{2\sqrt{\nu}},\frac{q}{2\sqrt{\nu}}\right)\right\}\sum_{\pm}\exp\left\{\pm i\left(p'+\frac{1}{2}\right)\frac{\pi}{2}\right\}\exp\left\{\pm i\left(p+\frac{1}{2}\right)\frac{\pi}{2}\right\}\\
&\quad\times\int_{t_1}^{t_2}\int_{s_1}^{s_2}\beta(s,t)\exp\left\{i\left[\pm\left(p'+\frac{1}{2}\right)\pi+A\right]s\right\}\exp\left\{i\left[\pm\left(p+\frac{1}{2}\right)\pi+B\right]t\right\}ds\,dt\\
&=\frac{1}{2\sqrt{\nu}}\exp\left\{i\nu\phi\left(\frac{q'}{2\sqrt{\nu}},\frac{q}{2\sqrt{\nu}}\right)\right\}\sum_{\pm}\exp\left\{\pm i\left(p'+\frac{1}{2}\right)\frac{\pi}{2}\right\}\exp\left\{\pm i\left(p+\frac{1}{2}\right)\frac{\pi}{2}\right\}\\
&\quad\times\hat{\beta}\left(A\pm\left(p'+\frac{1}{2}\right)\pi, B\pm\left(p+\frac{1}{2}\right)\pi\right).
\end{aligned} \quad (3.18)$$

Let us evaluate $\hat{\beta}(\xi,\eta)$, which is a Fourier transform of

$$\beta(x,y) = b(x)b(y)e^{ir^{\nu}_{qq'}(x,y)},$$

where $r^{\nu}_{qq'}(x,y)$ is the residual after the linear approximation (3.14).

We claim that there exists such a constant K that $|r^{\nu}_{qq'}(x,y)| \le K$ and

$$\left|\frac{\partial^{i+j}}{\partial x^i\partial y^j}r_{\nu}(x,y)\right| \le \frac{K}{(\sqrt{\nu})^{i+j}}. \quad (3.19)$$

We observe that the residual

$$r^{\nu}_{qq'}(x,y) = \nu\left[\phi(x,y) - \phi\left(\frac{q'}{2\sqrt{\nu}},\frac{q}{2\sqrt{\nu}}\right) - As - Bt\right]$$

has a leading term which is a sum of three terms,

$$N\left[a\left(\frac{s}{2\sqrt{\nu}}\right)^2+b\left(\frac{t}{2\sqrt{\nu}}\right)^2+c\frac{st}{4\nu}\right],$$

where

$$\begin{aligned} a&=\frac{1}{4}\frac{\partial^2\phi}{\partial q'^2}\left(\frac{q'}{2\sqrt{\nu}},\frac{q}{2\sqrt{\nu}}\right), \qquad b=\frac{1}{2}\frac{\partial^2\phi}{\partial q'\partial q}\left(\frac{q'}{2\sqrt{\nu}},\frac{q}{2\sqrt{\nu}}\right),\\ c&=\frac{1}{4}\frac{\partial^2\phi}{\partial q^2}\left(\frac{q'}{2\sqrt{\nu}},\frac{q}{2\sqrt{\nu}}\right). \end{aligned} \tag{3.20}$$

Each of the terms is a constant; therefore the residual is bounded. The derivative of the residual is a sum of four terms; each of them is the third partial derivative of ϕ (in t, s, or mixed) with an appropriate coefficient multiplied by

$$\frac{s^i t^j}{\sqrt{N}}, \qquad i+j=4.$$

The third derivatives of ϕ are bounded due to the assumptions; consequently

$$\frac{\partial r^{\nu}_{qq'}(s,t)}{\partial s}=O\left(\frac{1}{\sqrt{\nu}}\right).$$

Similarly all the partial derivatives of $r^{\nu}_{qq'}(s,t)$ of order k are $O(1/\nu^{k/2})$, which proves (3.19).

The function

$$\beta(s,t)=b(s)b(t)e^{ir^{\nu}_{qq'}(s,t)}$$

is bounded together with all the partial derivatives of β up to order M:

$$\left|\frac{\partial^{i+j}\beta(s,t))}{\partial s^i\partial t^j}\right|\le C, \qquad i+j\le M. \tag{3.21}$$

It is known that if a function f is bounded together with its derivatives

$$\left|\frac{\partial^{i+j}f(s,t)}{\partial s^i\partial t^j}\right|\le C, \qquad i+j\le M, \tag{3.22}$$

then the Fourier transforms of these derivatives are also bounded:

$$\left|\left(\widehat{\frac{\partial^{i+j}f(s,t)}{\partial s^i\partial t^j}}\right)\right|=\left|\hat{f}(\xi,\eta)\xi^i\eta^j\right|\le C_1. \tag{3.23}$$

Moreover,

$$\left|\hat{f}(\xi,\eta)\xi^i\eta^j\right|\le\frac{C_1}{1+\max(|\xi|,|\eta|)^M}.$$

Thus

$$\left|\hat{\beta}(\xi,\eta)\xi^i\eta^j\right|\le\frac{C_1}{1+\max(|\xi|,|\eta|)^M}.$$

This means that for large ξ, η the Fourier coefficients are negligible. For example, if ξ or $\eta > 10$ then

$$\left|\hat{\beta}(\xi, \eta)\right| \le \frac{C_1}{1 + 10^M},$$

which means that the coefficients corresponding to frequencies $|\xi| + |\eta| > 10$ are negligible. Thus there is only a constant number C_2 of significant Fourier coefficients for each of the two windows corresponding to q and q'. We have $2\sqrt{\nu}$ windows, so there are 4ν pairs of interactions for windows located at q and q' $(0 < q, q' \le 2\sqrt{\nu})$; not more than C_2 coefficients above a certain threshold for each interaction lead to $4\nu C_2$ significant coefficients among the total of $16\nu^2$ coefficients. Thus, the total number of coefficients above a threshold is $O(N)$, where N is a number of grid points in each direction.

We have already observed that nearly all the coefficients in the power of the exponents in (3.18) are between -10 and 10:

$$\left|\left(p' + \frac{1}{2}\right) \pm \frac{\sqrt{\nu}}{2\pi} \frac{\partial \phi}{\partial q'}\left(\frac{q'}{2\sqrt{\nu}}, \frac{q}{2\sqrt{\nu}}\right)\right| \le \frac{10}{\pi} \approx 3, \tag{3.24}$$

$$\left|\left(p + \frac{1}{2}\right) \pm \frac{\sqrt{\nu}}{2} \frac{\partial \phi}{\partial q}\left(\frac{q'}{2\sqrt{\nu}}, \frac{q}{2\sqrt{\nu}}\right)\right| \le \frac{10}{\pi} \approx 3. \tag{3.25}$$

Thus the neighborhood of the strongest interaction can be found: first, from (3.24) the domain of q can be computed (p', q' are given and we are looking for frequencies p and locations q which interact essentially with a chosen frequency p' and location q'); second, the computed domain for q is substituted into (3.25) and thus the domain for p can be evaluated. For finding p, q the nondegeneration condition is assumed:

$$\det \begin{pmatrix} \dfrac{\partial^2 \phi}{\partial x^2}(x, y) & \dfrac{\partial^2 \phi}{\partial x \partial y}(x, y) \\ \dfrac{\partial^2 \phi}{\partial x \partial y}(x, y) & \dfrac{\partial^2 \phi}{\partial y^2}(x, y) \end{pmatrix} \neq 0.$$

For each domain of pairs (q, p) we find domain of pairs (q', p') interacting with it. We deal not with points but with "clouds" (corresponding to close locations and frequencies). Such an approach is natural in quantum mechanics where we can follow not the points but only the clouds.

