Contemporary Mathematics
Volume **330**, 2003

Trigonometric Spectral Collocation Methods on Lattices

Dong Li and Fred J. Hickernell

This paper is dedicated to Stan Osher on the occasion of his 60th birthday.

Abstract. Trigonometric spectral collocation methods can be very accurate for computing approximate solutions to periodic problems on rectangular domains. They obtain their answers by sampling the input function on a grid. This article shows how spectral methods can be extended to the situation where the input function is sampled on the nodeset of an integration lattice, a generalization of a grid. The error analysis is derived for a general approximation problem. Numerical examples illustrate how spectral methods on rank-1 lattices can give higher accuracy than spectral methods on grids.

1. Introduction

Given any function $f : \mathbb{R}^s \to \mathbb{R}$ with unit period interval, i.e.,

$$f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) \qquad \text{for all } \mathbf{x} \in \mathbb{R}^s,\ \mathbf{y} \in \mathbb{Z}^s,$$

one may consider the general linear problem of finding the periodic function u such that

$$u(\mathbf{x}) = (\mathcal{A}f)(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{R}^s, \tag{1}$$

where $\mathcal{A}$ is some operator. The dimension s is assumed to be arbitrary, perhaps large. Furthermore, it is assumed that $\mathcal{A}(f)$ can be easily expressed in terms of the trigonometric Fourier coefficients of f. Suppose that f and u may both be expressed by absolutely convergent Fourier series:

$$\begin{aligned} f(\mathbf{x}) &= \sum_{\mathbf{k}\in\mathbb{Z}^s} F(\mathbf{k}) e^{2\pi\imath\mathbf{k}\cdot\mathbf{x}}, \quad \text{where} \quad F(\mathbf{k}) = \int_{[0,1)^s} f(\mathbf{x}) e^{-2\pi\imath\mathbf{k}\cdot\mathbf{x}}\, d\mathbf{x}, \\ u(\mathbf{x}) &= \sum_{\mathbf{k}\in\mathbb{Z}^s} U(\mathbf{k}) e^{2\pi\imath\mathbf{k}\cdot\mathbf{x}}, \quad \text{where} \quad U(\mathbf{k}) = \int_{[0,1)^s} u(\mathbf{x}) e^{-2\pi\imath\mathbf{k}\cdot\mathbf{x}}\, d\mathbf{x}, \end{aligned} \tag{2}$$

1991 *Mathematics Subject Classification.* Primary 65N35, 65T40; Secondary 65D05, 65D32.

Key words and phrases. Approximation, Integration lattices, Partial differential equations, Spectral methods, Trigonometric polynomials.

This research was supported in part by Hong Kong Baptist University Faculty Research Grants FRG/99-00/I-45 and FRG/00-01/II-62.

and $\imath = \sqrt{-1}$. It is assumed that some function $A : \mathbb{Z}^s \to \mathbb{R}$ exists such that

$$U(\mathbf{k}) = A(\mathbf{k})F(\mathbf{k}). \tag{3}$$

Examples of problems of this form include

$$\text{Integration:} \qquad \mathcal{A}f = \int_{[0,1)^s} f(\mathbf{x})\, d\mathbf{x}, \qquad A(\mathbf{k}) = \delta_{\mathbf{k},\mathbf{0}},$$

$$\text{Function approximation:} \quad (\mathcal{A}f)(\mathbf{x}) = f(\mathbf{x}), \qquad A(\mathbf{k}) = 1,$$

$$\text{Poisson's equation:} \quad (\mathcal{A}f)(\mathbf{x}) = \Delta^{-1} f, \quad A(\mathbf{k}) = \begin{cases} -(4\pi^2 \|\mathbf{k}\|_2^2)^{-1}, & \mathbf{k} \neq \mathbf{0}, \\ 0, & \mathbf{k} = \mathbf{0}, \end{cases}$$

where $\delta_{i,j}$ denotes the Kronecker delta function and Δ is the Laplacian operator.

Because f may not be known in closed form, or may be somewhat complicated, it is desirable to be able to compute an approximation to u given the N data points $\{f(\mathbf{x}_i) : i = 0, \ldots, N-1\}$. Trigonometric spectral methods often provide highly accurate approximations [**1, 2, 10**]. In these cases, the sampling points $\mathcal{P} = \{\mathbf{x}_i : i = 0, \ldots, N-1\}$ are typically chosen to form an s-dimensional rectangular grid s aligned parallel to the coordinate axes. For simplicity, such sets will be referred to as grid samples in this article.

Choosing $\mathcal{P}$ to be the node set of a lattice, a generalization of a grid, yields accurate approximations to high dimensional integrals that are much better than those obtained by grid sampling [**5, 8, 9**] for certain classes of integrands. Several scholars have proposed sampling the input function on the nodeset of a lattice when approximating a function or solving a partial differential equation [**5, 6, 7**], but few details have been given regarding an efficient algorithm and its error analysis. This article attempts to fill that gap. It is shown that trigonometric spectral collocation methods may also be implemented for general linear problems of the form (1) if the sampling points are chosen as the nodeset of an integration lattice. Here too, nodesets of non-grid lattices may yield higher accuracy than that obtained by using grids for certain classes of input functions, f and certain classes of problems.

The next section gives a brief review of spectral methods using Poisson's equation as an example. Section 3 introduces lattice sampling. The error using spectral collocation methods on lattices is derived in Section 4, and an upper bound on the convergence rate is derived in Section 5. A numerical example demonstrating the superiority of spectral methods using lattices is given in Section 6. This article ends with a discussion of the relative merits of different kinds of lattices.

2. Trigonometric Spectral Methods

To illustrate trigonometric spectral methods, consider Poisson's partial differential equation:

$$\nabla^2 u(\mathbf{x}) = f(\mathbf{x}) \quad \text{for all } \mathbf{x} = (x_1, x_2) \in (0,1)^2, \tag{4a}$$

$$u(x_1, 0) = u(x_1, 1), \qquad u(0, x_2) = u(1, x_2), \tag{4b}$$

$$\left.\frac{\partial u}{\partial x_2}(x_1, x_2)\right|_{x_2=0} = \left.\frac{\partial u}{\partial x_2}(x_1, x_2)\right|_{x_2=1}, \tag{4c}$$

$$\left.\frac{\partial u}{\partial x_1}(x_1, x_2)\right|_{x_1=0} = \left.\frac{\partial u}{\partial x_1}(x_1, x_2)\right|_{x_1=1}, \tag{4d}$$

where f, not only is periodic, but also has zero integral over the unit square. The analytical solution to this partial differential equation is easily found via the method of Fourier expansion. The solution is unique up to an additive constant, which is determined by requiring the integral of u over the unit square to vanish. Then

$$u(\mathbf{x}) = (\mathcal{A}f)(\mathbf{x}) = \Delta^{-1}f(\mathbf{x}) = \sum_{0\neq\mathbf{k}\in\mathbb{Z}^2} -\frac{F(\mathbf{k})}{4\pi^2(k_1^2+k_2^2)}e^{2\pi\imath\mathbf{k}\cdot\mathbf{x}}. \tag{5}$$

The basic idea of trigonometric spectral methods is to represent the approximate solution as an expansion in terms of a suitably chosen basis of trigonometric polynomials. Define the following grid of $N = n_1 n_2$ points,

$$\mathcal{P} := \{(i_1/n_1, i_2/n_2) : i_1 = 0, 1, \cdots, n_1 - 1, i_2 = 0, 1, \cdots, n_2 - 1\},$$

as shown in Figure 1 and the associated grid of wavenumbers,

$$\begin{aligned}(6)\quad \mathcal{K} := \{(k_1, k_2) : k_1 = 1 - \lceil n_1/2\rceil, \ldots, n_1 - \lceil n_1/2\rceil,\\ k_2 = 1 - \lceil n_2/2\rceil, \ldots, n_2 - \lceil n_2/2\rceil\},\end{aligned}$$

where $\lceil x\rceil$ denotes the smallest integer greater than or equal to x. One approximates f by a truncated Fourier series,

$$\hat{f}(\mathbf{x}) := \sum_{\mathbf{k}\in\mathcal{K}} \hat{F}(\mathbf{k})e^{2\pi\imath\mathbf{k}\cdot\mathbf{x}}.$$

The interpolation conditions, $\hat{f}(\mathbf{z}) = f(\mathbf{z})$ for all $\mathbf{z} \in \mathcal{P}$, determine the approximations, $\hat{F}(\mathbf{k})$, to the true Fourier coefficients. Noting that

$$\frac{1}{N}\sum_{\mathbf{z}\in\mathcal{P}} e^{2\pi\imath(\mathbf{l}-\mathbf{k})\cdot\mathbf{z}} = \delta_{\mathbf{k},\mathbf{l}} \quad \text{for all } \mathbf{k}, \mathbf{l} \in \mathcal{K}, \tag{7}$$

it follows that

$$\begin{aligned}(8)\qquad \hat{F}(\mathbf{k}) &= \frac{1}{N}\sum_{\mathbf{l}\in\mathcal{K}}\sum_{\mathbf{z}\in\mathcal{P}} \hat{F}(\mathbf{l})e^{2\pi\imath(\mathbf{l}-\mathbf{k})\cdot\mathbf{z}} = \frac{1}{N}\sum_{\mathbf{z}\in\mathcal{P}} \hat{f}(\mathbf{z})e^{-2\pi\imath\mathbf{k}\cdot\mathbf{z}}\\ &= \frac{1}{N}\sum_{\mathbf{z}\in\mathcal{P}} f(\mathbf{z})e^{-2\pi\imath\mathbf{k}\cdot\mathbf{z}} \quad \text{for all } \mathbf{k} \in \mathcal{K}.\end{aligned}$$

The above sum can be evaluated efficiently by the Fast Fourier Transform (FFT) for suitable N, e.g. $N = 2^m$. The approximation to $u(\mathbf{x})$ is found by replacing $F(\mathbf{k})$ in (5) by $\hat{F}(\mathbf{k})$:

$$\hat{u}(\mathbf{x}) = \sum_{0\neq\mathbf{k}\in\mathcal{K}} -\frac{\hat{F}(\mathbf{k})}{4\pi^2(k_1^2+k_2^2)}e^{2\pi\imath\mathbf{k}\cdot\mathbf{x}}.$$

The error in this approximation arises from the error in $\hat{F}(\mathbf{k})$. From (8) it follows that

$$\begin{aligned}(9)\qquad \hat{F}(\mathbf{k}) &= \frac{1}{N}\sum_{\mathbf{z}\in\mathcal{P}} f(\mathbf{z})e^{-2\pi\imath\mathbf{k}\cdot\mathbf{z}} = \frac{1}{N}\sum_{\mathbf{l}\in\mathbb{Z}^2}\sum_{\mathbf{z}\in\mathcal{P}} F(\mathbf{l})e^{2\pi\imath(\mathbf{l}-\mathbf{k})\cdot\mathbf{z}}\\ &= \sum_{\mathbf{l}\in L^\perp} F(\mathbf{k}+\mathbf{l}) = F(\mathbf{k}) + \sum_{0\neq\mathbf{l}\in L^\perp} F(\mathbf{k}+\mathbf{l}),\end{aligned}$$

where $L^\perp = \{(m_1n_1, m_2n_2) : (m_1, m_2) \in \mathbb{Z}^2\}$. That is, $\hat{F}(\mathbf{k})$ is contaminated by terms $F(\mathbf{k}+\mathbf{l})$ with $\mathbf{0} \neq \mathbf{l} \in L^\perp$. If these terms are small because $F(\mathbf{l})$ decays quickly with increasing $\mathbf{l}$, then the error is small.