In Section 3.4 the number of significant coefficients after LFB of the function

$$\beta(x, y) = b(x)b(y)e^{ir^{\nu}_{qq'}(x,y)}$$

is evaluated numerically.

3.4. *Numerical Estimates for the LFB of the Residual: Dependence on the Bell and the Form of the Curve*

We evaluate numerically the following integrals:

$$G_{(p,q),(p',q')} = \int_0^1 b(x)b(y)e^{i(p+1/2)x}e^{i(p'+1/2)y}e^{ir^{\nu}_{qq'}(x,y)}\,dx\,dy, \tag{3.26}$$

which is the LFB of the residual.

TABLE 9
Number of Coefficients with an Imaginary Part above a Threshold in Blocks 128×128, with 64 Points Overlapping, $\nu = 2048$

Location of the block	10^{-5}	10^{-6}	10^{-7}	10^{-8}
At the diagonal	804	1264	1543	2271
Below or above the diagonal	169	662	1068	1747
Other blocks	15–30	27–50	50–65	150–350

As a model problem we take the case as in [6], when the total number of points in each direction is $128^2 = 10^{14} = 16384$, 128^2 windows, and the overlapping of 64 points from each end of the window. The chosen bell from the family (3.5) is $k = 10$. Most of the significant coefficients describe the interaction of a window with itself, which belongs to the diagonal blocks or to the blocks adjacent to the diagonal. Inside the window the significant coefficients correspond to pp' being small (see (3.23)).

First consider a circle:

$$z(s) = \cos(2\pi s) + i \sin(2\pi s). \tag{3.27}$$

The number of significant coefficients in (3.26) varies and it depends on the distance of a block from the diagonal.

The total number of coefficients with the imaginary part above the threshold 10^{-5} is about 0.18%, above 10^{-6} it is less than 0.5% of all coefficients, and for 10^{-7} it is less than 0.8%. The absolute value of the real parts is even less, so the same estimates are valid for both of them.

If the residual does not contain the second derivatives in x and in y, the matrix obtained after the application of LFB is sparser: 0.18% of the coefficients above the threshold 10^{-5}, 0.35% above 10^{-6}, and 0.55% above 10^{-7} (see Table 9).

Table 10 describes the dependence of the number of coefficients above a threshold on the steepness of the bell [15] for a block on the diagonal.

We observe that with a steeper bell function the number of significant coefficients is less. However, for a smooth bell coefficients of the LFB transform of the residual are a smooth function or can be easily reduced to a smooth function which can be additionally compressed, for example, by application of wavelets. If a bell is steep, the LFB transform of the bell is an oscillatory function and so is the LFB transform of the residual. A bell with

TABLE 10
Number of Coefficients with an Imaginary Part above a Threshold in a Diagonal Block 128×128, with 64 Points Overlapping, $\nu = 2048$ for Various Steepness of the Bell

Steepness of the bell	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
2	1225	7379	16384	16384	16384	16384	16384
4	303	755	2090	15490	16383	16384	16384
7	36	292	776	1048	1938	8926	16384
10	36	67	332	804	1264	1543	2271
11	38	63	266	773	1054	1529	1994

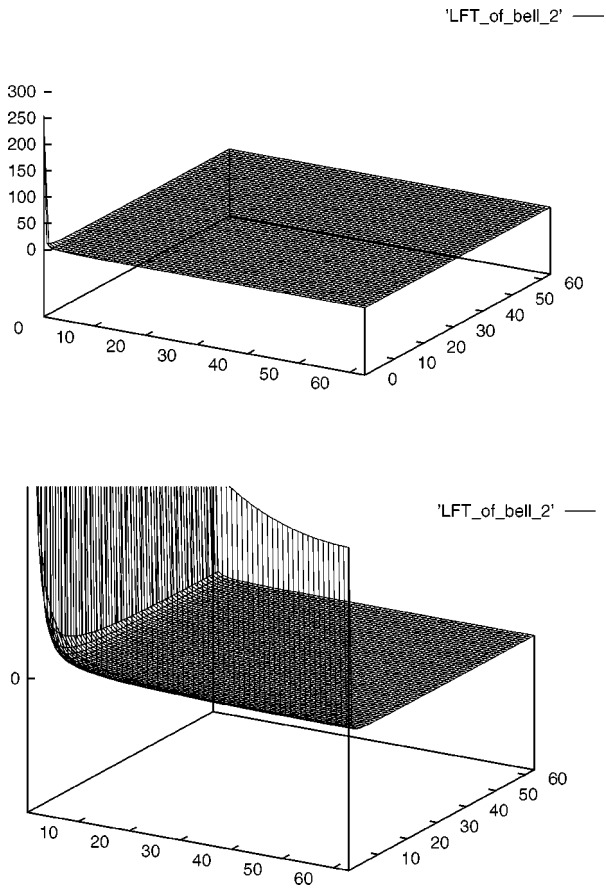


FIG. 2. The imaginary part after the expansion in LFB of the residual of the diagonal block for a bell with steepness $k = 2$ for the full range of coefficients and the range $[-0.1; 0.1]$.

a small steepness number is not smooth (see Fig. 1); however, its local Fourier transform is easily reduced to smooth functions: for example, the coefficients located at odd/even places in each direction form four (odd in x/odd in y, even/odd, odd/even, even/even) smooth functions. Figures 2 and 3 present the odd/odd coefficients of the local Fourier transform (imaginary part) of the residual for the diagonal block: first, in all the values of coefficients and second, for the range of values from -0.1 to 0.1.

Next we consider the dependence of the number of coefficients above a threshold on the number of grid points N, where $v = N/8$. Table 11 presents the number of coefficients above a threshold for a diagonal block, while Table 12 presents it for a block which is far away from the diagonal.

The theoretical estimate for the number of significant coefficients was $O(N)$ (see Section 3.3). Here we observe that for a diagonal block the number of significant coefficients is doubled when the number of points $\sqrt{N}$ in each direction is doubled. Besides, the number of diagonal blocks grows as $\sqrt{N}$. Thus, the contribution of the diagonal blocks (which is the same for each block on the diagonal due to the symmetry of the circle) has an order of $O(N)$. The same estimate is valid for the blocks near the diagonal. The number of coefficients above a certain threshold in blocks, which are

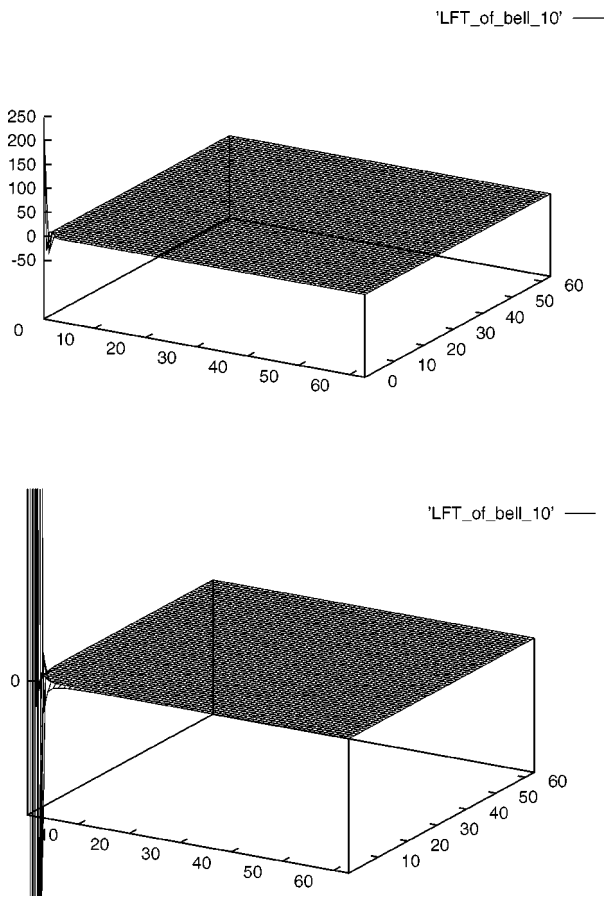


FIG. 3. The imaginary part of the local Fourier transform of the residual for the bell with steepness $k = 10$ for the full range of coefficients and the range $[-0.1; 0.1]$. The transform was applied to the diagonal block.

nonadjacent to the diagonal, is approximately constant (in fact, it even decreases for large threshold with the growth of N). The total number of blocks is $\sqrt{N}\sqrt{N} = N$. Therefore, the contribution of these blocks does not exceed $O(N)$; in fact it is even less. Due to the slow growth in the number of significant coefficients, the percentage of such coefficients

TABLE 11
The Number of Coefficients with an Imaginary Part above a Threshold in Diagonal Blocks of Various Sizes, where the Overlapping is Half a Window, $\nu = N/8$

Block size \ Threshold	ν	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
$16 \times 16 = 256$	32	36	67	106	141	170	218
$32 \times 32 = 1024$	128	47	91	170	249	339	448
$64 \times 64 = 4096$	512	57	183	404	598	749	1006
$128 \times 128 = 16384$	2048	67	332	804	1264	1543	2271
$256 \times 256 = 65536$	8192	87	606	1589	2570	3198	4827

TABLE 12
The Number of Coefficients with an Imaginary Part above a Threshold in Blocks of Various Sizes Which Are Far Away from the Diagonal, where the Overlapping is Half a Window, $\nu = N/8$

Block size \ Threshold	ν	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
$16 \times 16 = 256$	32	12	20	31	62	101	134
$32 \times 32 = 1024$	128	7	16	27	46	90	176
$64 \times 64 = 4096$	512	5	12	22	37	71	241
$128 \times 128 = 16384$	2048	5	11	18	35	59	321
$256 \times 256 = 65536$	8192	3	8	17	26	49	115

decays quickly: from 24% ($N = 256$) to 0.03% ($N = 65{,}536$), if a block is distant from the diagonal, and from 55% ($N = 256$) to 3.9% ($N = 65{,}536$) for diagonal blocks.

Consider now an elliptic curve. Here there is no leading contribution by the diagonal or adjacent to the diagonal blocks. As a model example the following ellipse was chosen:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2}, \qquad a = 3, \qquad b = 0.5.$$

All the blocks can be classified into two types: those corresponding to strongly interacting windows and those corresponding to weakly interacting windows.

The total number of coefficients with an imaginary part above the threshold 10^{-5} is less than 0.2%, above 10^{-6} it is less than 0.4% of all coefficients, and for 10^{-7} it is not more than 0.7% (see Table 13).

Table 14 presents the number of significant coefficients for a certain block with strong interaction for various eccentricities of the ellipse.

3.5. Smoothing of Functions by Local Fourier Transform

We consider LFB expansion of the residual (3.26). We observe that it is still oscillatory (see Fig. 3). There are at least two main reasons why oscillations arise: first, the bell is shifted with respect to the basis function, and second, the Fourier transform of the bell is not, generally speaking, a smooth function (see Fig. 4).

The first source of oscillations can be eliminated as follows,

$$G_{(p,q),(p',q')} = \int_0^1 \int_0^1 b\big(2\sqrt{\nu}(x - d_{q'})\big) e^{i(p'+1/2)\pi(x-c_{q'})} \times b\big(2\sqrt{\nu}(y - d_q)\big) e^{i(p'+1/2)\pi(y-c_q)} e^{i r^{\nu}_{qq'}(x,y)}\, dx\, dy, \tag{3.28}$$

TABLE 13
The Number of Coefficients with an Imaginary Part above the Threshold in Blocks 128×128**, where Overlapping is 64 Points,** $\nu = 2048$**, and the Eccentricity of the Ellipse is** $3/0.5$

Type of block	10^{-5}	10^{-6}	10^{-7}	10^{-8}
Strong interaction	80–150	100–200	150–250	350–600
Weak interaction	0–40	0–70	3–90	60–300

TABLE 14
The Number of Coefficients with an Imaginary Part above the Threshold in Blocks 128×128, where Overlapping is 64 Points, $\nu = 2048$, for Ellipses with Various Eccentricities a/b

$a:b$	10^{-5}	10^{-6}	10^{-7}	10^{-8}
1 : 1	18	33	49	251
1.2 : 0.8	35	51	86	370
2 : 0.7	76	98	137	318
3 : 0.5	114	160	220	370
4 : 0.4	150	214	299	564
5 : 0.1	192	280	383	758
3 : 2	71	101	154	405
3 : 1.5	95	130	178	553
3 : 1	104	151	202	398
3 : 0.5	114	160	220	370
3 : 0.1	107	161	419	223

where $c_q = 1/(2\sqrt{\nu})$ is the left end of the qth window and $d_q = c_q + 1/(4\sqrt{\nu})$ is the center of the qth window. Integral (3.28) can be rewritten as follows,

$$G_{(p,q),(p',q')} = e^{i(p'+1/2)\pi(d_{q'}-c_{q'})}\left(\int_0^1\int_0^1 b\big(2\sqrt{\nu}(x-d_{q'})\big)e^{i(p'+1/2)\pi(x-d_{q'})} \times b\big(2\sqrt{\nu}(y-d_q)\big)e^{i(p'+1/2)\pi(y-d_q)}e^{ir^{\nu}_{qq'}(x,y)}\right) \times e^{i(p+1/2)\pi(d_q-c_q)}\,dx\,dy, \tag{3.29}$$