The above procedure is not restricted to solving this particular partial differential equation. Any linear equation that can be represented in Fourier space as (3) can be solved in the same way. Moreover, the sets $\mathcal{P}$ and $\mathcal{K}$ need not be grids. The key property they need to satisfy is (7). Also, there should be an efficient method for evaluating the sum in (8) giving the $\hat{F}(\mathbf{k})$ from the data. Integration lattices are generalizations of grids that have these desired properties.

3. Integration Lattices

Integration lattices [**3, 5, 8, 9**] have been used for over forty years for multidimensional integration, and they are still an area of active research. An s-dimensional integration lattice, L, is a set satisfying the following:

$$\mathbb{Z}^s \subseteq L \subset \mathbb{R}^s, \qquad \mathbf{y}, \mathbf{z} \in L \text{ implies } \mathbf{y} \pm \mathbf{z} \in L.$$

Furthermore, the nodeset of a lattice, $\mathcal{P} := L \cap [0,1)^s$ is assumed to be finite. Shifted lattices are also of interest. A shifted lattice is defined as $L_{\mathbf{\Delta}} := \{\mathbf{z} + \mathbf{\Delta} : \mathbf{z} \in L\}$, where the shift $\mathbf{\Delta} \in [0,1)^s$ is often chosen randomly to eliminate the bias in the algorithm.

A rectangular grid is one example of a lattice, but there are many other possibilities. Examples of three two-dimensional lattices and their nodesets are given below:

$$L = \{(i_1/4, i_2/4) : i_1, i_2 \in \mathbb{Z}\}, \quad \mathcal{P} = \{(i_1/4, i_2/4) : i_1, i_2 = 0, 1, 2, 3\}, \tag{10}$$

$$L = \{(1,7)i_1/16 + (0,1)i_2 : i_1, i_2 \in \mathbb{Z}\}, \tag{11a}$$

$$\mathcal{P} = \{(1,7)i_1/16 \mod 1 : i_1 = 0, \dots, 15\}, \tag{11b}$$

$$L = \{(1,3)i_1/8 + (0,1)i_2/2 : i_1, i_2 \in \mathbb{Z}\}, \tag{12a}$$

$$\mathcal{P} = \{(1,3)i_1/8 + (0,1)i_2/2 \mod 1 : i_1 = 0, \dots, 7,\ i_2 = 0, 1\}, \tag{12b}$$

These three nodesets are plotted in Figures 1–3. The lattices in (10) and (12a) are rank-2 lattices because they require two generating vectors, whereas the lattice in (11a) is a rank-1 lattice. For details on determining the rank of a lattice see [**9**].

General integration lattices share several important properties with grids that make them amenable to spectral methods. Suppose that the nodeset of an unshifted lattice, $\mathcal{P} = L \cap [0,1)^s$, has cardinality N. The dual lattice, $L^\perp$ is defined as the set of $\mathbf{k} \in \mathbb{Z}^s$ satisfying

$$\frac{1}{N} \sum_{\mathbf{z} \in \mathcal{P}} e^{2\pi \imath \mathbf{k} \cdot \mathbf{z}} = 1.$$

It then follows that for the nodeset of a shifted lattice, $\mathcal{P} = L_{\mathbf{\Delta}} \cap [0,1)^s$,

$$\frac{1}{N} \sum_{\mathbf{z} \in \mathcal{P}} e^{2\pi \imath \mathbf{k} \cdot \mathbf{z}} = \begin{cases} e^{2\pi \imath \mathbf{k} \cdot \mathbf{\Delta}}, & \text{for } \mathbf{k} \in L^\perp, \\ 0, & \text{for } \mathbf{k} \in \mathbb{Z}^s,\ \mathbf{k} \notin L^\perp. \end{cases} \tag{13}$$

The set of all integer wavenumber vectors, $\mathbb{Z}^s$ may be written as $\mathcal{K} \oplus L^\perp$, where $\oplus$ denotes the direct sum, and $\mathcal{K}$ is a set of N integer vectors having the property

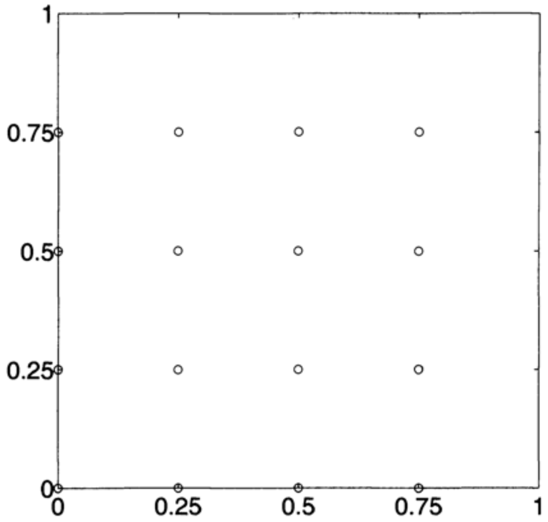

FIGURE 1. The nodeset (10).

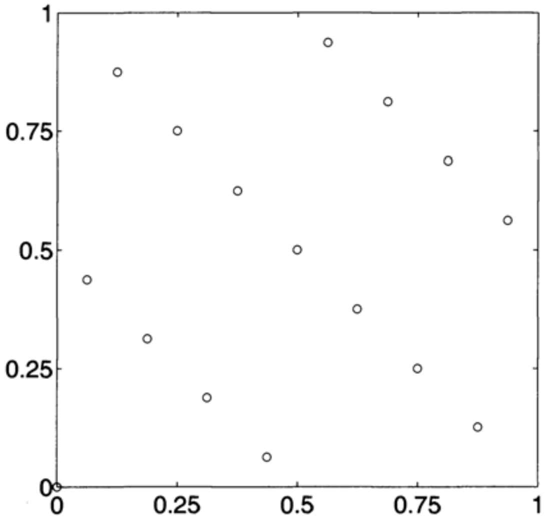

FIGURE 2. The nodeset (11b).

that any two distinct vectors in $\mathcal{K}$ differ by a vector that is not in the dual lattice, i.e.,

$$\mathbf{k}, \mathbf{l} \in \mathcal{K} \text{ and } \mathbf{k} \neq \mathbf{l} \quad \Longrightarrow \quad \mathbf{k} - \mathbf{l} \notin L^{\perp}. \tag{14}$$

There are many ways to choose $\mathcal{K}$ such that $\mathbb{Z}^s = \mathcal{K} \oplus L^{\perp}$. In (6) the vectors in $\mathcal{K}$ were chosen to be the smallest possible in terms of the Euclidean norm. For general lattices the condition to uniquely determine $\mathcal{K}$ from L, or its nodeset $\mathcal{P}$, also assumes that the $\mathbf{k}$ in $\mathcal{K}$ are the smallest in some sense, but does not necessarily use the Euclidean norm. Details are provided in the following sections.

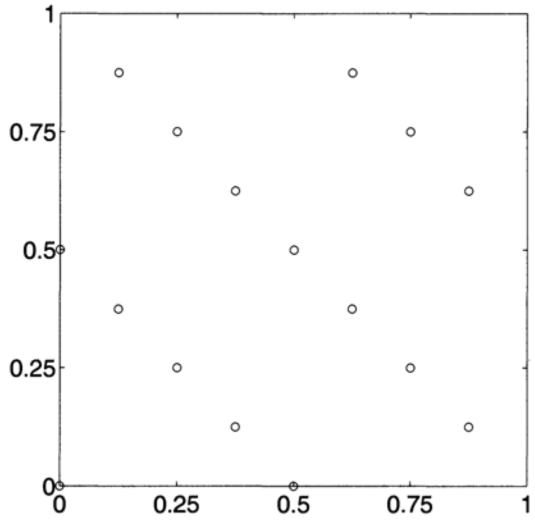

FIGURE 3. The nodeset (12b).

Because of property (13) it is possible to perform spectral methods using nodesets of general shifted integration lattices (not only grids) by following the procedure in Section 2. The approximate solution to the general problem (1) may be written as

$$\hat{u}(\mathbf{x}) := \sum_{\mathbf{k}\in\mathcal{K}} A(\mathbf{k})\hat{F}(\mathbf{k})e^{2\pi\imath\mathbf{k}\cdot\mathbf{x}}, \tag{15a}$$

where $\mathcal{K}$ is some wavenumber set satisfying (14) and $\hat{F}(\mathbf{k})$ is the approximation to $F(\mathbf{k})$ given by a lattice rule:

$$\hat{F}(\mathbf{k}) := \frac{1}{N}\sum_{\mathbf{z}\in\mathcal{P}} f(\mathbf{z})e^{-2\pi\imath\mathbf{k}\cdot\mathbf{z}} \quad \text{for all } \mathbf{k}\in\mathcal{K}. \tag{15b}$$

Analogous to (9), it is true for general shifted lattices that $\hat{F}(\mathbf{k})$ is contaminated by terms proportional to $F(\mathbf{k}+\mathbf{l})$ for $\mathbf{l}\in L^{\perp}$:

$$\hat{F}(\mathbf{k}) = F(\mathbf{k}) + \sum_{\mathbf{0}\neq\mathbf{l}\in L^{\perp}} F(\mathbf{k}+\mathbf{l})e^{2\pi\imath\mathbf{l}\cdot\mathbf{\Delta}} \quad \text{for all } \mathbf{k}\in\mathcal{K}. \tag{16}$$

The quality of the approximation depends on how fast $\hat{F}(\mathbf{l})$ decays as $\mathbf{l}$ increases in size, and also on the quality of $\mathcal{P}$ and $\mathcal{K}$. To understand this dependence the worst-case error is now analyzed.

4. Worst-Case Error Analysis

Before deriving the worst-case error for the general linear problem (1), the known error analysis for integration is reviewed [**3, 8, 9**]. For an integrand of the form (2) and a numerical integration rule that takes the mean of the integrand sampled at the nodeset of a shifted lattice, the error is

$$\int_{[0,1)^s} f(\mathbf{x})\,d\mathbf{x} - \frac{1}{N}\sum_{\mathbf{z}\in\mathcal{P}} f(\mathbf{z}) = -\sum_{\mathbf{0}\notin\mathbf{k}\in L^{\perp}} F(\mathbf{k})e^{2\pi\imath\mathbf{k}\cdot\mathbf{\Delta}}.$$

This formula is derived by noting that the integral of $e^{2\pi\imath\mathbf{k}\cdot\mathbf{x}}$ vanishes for all nonzero wavenumbers, and the numerical integration rule integrates $e^{2\pi\imath\mathbf{k}\cdot\mathbf{x}}$ correctly for all wavenumbers not in the dual lattice. For wavenumbers in the dual lattice $e^{2\pi\imath\mathbf{k}\cdot\mathbf{x}}$ is aliased with the constant term, which is integrated exactly.