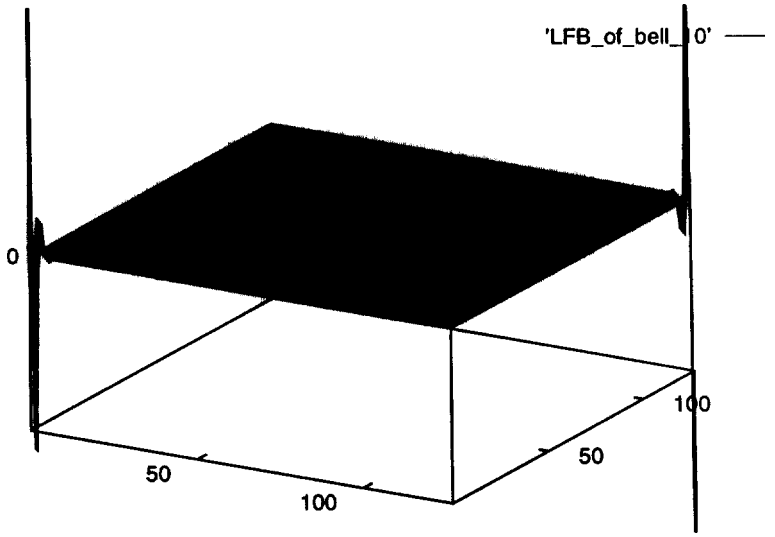


FIG. 4. The real part of the LFB transform of the kernel (2.1) $K(t,s) \equiv 1$ and the bell with steepness 10 in the range of values $[-0.1; 0.1]$.

where in the integral the bells and the basis functions are centered at the same point. Denote

$$a_{pq} = e^{i(p+1/2)\pi(d_q - c_q)}.$$

We can represent the matrix of coefficients $G = G_{(p,q),(p',q')}$ as a product of three matrices,

$$G = AUA, \tag{3.30}$$

where A is a diagonal matrix with a_{pq} at the diagonal. The components of U are

$$u_{(p,q),(p',q')} = \int_0^1 \int_0^1 b\big(2\sqrt{\nu}(x - d_{q'})\big) e^{i(p'+1/2)\pi(x-d_{q'})} \times b\big(2\sqrt{\nu}(y - d_q)\big) e^{i(p'+1/2)\pi(y-d_q)} e^{ir^{\nu}_{qq'}(x,y)}\, dx\, dy. \tag{3.31}$$

For these integrals the centers of the bell and basis functions coincide. If the transform of the bell is smooth and the residual is small, then the matrix obtained is smooth and can be additionally compressed, for example, by wavelets. This is the case of the bell of a Gaussian type (see [6]) or a bell which is considered here with a low steepness k. For example, the LFB transform of the diagonal block with the bell corresponding to $k = 2$ in Fig. 2 was obtained for a bell centered at the centers of basis functions. This is equivalent to the multiplication of G in (3.30) by the diagonal matrix A^{-1} from the left and from the right. Such a matrix can be additionally compressed, for example, if the transform of the diagonal block is expanded with the biorthogonal wavelets [7] (odd and even elements separately). Table 15 presents the number of coefficients above thresholds 10^{-5} and 10^{-6} after LFB expansion with $k = 2$ and additionally compressed by orthogonal wavelets. The matrix, which is expanded, is a diagonal block of the kernel corresponding to a circle and the ellipse with axes 3 and 0.5.

However, for bells with a small k the number of coefficients above thresholds 10^{-5} to 10^{-8} is large. Even after the compression by wavelets there are more significant coefficients than after LFB expansion with a steep bell with $k = 10$ (see Table 9).

On the contrary, if we choose a steep bell ($k = 10, 11$), the number of coefficients above thresholds 10^{-6} to 10^{-8} is small (0.5–1.5%). However, the matrix obtained is oscillatory even if the bells and the basis functions are centered at the same points. This is stipulated

TABLE 15

The Number of Coefficients above a Threshold in a Diagonal Block after LFB Expansion with $k = 2$ and after the Additional Expansion with Biorthogonal Wavelets, where the Curve is either a Circle or an Ellipse and the Eccentricity of the Ellipse is $3/0.5$

Type of curve	Threshold ε	Number of coefficients after LFB	Number of coefficients after LFB+wavelets
circle	10^{-5}	6202	755
	10^{-6}	8290	1312
ellipse	10^{-5}	6058	607
3.0/0.5	10^{-6}	8688	1504

by the oscillatory character of the transform of the bell. Denote by $\hat{b}$

$$\hat{b}(p) = \int_0^1 b\big(2\sqrt{\nu}(x-d_q)\big)e^{i(p+1/2)\pi(x-d_q)}\,dx \tag{3.32}$$

the local Fourier transform of the bell, which obviously does not depend on q. Let us choose a smooth function $\varphi(p)$ such that the ratio $\varphi(p)/\hat{b}(p)$ is not very big and not very small; for example,

$$0.01 < \left|\frac{\varphi(p)}{\hat{b}(p)}\right| < 100.$$

If the residual is small, then the coefficients u in (3.31) are close to the transform of the bell; thus the coefficients

$$\begin{aligned} v_{(p,q),(p',q')} = {} & \frac{\varphi(p)}{\hat{b}(p)} \int_0^1 \int_0^1 b\big(2\sqrt{\nu}(x-d_{q'})\big)e^{i(p'+1/2)\pi(x-d_{q'})}\,dx \\ & \times \frac{\varphi(p')}{\hat{b}(p')} b\big(2\sqrt{\nu}(y-d_q)\big)e^{i(p'+1/2)\pi(y-d_q)}e^{ir^{\nu}_{qq'}(x,y)}\,dy \end{aligned} \tag{3.33}$$

form a smooth function. Consequently, (3.30) can be rewritten as

$$G = AFVFA, \tag{3.34}$$

where the elements $v_{(p,q),(p',q')}$ of the matrix V are defined by (3.33) and F is a diagonal matrix with

$$f_p = \frac{\hat{b}(p)}{\varphi(p)} \tag{3.35}$$

at the diagonal (the same for each q). Figure 5 compares the matrices U and V for a nondiagonal block. The number of points in the window is 128×128. U is oscillatory while V is smooth. In Fig. 6 matrices U and V are shown at a higher resolution: between -10^{-6} and 10^{-6}.

There are two open problems connected with the smoothing of matrix V obtained by LFB expansion. The first problem is how a diagonal block or a block corresponding to a significant residual can be smoothed. The second problem is how smooth matrix U can be additionally compressed. For blocks which are far away from the diagonal the number

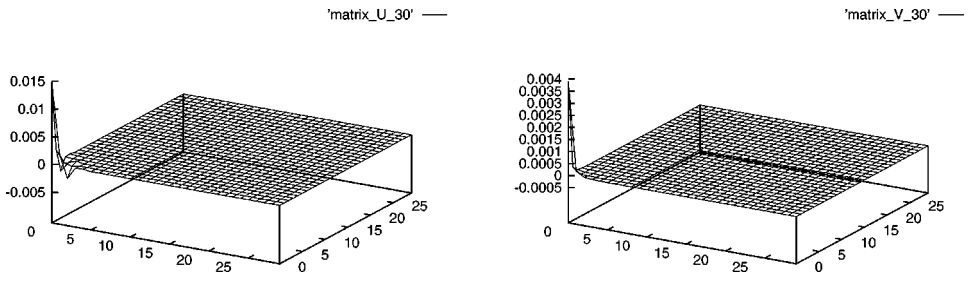


FIG. 5. The real part of the left-upper corners (30 points in each direction) of matrices U and V which are defined by the bell with steepness $k = 10$. The transform was applied to a block far away from the diagonal.