To separate the error dependence on the integrand from the dependence on the lattice, Hölder's inequality is applied to the formula above. First one defines the semi-norm

$$V_{\alpha,p}(f) := \left\|F(\mathbf{k})r(\mathbf{k})^{\alpha} \cdot 1_{\mathbf{k}\neq\mathbf{0}}\right\|_p, \quad \alpha > 0,\ 1 \leq p \leq \infty, \tag{17}$$

where

$$r(\mathbf{k}) = r(k_1, \cdots, k_s) = \prod_{j=1}^{s} \max(|k_j|, 1).$$

The exponent α measures how fast the Fourier coefficients of f are expected to decay. Now, Hölder's inequality implies that

$$\left|\int_{[0,1)^s} f(\mathbf{x})\, d\mathbf{x} - \frac{1}{N}\sum_{\mathbf{z}\in\mathcal{P}} f(\mathbf{z})\right| \leq [D_{\alpha q}(\mathcal{P})]^{\alpha} V_{\alpha,p}(f), \quad \frac{1}{p} + \frac{1}{q} = 1,\ \alpha q > 1,$$

where

$$D_q(\mathcal{P}) := \left\|\frac{1_{\mathbf{0}\neq\mathbf{k}\in L^{\perp}}}{r(\mathbf{k})}\right\|_q, \quad 1 < q \leq \infty. \tag{18}$$

Furthermore, this error bound is tight. The quantity $D_q(\mathcal{P})$ may be called a discrepancy because it may be interpreted as some norm of the difference between the uniform distribution over $[0,1)^s$ and the empirical distribution of $\mathcal{P}$. In this way it is related to the discrepancy of Weyl [**11**]. Upper bounds on $D_q(\mathcal{P})$ that may be achieved by good choices of lattices are given by [**4, 5, 8, 9**].

One may wonder about the choice of $r(\mathbf{k})$ that arises in the definition of the semi-norm $V_{\alpha,p}$. The subsets of $\mathbb{Z}^2$ defined by $r(\mathbf{k}) \leq c$ take the shape of hyperbolic crosses. The reason for such a definition is that for multidimensional problems one might expect f to vary strongly in only one or two dimensions. Defining $V_{\alpha,p}$ as above, and then choosing $\mathcal{P}$ with small $D_{\alpha q}(\mathcal{P})$ ensures that $\mathcal{P}$ samples f at many different values in each coordinate direction. This is why nodesets of the form (11b) tend to be better for numerical integration than grids with the same numbers of points for this class of periodic integrands.

To analyze the error for the general linear problem let ω denote a nonnegative weight function on $\mathbb{Z}^s$. The error of the approximate solution $\hat{u}$ may be measured by the semi-norm

$$\|u - \hat{u}\|_{\omega} := \sum_{\mathbf{k}\in\mathbb{Z}^s} |\omega(\mathbf{k})(u(\mathbf{k}) - \hat{u}(\mathbf{k}))| = \sum_{\mathbf{k}\in\mathbb{Z}^s} \left|\omega(\mathbf{k})A(\mathbf{k})(F(\mathbf{k}) - \hat{F}(\mathbf{k}))\right|.$$

For each integer vector $\mathbf{k}$ in $\mathbb{Z}^s$, let $\mathbf{k}/\mathcal{K}$ denote the unique vector in $\mathcal{K}$ such that $\mathbf{k} = \mathbf{k}/\mathcal{K} + \mathbf{l}$ for some $\mathbf{l} \in L^{\perp}$ (see (14)). It is assumed that $\mathbf{0} \in \mathcal{K}$. By the aliasing property (16) it follows that the error is bounded by

$$\begin{aligned} \|u - \hat{u}\|_{\omega} &\leq \sum_{\mathbf{k}\in\mathcal{K}} \sum_{\mathbf{0}\neq\mathbf{l}\in L^{\perp}} \left|F(\mathbf{k}+\mathbf{l})e^{2\pi\imath\mathbf{l}\cdot\boldsymbol{\Delta}}\omega(\mathbf{k})A(\mathbf{k})\right| + \sum_{\mathbf{k}\notin\mathcal{K}} |F(\mathbf{k})\omega(\mathbf{k})A(\mathbf{k})| \\ &\leq \sum_{\mathbf{k}\notin\mathcal{K}} |F(\mathbf{k})| \left\{|\omega(\mathbf{k}/\mathcal{K})A(\mathbf{k}/\mathcal{K})| + |\omega(\mathbf{k})A(\mathbf{k})|\right\}. \end{aligned} \tag{19}$$

Applying Hölder's inequality yields the following upper bound on the error.

THEOREM 1. *Consider problems of the form* (1) *and spectral method approximations of the form* (15). *The error is bounded by*

$$\|u-\hat{u}\|_{\omega} \le D_{\alpha,q}(\mathcal{P};\mathcal{K},\mathcal{A})V_{\alpha,p}(f), \quad \frac{1}{p}+\frac{1}{q}=1,\ 1\le p,q\le\infty.$$

where the semi-norm of f *was defined in* (17) *and*

$$D_{\alpha,q}(\mathcal{P};\mathcal{K},\mathcal{A}) := \left\| \frac{|\omega(\mathbf{k}/\mathcal{K})A(\mathbf{k}/\mathcal{K})| + |\omega(\mathbf{k})A(\mathbf{k})|}{r(\mathbf{k})^{\alpha}} \cdot 1_{\mathbf{k}\notin\mathcal{K}} \right\|_q . \tag{20}$$

The quantity defined in (20) depends on the node set of the lattice, $\mathcal{P}$, the wavenumber set, $\mathcal{K}$, and the norm on the space of input functions. However, $D_{\alpha,q}(\mathcal{P};\mathcal{K},\mathcal{A})$ does not depend on the particular function f. Good $\mathcal{P}$ and $\mathcal{K}$ are those that make $D_{\alpha,q}(\mathcal{P};\mathcal{K},\mathcal{A})$ as small as possible.

One drawback of $D_{\alpha,q}(\mathcal{P};\mathcal{K},\mathcal{A})$ is that it does depend on the particular kind of problem being solved, $\mathcal{A}$. To remove this dependence, suppose that there exist constants $C>0$ and β such that

$$|\omega(\mathbf{k})A(\mathbf{k})| \le C[r(\mathbf{k})]^{\beta}. \tag{21}$$

In the Poisson's equation example of Section 2 $A(\mathbf{k}) = 1/[4\pi^2(k_1^2+k_2^2)]$ for $\mathbf{k}\neq\mathbf{0}$. If $\omega(\mathbf{k}) = A(\mathbf{k})^{-t}$ for some number t and for $\mathbf{k}\neq\mathbf{0}$, and $\omega(\mathbf{0})=1$, then this inequality takes the form

$$|\omega(\mathbf{k})A(\mathbf{k})| = [4\pi^2(k_1^2+k_2^2)]^{t-1} \le \begin{cases} [4\pi^2 r(\mathbf{k})]^{t-1}, & t<1,\\ 1, & t=1,\\ [2\sqrt{2}\pi r(\mathbf{k})]^{2(t-1)}. & t>1 \end{cases}$$

This assumption is discussed further at the end of this article.

Furthermore, assume that $\mathcal{K}$ is chosen to satisfy the condition

$$r(\mathbf{k}) \ge r(\mathbf{k}/\mathcal{K}) \quad \text{for all } \mathbf{k}\in\mathbb{Z}^s. \tag{22}$$

This may be done by ranking wavenumbers in terms of their $r(\mathbf{k})$ values, and then choosing the first N that have distinct values of $\mathbf{k}/\mathcal{K}$. Then $D_{\alpha,q}(\mathcal{P};\mathcal{K},\mathcal{A}) \le 2C[\tilde{D}_{(\alpha-\beta)q}(\mathcal{P})]^{\alpha-\beta}$, where

$$\tilde{D}_q(\mathcal{P}) := \left\| \frac{1_{\mathbf{k}\notin\mathcal{K}}}{r(\mathbf{k})} \right\|_q = \begin{cases} \left\{ [1+2\zeta(q)]^s - \left\| \frac{1_{\mathbf{k}\in\mathcal{K}}}{r(\mathbf{k})} \right\|_q^q \right\}^{1/q}, & 1<q<\infty,\\ \sup_{\mathbf{k}\notin\mathcal{K}} r(\mathbf{k}), & q=\infty, \end{cases} \tag{23}$$

and ζ is the Riemann zeta function. The last expression above is more suited to numerical evaluation since it involves looking at only a finite number of terms.

COROLLARY 2. *Under the same assumptions as in Theorem 1 the worst-case error has the somewhat looser error bound of*

$$\|u-\hat{u}\|_{\omega} \le 2C[\tilde{D}_{(\alpha-\beta)q}(\mathcal{P})]^{\alpha-\beta}V_{\alpha,p}(f), \quad \frac{1}{p}+\frac{1}{q}=1,\ (\alpha-\beta)q>1,$$

where C *and* β *are absolute constants arising in* (21) *that depend on the problem being solved.*

5. Asymptotic Bounds on the Worst-Case Error for Large N

The quantity $\tilde{D}_q(\mathcal{P})$ defined in (23) is similar in form to $D_q(\mathcal{P})$ defined in (18). There have been many studies of the best obtainable convergence rate for the discrepancy or integration error for lattice rules, $D_q(\mathcal{P})$, but none yet for $\tilde{D}_q(\mathcal{P})$. The following theorem relates these two.

THEOREM 3. *For any $\mathcal{P}$ that is the node set of a shifted lattice,*

$$\tilde{D}_q(\mathcal{P}) \le 2^{s/2} N^{1/q} [D_{q/2}(\mathcal{P})]^{1/2} \quad \textit{for all } q > 2.$$

The key fact needed to prove this theorem is the following lemma.

LEMMA 4. *For any integer wave number vectors* $\mathbf{k}$ *and* $\mathbf{l}$ *it follows that*

$$\frac{r(\mathbf{l})}{r(\mathbf{k}+\mathbf{l})} \le 2^s r(\mathbf{k}). \tag{24}$$

PROOF. The proof is given for the case $s = 1$. The case of general s follows immediately. If $(k+l)kl = 0$, then (24) holds automatically, so, suppose that $(k+l)kl \ne 0$. Under this condition $r(k) = |k|$, $r(l) = |l|$, and $r(k+l) = |k+l| \ge 1$. Therefore,

$$\frac{r(l)}{r(k+l)} = \frac{|l|}{|k+l|} \le \frac{|k+l|+|k|}{|k+l|} \le 1 + |k| \le 2r(k),$$

establishing the lemma. Note that equality holds above when $k = 1$ and $l = -2$. □