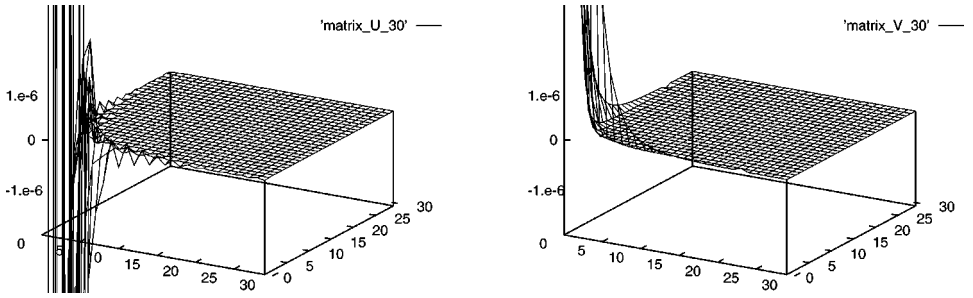


FIG. 6. The real part of the left-upper corners (30 points in each direction) of matrices U and V which are defined by bell with steepness $k = 10$ for the range of coefficients $[-10^{-6}, 10^{-6}]$. The transform was applied to a block far away from the diagonal.

of coefficients above 10^{-7} is about 55 (of 16,384) and matrix function V is very steep; therefore V cannot be additionally compressed by a wavelet or some other expansion.

4. MULTISCALE DECOMPOSITION WITH LOCAL FOURIER BASES

Multiscale LFB may transcribe a better compact description of an oscillatory kernel by utilization of the best-basis mechanism of the multiscale LFB. In this section purely real methods are developed for fast computation of oscillatory integrals. The effective representation is achieved by constructing an adaptive basis.

4.1. Adaptive Local Cosine for Kernel Compression

The smooth local trigonometric basis is used in the algorithm in the following way: Initially, we apply the local trigonometric transform (or basis). Then, the best-basis methodology [12] is applied on a tree of local trigonometric expansions.

4.1.1. Adaptive local trigonometric bases. We are interested in a local time-frequency analysis of the kernel. In order to analyze the local frequency content of the kernel, we first cut the support of the kernel into adjacent blocks. Then a local Fourier analysis is performed inside each block. To obtain a better frequency localization, we do not cut abruptly the signal, but we use a smooth window function to localize the segment of interest.

First we review the construction of one-dimensional smooth localized trigonometric bases [10, 2]. These functions are cosines with good localization in both position and Fourier space. We consider a cover $R = \bigcup_{n=-\infty}^{n=+\infty}[a_n, a_{n+1}[$. We write $l_n = a_{n+1} - a_n$ and $c_n = (a_n + a_{n+1})/2$. Around each a_n we define a neighborhood of radius ε. Let r be a ramp function such that

$$r(t) = \begin{cases} 0 & \text{if } t \le -1, \\ 1 & \text{if } t \ge 1 \end{cases} \tag{4.1}$$

and

$$r^2(t) + r^2(-t) = 1, \qquad \forall t \in R. \tag{4.2}$$

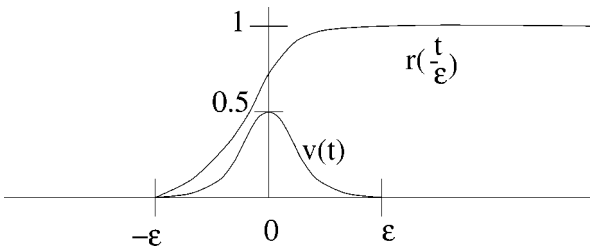


FIG. 7. Ramp function r, and bump function v.

We define the smooth orthogonal projection $P_{I_n} f$ of a one-dimensional signal f as follows [2, 10]:

$$P_{I_n} f(x) = \begin{cases} r^2(x-a_n) f(x) + r(x-a_n) r(a_n - x) \\ \qquad \text{if } a_n - \varepsilon \le x < a_n + \varepsilon f(2a_n - x) \\ f(x) \qquad \text{if } a_n + \varepsilon \le x < a_{n+1} - \varepsilon \\ r^2(a_{n+1} - x) f(x) - r(x - a_{n+1}) r(a_{n+1} - x) f(2a_{n+1} - x) \\ \qquad \text{if } a_{n+1} - \varepsilon \le x < a_{n+1} + \varepsilon. \end{cases} \tag{4.3}$$

We use the result that the projection of contiguous intervals is equal to the sum of the projections:

$$P_{I_n} \oplus P_{I_{n+1}} = P_{I_n \cup I_{n+1}}. \tag{4.4}$$

Consequently,

$$L^2(R) = \bigoplus_{n \in \mathbb{Z}} P_{I_n} L^2(R). \tag{4.5}$$

Furthermore, the projection is as smooth as the original function: if $f \in C^d(R)$, then $P_{I_n} f$ has a unique continuous extension in $C^d(R)$ that is supported in the interval $[a_n - \varepsilon, a_{n+1} + \varepsilon]$ [2] (see Scheme 1). We consider the collection of cosine functions:

$$c_{j,n} = \cos\left[\frac{2k+1}{2} \frac{\pi}{|I_n|} (x - a_n)\right].$$

In their discrete form these functions correspond to the DCT-IV:

$$X(k) = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} x(n) \cos\left[\frac{(2n+1)(2k+1)\pi}{4N}\right], \qquad k = 0, \ldots, N-1. \tag{4.6}$$

If we multiply the $c_{j,n}$ by the ramp, on each side of I_n, then we obtain a basis of $P_{I_n} L^2(R)$:

$$w_{k,n} = \sqrt{\frac{2}{|I_n|}} r(x - a_n) r(a_{n+1} - x) \cos\left[\frac{2k+1}{2} \frac{\pi}{|I_n|} (x - a_n)\right]. \tag{4.7}$$

To get better performance we use the bells that were developed by [15]. This bell generalizes the Coifman–Meyer (CM) LFB [10] by extending the choice of the bell functions. The orthonormality of the CM bells was dropped, as was done in [3].

This produces a family of smooth local trigonometric bases that efficiently compresses trigonometric functions. Any such basis is, in general, not orthogonal, but any element of the dual basis differs from the corresponding element of the original basis only by the shape of the bell. Numerical examples demonstrate that in many cases the proposed bases provide substantially better (up to a factor of two) compression than the standard CM bells. This is certainly true for seismic data [4]. We used the bell (3.5) from [15].