PROOF OF THEOREM 3. By (23) it follows that

$$\tilde{D}_q(\mathcal{P}) = \left\| \frac{1_{\mathbf{k}\notin\mathcal{K}}}{r(\mathbf{k})} \right\|_q = \left[\sum_{\mathbf{k}\notin\mathcal{K}} \frac{1}{r(\mathbf{k})^q} \right]^{1/q} = \left[\sum_{\mathbf{k}\in\mathcal{K}} \sum_{\mathbf{0}\ne\mathbf{l}\in L^\perp} \frac{1}{r(\mathbf{k}+\mathbf{l})^q} \right]^{1/q}.$$

For any $\mathbf{k} \in \mathcal{K}$ and $\mathbf{l} \in L^\perp$ the lemma above and condition (22) imply that

$$\frac{1}{r(\mathbf{k}+\mathbf{l})} \le \frac{2^s r(\mathbf{k})}{r(\mathbf{l})} \le \frac{2^s r(\mathbf{k}+\mathbf{l})}{r(\mathbf{l})}, \qquad \text{so} \qquad \frac{1}{r(\mathbf{k}+\mathbf{l})} \le \left[\frac{2^s}{r(\mathbf{l})}\right]^{1/2}.$$

Thus,

$$\tilde{D}_q(\mathcal{P}) \le \left[\sum_{\mathbf{k}\in\mathcal{K}} \sum_{\mathbf{0}\ne\mathbf{l}\in L^\perp} \left\{ \frac{2^s}{r(\mathbf{l})} \right\}^{q/2} \right]^{1/q} = N^{1/q} 2^{s/2} [D_{q/2}(\mathcal{P})]^{1/2}.$$

□

It is known by [**4**] and the references therein that there exist extensible integration lattices with $D_q(\mathcal{P}) = \mathcal{O}(N^{-1+\epsilon})$ for any $\epsilon > 0$. Theorem 3 then immediately leads to the following corollary.

COROLLARY 5. *For $2 < q \le \infty$ and any $\epsilon > 0$ there exist extensible integration lattices with $\tilde{D}_q(\mathcal{P}) = \mathcal{O}(N^{-1/2+1/q+\epsilon})$.*

Referring to Corollary 2 it follows that the worst-case error for problems of the form (1) is $\|u - \hat{u}\|_\omega = \mathcal{O}(N^{(\alpha-\beta)(-1/2+1/q)+\epsilon})$. The parameter α indicates the smoothness of f. The larger α is, the higher the convergence rate. If all derivatives of f exist, then α can be made arbitrarily large. The parameter β depends on the particular problem and the function ω used to define the norm of the true solution minus its approximation. Notice that for the class of functions considered the convergence rate using good lattices is basically independent of the dimension,

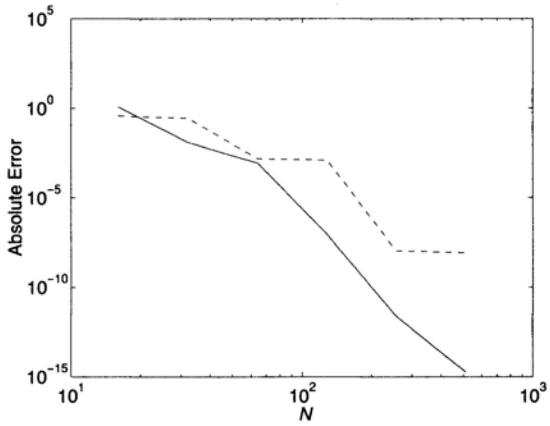

FIGURE 4. A comparison of errors using spectral methods to solve (25) on a rank-1 lattice (solid) and on a grid (dashed).

s. In contrast, using grids one has $\tilde{D}_q(\mathcal{P}) = \mathcal{O}(N^{1/s})$, since one can sample at only $N^{1/s}$ different values in each coordinate direction. For large s this may lead to very slow convergence of the spectral method.

6. Numerical Example

To illustrate the advantages of using general integration lattices the example of the two-dimensional Poisson's equation in (4) is revisited with

$$\Delta u = f(\mathbf{x}) = -4\pi^2 \sin(2\pi x_1) \left\{ e^{\sin(2\pi x_1)} \sin(2\pi x_2) \left[1 + \sin(2\pi x_1)\right] + e^{\cos(2\pi x_2)} \cos(2\pi x_2) \left[1 + \cos(2\pi x_2)\right] \right\}, \tag{25a}$$

and the solution

$$u(\mathbf{x}) = e^{\sin(2\pi x_1)} \sin(2\pi x_2) + e^{\cos(2\pi x_2)} \sin(2\pi x_1). \tag{25b}$$

For $N = 16, 32, \ldots, 512$, the solution is computed using spectral methods where $\mathcal{P}$ is an $n_1 \times n_2$ grid with $n_1 := 2^{\lceil (\log_2 N)/2 \rceil} \geq n_2$. Spectral method approximations are also computed using the nodeset of a rank-1 lattice,

$$\mathcal{P} = \{(1, a)i/N \mod 1 : i = 0, \ldots, N\}, \tag{26}$$

with

a	7	7	19	47	75	149
N	16	32	64	128	256	512

These values of a are those for which $\mathcal{P}$ of the form (26) minimize $\tilde{D}_\infty(\mathcal{P})$. The error of the approximate solution is computed as $\sup_{\mathbf{z}\in\mathcal{P}} |u(\mathbf{z}) - \hat{u}(\mathbf{z})|$ and plotted in Figure 4. For this example the error using the rank-1 lattice is substantially less than that using the grid for some values of N.

7. Discussion

The quality of the algorithm described here depends on the set of wavenumbers $\mathcal{K}$, which should include those wavenumbers which are more important. As seen in (19) this depends upon three factors: i) which Fourier coefficients of the input function, $F(\mathbf{k})$, are large, ii) the particular problem as defined through $A(\mathbf{k})$, and iii) the norm defining the error as given through $\omega(\mathbf{k})$. If all these factors are all well understood, then one can design the integration lattice and its nodeset, grid or not grid, to fit the problem. An interesting topic for further research is to find a correspondence between the rank of a lattice (1 through s) and the possible shapes of the set $\mathcal{K}$.

If a priori knowledge is difficult to obtain, e.g., the input function is quite complicated, then it is better to find a lattice nodeset and corresponding $\mathcal{K}$ that work for a wide variety of problems. This is the spirit of work presented here and the reason behind several of the choices made above. The theory above has been directed to choose sets $\mathcal{K}$ with $r(\mathbf{k})$, specifically to minimize a figure of merit such as $\tilde{D}_\infty(\mathcal{P})$. An intuitive rationale for minimizing $r(\mathbf{k})$ rather than, say, the ℓ_2-norm of $\mathbf{k}$ is now explained.

For spectral method problems using sampling on grids, the grids are not always square because there may be known strong dependencies on one or another of the coordinates. The input function may have a strong dependence on a particular coordinate or not even depend on some of the coordinates. Also, the operator $\mathcal{A}$ may be non-isotropic, as in the problem

$$100\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = f(\mathbf{x}).$$

If one samples on the grid,

$$\mathcal{P}_{\text{grid}} = \{(i_1/n_1, \ldots, i_s/n_s) : i_1 = 0, 1, \ldots, n_1 - 1, \ldots i_s = 0, 1, \ldots, n_2 - 1\},$$

with $n_1 \cdots n_s = N$, then the natural choice for $\mathcal{K}$ is the following rectangular box with volume N:

$$\begin{aligned}\mathcal{K}_{\text{grid}} = \{(k_1, \ldots, k_s) : k_1 = 1 - \lceil n_1/2\rceil, \ldots, n_1 - \lceil n_1/2\rceil, \ldots,\\ k_s = 1 - \lceil n_s/2\rceil, \ldots, n_s - \lceil n_s/2\rceil\}.\end{aligned}$$

Now consider the alternative, $\mathcal{K}_{\text{hyp}}$, the set of N wavenumbers that minimizes $\tilde{D}_\infty(\mathcal{P})$. This set, whose two dimensional projections are hyperbolic crosses, includes all $\mathbf{k}$ with $r(\mathbf{k}) \leq M$ for M as large as possible. Corollary 5 implies that there exist lattices with $M = \mathcal{O}(N^{1/2-\epsilon})$ for any $\epsilon > 0$. Moreover, the set $\mathcal{K}_{\text{hyp}}$ automatically includes all rectangular boxes $\mathcal{K}_{\text{grid}}$ with volume $\leq M = \mathcal{O}(N^{1/2-\epsilon})$. By contrast, if one chooses a fixed $\mathcal{K}_{\text{grid}}$ with volume N, it is only guaranteed to contain all other $\mathcal{K}_{\text{grid}}$ with volume equal to the smallest side of the original box, i.e., at best $\mathcal{O}(N^{1/s})$. In summary, for an isotropic problem, a square grid is a good choice, but if one may have unknown anisotropy, then a rank-1 lattice with wavenumber set of the form $\mathcal{K}_{\text{hyp}}$ is a safer choice. Note that the example in the previous section has not been chosen to be particularly anisotropic, however, the rank-1 lattice still outperforms the grid.

Acknowledgements

The authors would like to thank Profs. Bengt Fornberg, Alexander Keller and Tao Tang and two anonymous referees for their helpful suggestions that contributed to this manuscript.

References

1. B. Fornberg, *A practical guide to pseudospectral methods*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, 1995.
2. D. Gottlieb and S. A. Orszag, *Numerical analysis of spectral methods: Theory and applications*, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1977.
3. F. J. Hickernell, *Lattice rules: How well do they measure up?*, Random and Quasi-Random Point Sets (P. Hellekalek and G. Larcher, eds.), Lecture Notes in Statistics, vol. 138, Springer-Verlag, New York, 1998, pp. 109–166.
4. F. J. Hickernell and H. Niederreiter, *The existence of good extensible rank-1 lattices*, J. Complexity (2003), to appear.
5. L. K. Hua and Y. Wang, *Applications of number theory to numerical analysis*, Springer-Verlag and Science Press, Berlin and Beijing, 1981.
6. A. Keller, *Random fields on rank-1 lattices*, Technical Report 307/01, University of Kaiserslautern, 2001.
7. N. M. Korobov, *Exponential sums and their applications*, Kluwer Academic Publishers, Dordrecht, 1992.
8. H. Niederreiter, *Random number generation and quasi-Monte Carlo methods*, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1992.
9. I. H. Sloan and S. Joe, *Lattice methods for multiple integration*, Oxford University Press, Oxford, 1994.
10. L. N. Trefethen, *Spectral methods in MATLAB*, Software, Environments, Tools, SIAM, Philadelphia, 2000.
11. H. Weyl, *Über die Gleichverteilung der Zahlen mod Eins*, Math. Ann. **77** (1916), 313–352.

Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China

Current address: Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544-1000, United States of America

E-mail address: `dongli@princeton.edu`

Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China

E-mail address: `fred@hkbu.edu.hk`