4.1.2. Implementation by folding. In practice, in order to expand a function f into the basis $w_{j,n}$ we do not calculate the correlation between f and the basis $\{w_{j,n}\}$. Instead we transform f restricted to $[a_n-\varepsilon, a_{n+1}+\varepsilon]$ into a smooth periodic function onto $[a_n, a_{n+1}]$ and expand it into the basis $\{c_{j,n}\}$. To do this we fold the overlapping parts of the window b_n and of the bump v back into the interval, across the endpoints of the interval, with some folding and unfolding operators. The advantage of the procedure is that we can preprocess the data with the folding operators and then use a fast DCT to calculate the expansion into the basis $\{c_{j,n}\}$.

UNITARY FOLDING AND UNFOLDING. We define the unitary folding operator U_{a_n} and its adjoint, the unfolding operator $U^*_{a_n}$, as follows:

$$U_{a_n} f(t) = \begin{cases} r(\frac{a_n - t}{\varepsilon}) f(t) - r(\frac{t-a_n}{\varepsilon}) f(2a_n - t), & \text{if } a_n - \varepsilon < t < a_n, \\ r(\frac{t-a_n}{\varepsilon}) f(t) + r(\frac{a_n - t}{\varepsilon}) f(2a_n - t), & \text{if } a_n < t < a_n + \varepsilon, \\ f(t), & \text{otherwise;} \end{cases} \tag{4.8}$$

$$U^*_{a_n} f(t) = \begin{cases} r(\frac{a_n - t}{\varepsilon}) f(t) + r(\frac{t-a_n}{\varepsilon}) f(2a_n - t), & \text{if } a_n - \varepsilon < t < a_n, \\ r(\frac{t-a_n}{\varepsilon}) f(t) - r(\frac{a_n - t}{\varepsilon}) f(2a_n - t), & \text{if } a_n < t < a_n + \varepsilon, \\ f(t), & \text{otherwise.} \end{cases} \tag{4.9}$$

Figures 8 and 9 show the result of the folding and unfolding operators.

The algorithm to expand a function f into the basis $w_{k,n}$ is thus:

- Apply the transform U_{a_n} to f.
- Apply $U_{a_{n+1}}$ to $U_{a_n} f$.
- Expand $U_{a_{n+1}} U_{a_n} f$, which is now living on $[a_n, a_{n+1}]$, into the DCT IV basis.

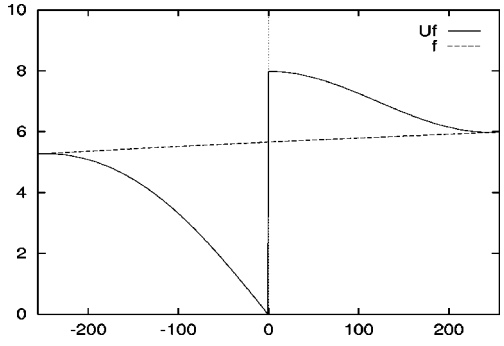


FIG. 8. Result of the folding operator.

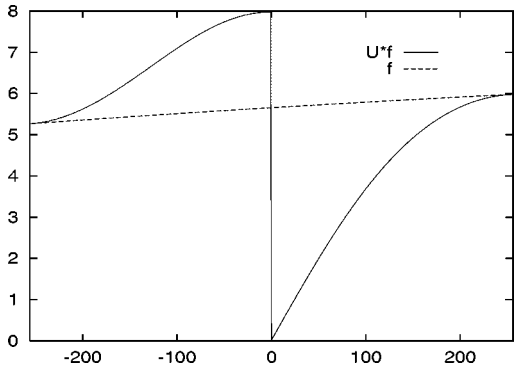


FIG. 9. Result of the unfolding operator.

The reconstruction algorithm that recovers the function f is:

– Apply an inverse DCT-IV to the coefficients. This gives $U_{a_{n+1}} U_{a_n} f$, which is living on $[a_n, a_{n+1}]$, the DCT IV basis.
– Apply the transform $U^*_{a_n}$.
– Apply $U^*_{a_{n+1}}$.

4.1.3. Adaptive segmentation. As explained in [2] we can adaptively select the size and location of the windows $[a_n, a_{n+1})$ with the best basis algorithm. We consider a function f with finite support. We divide the support into two intervals, and we consider the local cosine basis associated with each interval. We then further decompose each interval into two subintervals and consider the local cosine bases associated with this finer subdivision.

By applying this decomposition recursively we obtain a homogeneous binary tree-structured decomposition as shown in Fig. 10.

For each interval, or node in the binary tree, we calculate the set of coefficients in the subblock. If we associate a cost for each node of the tree, based on the set of coefficients, then we can find an optimal segmentation. Using a divide and conquer algorithm, groups of connected nodes are pruned if their total cost is greater than the cost of their father [12].

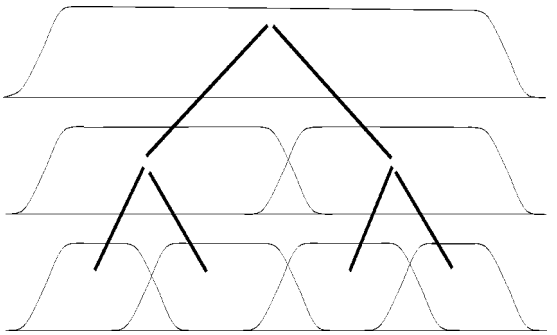


FIG. 10. One-dimensional binary tree decomposition. Within each interval the signal is expanded into a local DCT. The optimal segmentation is then searched for.

The process is recursively applied from the bottom to the top of the tree, and a global optimal tree is then found.

4.1.4. Two-dimensional case. In the two-dimensional case we define two partitions of R,

$$\bigcup_{n=-\infty}^{n=+\infty} [a_n, a_{n+1}[\qquad \text{and} \qquad \bigcup_{k=-\infty}^{k=+\infty} [b_m, b_{k+1}[.$$

We write $h_n = a_{n+1} - a_n$ and $l_m = b_{k+1} - b_m$. We then consider the tiling obtained by the lattice cubes $[a_n, a_{n+1}[\otimes[b_m, b_{k+1}[$. We consider the separable tensor products of bases $w_{m,j}$ and $w_{n,k}$. The sequence $w_{m,j} \otimes w_{n,k}$ is an orthonormal basis for $L^2(R^2)$.

4.1.5. Adaptive tiling of the image. As explained above we can adaptively select the size and location of the windows

$$[a_j, a_{j+1}[\otimes[b_m, b_{k+1}[$$

with the best basis algorithm. We consider only tilings that can be generated from separable bases. We divide the image into four subsquares, and we consider the local cosine basis associated with this tiling. We then further decompose each square into four subsquares and consider the local cosine bases associated with this finer tiling.

By applying this decomposition recursively we obtain a homogeneous quadtree-structured decomposition as shown in Fig. 11.

As in the one-dimensional case, for each subblock, or node of the quadtree, we calculate the set of coefficients in the subblock. We associate a cost for each node of the tree, based on the set of coefficients, and we find an optimal segmentation of the kernel.

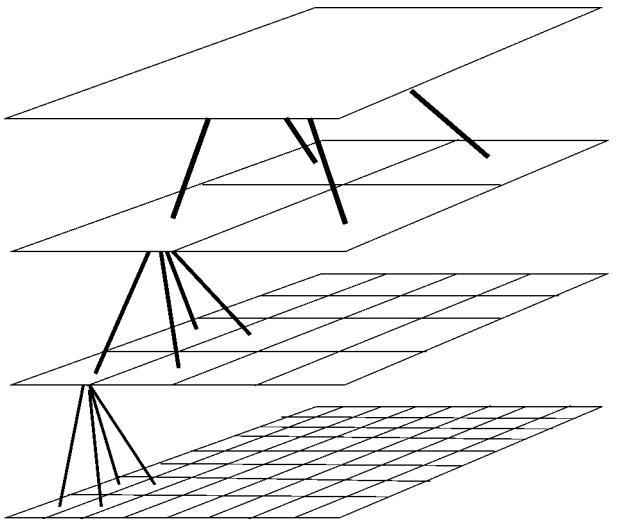


FIG. 11. Quadtree decomposition of the image. Within each block the image is expanded into a local DCT. The optimal tiling is then searched for.

TABLE 16
Parameters for the Construction of the Multiscale LFB

Number of points	Size of the window	Overlapping	Steepness of the bell	Number of levels
64×64	16	16	2	2
256×256	32	32	10	3
1024×1024	64	64	10	3
2048×2048	128	128	10	3
4096×4096	128	128	10	3

4.2. Numerical Results from the Application of the Multiscale Local Fourier Bases

The input parameters used are shown in Table 16.

Table 17 compares between the sparsity levels between the matrix kernel after using LFB (no multiresolution) and the matrix expanded in multiscale LFB (MLFB).

We observe the efficiency of the multiresolution approach. To illustrate the idea of the adaptive multiscale bases, we present the best basis grid chosen for 1024×1024 discretization points. The window is of size 32 with an overlapping of 16 points.

The chosen adaptive grid is symmetric with respect to the main diagonal. The highest resolution (the smallest boxes) correspond to the windows at the opposite sides of the same diameter, which have the strongest interaction (Fig. 12).

Table 18 illustrates the dependence of the number of coefficients above a threshold on the number of points N in each direction when ν in the exponent is constant.

TABLE 17
Comparison between the Sparsity Levels of the Oscillatory Kernel ($N = 8\nu$ is the Number of Grid Points in Each Direction) in Local Fourier Basis and in Multiresolution Local Fourier Basis with Best-Basis Methodology

Threshold ε	ν	N	Number of coefficients $> \varepsilon$ after LFB	Number of coefficients $> \varepsilon$ after MLFB	Number of coefficients per row after LFB	Number of coefficients per row after MLFB	Ratio LFB / MLFB
1e−3	8	64	2067	254	32.3	4.0	8.1
	32	256	7269	1706	28.4	6.7	4.3
	128	1024	17654	7736	17.2	7.6	2.3
1e−4	8	64	3381	906	52.8	14.2	3.7
	32	256	20304	3992	79.3	15.6	5.1
	128	1024	66307	21836	64.8	21.3	3.0
1e−5	8	64	3932	2658	61.4	41.5	1.5
	32	256	33947	10984	132.6	42.9	3.1
	128	1024	159943	48028	156.2	46.9	3.3
1e−6	8	64	4073	3954	63.6	61.8	1.0
	32	256	50217	17208	196.2	67.2	2.9
	128	1024	300998	78332	294	76.5	3.8

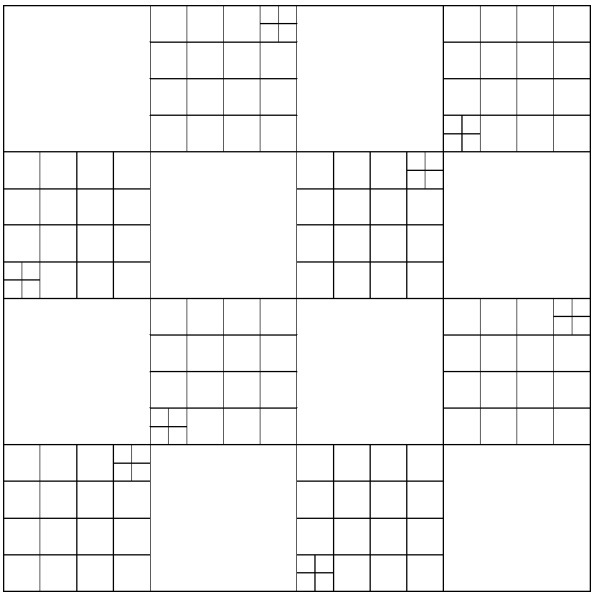


FIG. 12. Grid of the best-basis description of the multiscale LFB with 1024 points in each direction where the curve is a circle.

Thus, for $\nu = 128$ and $N = 4096$ there are less than 2% coefficients above 10^{-7}. The number of significant coefficients approximately doubles when the total number of grid points is multiplied by four (for small thresholds; for 10^{-3} to 10^{-4} the growth is slower).

Now consider the elliptic curve which corresponds to the following integral:

$$G(t,s) = \sin\left(\nu\sqrt{a^2(\cos(t)-\cos(s))^2 + b^2(\sin(t)-\sin(s))^2}\right), \qquad t,s \in [0,2\pi].$$

It is expected that due to the partial loss of symmetry the representation of this kernel in multiresolution local cosine bases will be less sparse than for the circle. Below are the results for various sizes and eccentricities of the ellipse. For discretization a simple trapezoidal quadrature was applied (without a special treatment for the discontinuity of the first derivative on the diagonal). The number of points in the kernel for the following

TABLE 18
Number of Coefficients above Thresholds 10^{-3} to 10^{-7} for Constant in the Exponential Function (2.1) $\nu = 128$ and Number of Grid Points in Each Direction $N \sim 512-4096$

		Bell width	Steepness		Number of coefficients above threshold				
N	ν	+ overlap	of bell	Levels	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
512	128	64 + 64	10	3	4044	10216	23708	34047	43359
1024	128	64 + 64	10	3	7736	21836	48028	78332	101840
2048	128	128 + 128	10	3	7732	22144	48364	95940	135780
4096	128	128 + 128	10	2	12669	47144	109792	198890	327938

TABLE 19
Sparsity of the Representation of the Kernel for Ellipse in Multiresolution LCB for 1024 Grid Points in Each Direction, $\nu = 128$, where Window Size is 64 Points with 32 Points Overlapping

Coefficients			Threshold ε				
a	b		10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
2.0	1.0	Total	12748	35068	102510	289222	529806
		%	1.2	3.4	10.0	28.2	51.7
0.8	0.6	Total	8774	24470	80340	251682	467298
		%	0.84	2.3	7.8	24.5	44.6
0.7	0.5	Total	7832	21740	72456	234868	440712
		%	0.7	2.1	7.0	22.4	42.0
0.3	0.1	Total	5856	15324	59542	196730	378236
		%	0.6	1.5	5.7	18.8	36.1
0.3	0.2	Total	4480	11261	50210	180244	361072
		%	0.4	1.1	4.8	17.2	34.4
0.1	0.05	Total	2768	7450	39150	134898	282958
		%	0.3	0.7	3.7	13.1	27.0
0.08	0.06	Total	2066	6372	23861	104115	249833
		%	0.2	0.6	2.3	9.9	23.8
0.02	0.01	Total	756	2768	11124	51256	145177
		%	0.07	0.3	1.0	5.0	13.8

example was 1024×1024, $\nu = 128$, the window size was 64 in both directions, overlapping was 32, steepness equal to 2, with four resolution levels.

5. SUMMARY AND DISCUSSION

The results of the present paper can be summarized as follows.

1. The kernel of the oscillatory integral (1.4) has a discontinuous derivative. For accurate computation of this integral the corrected trapezoidal rules developed in [20, 19] were applied. As a result an accuracy of 10^{-7} can be achieved for $N = 1024$, $\nu = 128$ (i.e., four grid points per oscillation).

2. For fast evaluation, the matrix, which was obtained after the discretization, is represented with local Fourier basis. The theoretical estimate of the number of coefficients above a threshold in the matrix representation is $O(N)$. The numerical results verify this estimate and demonstrate that LFB should be applied (rather than multiwavelets) to achieve a sparse representation of oscillatory kernels.

3. A variety of bells for LFB was considered. The numerical experiments show that the number of coefficients above a threshold in the LFB expansion decreases when the bell becomes smoother and steeper. Thus, for one of the steepest bells in the family [15], which has a smooth matching with 1 in the middle and zero in the sides, $N = 16384$, $\nu = 4096$ there are only about 0.18% of all the coefficients above 10^{-5}, 0.35% are above 10^{-6}, and 0.55% are above 10^{-7}. For bells which are not steep and smooth there are from 25 to 90% of coefficients above 10^{-7}.

4. In [6] it was illustrated that the LFB expansion of the oscillatory kernel is a smooth function. This phenomenon seems to be valid for bells which are not steep (their first derivatives are not large). For smooth and steep bells the LFB expansion is oscillatory. A smoothing procedure was developed for such oscillatory matrices.

5. Multiscale local Fourier bases provide adaptive representation of oscillatory kernels. Numerical examples demonstrate that for thresholds 10^{-4} to 10^{-6}, a number of grid points 1024×1024 and $\nu = 128$ the above expansion contains three times fewer points above a threshold than the LFB expansion.

Further work is in progress along the following lines:

1. The same method is applied via tensor product to 2-D and 3-D operators.

2. We investigate interactions among regions where the integral is defined. In the fast multipole method (FMM) the connections among the regions were applied to generate faster algorithms by compact coding of the coefficients. We strive to achieve effective coding of oscillatory integrals similar to FMM.

3. A more general class of bases [8, 16] with more parameters can be used.

REFERENCES

1. B. Alpert, G. Beylkin, R. Coifman, and V. Rokhlin, Wavelet-like bases for the fast solution of second-kind integral equations, *SIAM J. Sci. Comput.* **14** (1993), 159–184.
2. P. Auscher, G. Weiss, and M. V. Wickerhauser, Local sine and cosine bases of Coifman and Meyer and the construction of smooth wavelets, *in* "Wavelets—A Tutorial in Theory and Applications" (C. K. Chui, Ed.), pp. 237–256, Academic Press, New York, 1992.
3. A. Averbuch, G. Aharoni, R. Coifman, and M. Israeli, Local cosine transform—A method for the reduction of the blocking effect in JPEG, *J. Math. Imaging Vision*, Special Issue on Wavelets, **3** (1993), 7–38.
4. A. Averbuch, F. Meyer, O. Strömberg, R. Coifman, and A. Vassiliou, Low bit-rate efficient compression for seismic data, in preparation.
5. G. Beylkin, R. Coifman, and V. Rokhlin, Fast wavelet transforms and numerical algorithms, *Comm. Pure Appl. Math.* **44** (1991), 141–183.
6. B. Bradie, R. Coifman, and A. Grossman, Fast numerical computations of oscillatory integrals related to acoustic scattering, I, *Appl. Comput. Harmonic Anal.* **1** (1993), 94–99.
7. A. Cohen, I. Daubechies, and J. Feauveau, Biorthogonal bases of compactly supported wavelets, *Comm. Pure Appl. Math.* **45** (1996), 485–560.
8. R. Coifman and Y. Meyer, Gaussian bases, *Appl. Comput. Harmonic Anal.* **2** (1995), 299–302.
9. R. R. Coifman and Y. Meyer, Orthonormal wave packet bases, preprint, Dept. of Mathematics, Yale University, New Haven, 1990.
10. R. R. Coifman and Y. Meyer, Remarques sur l'analyse de Fourier à fenêtre, *C.R. Acad. Sci. Paris I* **312** (1991), 259–261.
11. R. Coifman, G. Matviyenko, and Y. Meyer, Modulated Malvar–Wilson bases, *Appl. Comput. Harmonic Anal.* **4** (1997), 58–61.
12. R. R. Coifman and M. V. Wickerhauser, Entropy-based algorithms for best basis selection, *IEEE Trans. Inform. Theory* **38** (1992), 713–718.
13. R. R. Coifman and Y. Meyer, Size properties of wavelet packets, *in* "Wavelets and Their Applications" (M. B. Ruskai, G. Beylkin, R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, and L. Rafael, Eds.), pp. 125–150, Jones & Bartlett, Boston, 1992.
14. I. Daubechies, Orthogonal bases of compactly supported wavelets, *Comm. Pure and Appl. Math.* **41** (1988), 909–996.
15. G. Matviyenko, Optimized local trigonometric bases, *Appl. Comput. Harmonic Anal.* **3** (1996), 301–323.

16. F. G. Meyer and R. R. Coifman, Brushlets: a tool for directional image analysis and image compression, *Appl. Comput. Harmonic Anal.* **4** (1997), 147–187.
17. A. Ralston and P. Rabinowitz, "A First Course in Numerical Analysis," Int. Series in Pure and Appl. Math., Vol. XIX, McGraw-Hill, KOGAKUSHA, Tokyo, 1978.
18. V. Rokhlin, Diagonal forms of translation operators for the Helmholtz equation in three dimensions, *Appl. Comput. Harmonic Anal.* **1** (1993), 82–93.
19. A. Sidi and M. Israeli, Quadrature methods for periodic singular and weakly singular Fredholm integral equations, *J. Sci. Comput.* **3** (1988), 201–231.
20. A. Sidi and J. A. Pennline, Improved accuracy of quadrature method solutions of Fredholm integral equations that arise from nonlinear two-point boundary value problems, *J. Integral Equations Appl.* **11**, 1 (1999), 103–139.
21. G. G. Stokes, On the numerical calculation of a class of definite integrals and infinite series, *Camb. Philos. Trans.* **9** (1856), 166–187.
22. G. N. Watson, The limits of applicability of the principle of stationary phase, *Proc. Camb. Philos. Soc.* **19** (1918), 49–55.