Acta Numerica (1994), *pp.* 203–267

A review of pseudospectral methods for solving partial differential equations

Bengt Fornberg
Corporate Research
Exxon Research and Engineering Company
Annandale, NJ 08801, USA
E-mail: bfornbe@erenj.com

David M. Sloan
Department of Mathematics
University of Strathclyde
Glasgow G1 1XH, Scotland
E-mail: caas10@computer-centre-sun.strathclyde.ac.uk

CONTENTS

1. Introduction

Finite Difference (FD) methods approximate derivatives of a function by *local* arguments (such as $\mathrm{d}u(x)/\mathrm{d}x \approx (u(x+h) - u(x-h))/2h$, where h is a small grid spacing) – these methods are typically designed to be exact for polynomials of low orders. This approach is very reasonable: since the derivative is a local property of a function, it makes little sense (and is costly) to invoke many function values far away from the point of interest.

In contrast, spectral methods are *global*. The traditional way to introduce

them starts by approximating the function as a sum of very smooth basis functions:

$$u(x) = \sum_{k=0}^{N} a_k \Phi_k(x),$$

where the $\Phi_k(x)$ are, for example, Chebyshev polynomials or trigonometric functions – and then differentiate these exactly. In the context of solving time-dependent Partial Differential Equations (PDEs), this approach has notable strengths:

1 for analytic functions, errors typically decay (as N increases) at an *exponential* rather than at a (much slower) *polynomial* rate;
2 the method is virtually dissipation-free (in the context of solving high Reynolds number fluid flows, the low physical dissipation will not be overwhelmed by large numerical dissipation);
3 the approach is (surprisingly) powerful for many cases with nonsmooth or even discontinuous functions;
4 especially in several space dimensions, the relatively coarse grids which suffice for most accuracy requirements allow very time- and memory-effective calculations.

However, there are also factors which might cause difficulties or inefficiencies: certain boundary conditions; irregular domains; strong shocks; variable resolution requirements in different parts of a large domain; partly incomplete theoretical understanding.

In many applications where these disadvantages are not present (or they can somehow be overcome), FD or FE (Finite Element) methods do not even come close in efficiency. However, the situation in most major applications turns out less clear-cut than this. At present, spectral methods are highly successful in several areas such as turbulence modelling, weather prediction, nonlinear waves, seismic modelling etc. and the list is growing (see, for example, Boyd (1989) for examples and references).

Spectral methods have been a tool for analytic studies of differential equations since the days of Fourier (1822). The idea of using them for numerical solutions of ordinary differential equations (ODEs) goes back at least to Lanczos (1938). Their current popularity for PDEs dates back to the early 1970s and the works of Kreiss and Oliger (1972) and Orszag (1972) – facilitated by the Fast Fourier Transform (FFT) algorithm presented by Cooley and Tukey (1965).

Although spectral methods are normally introduced in the way we have indicated (through expansions using smooth global functions – the topic of Section 2), there is a useful alternative: pseudospectral (PS) methods can be seen as a special case of high-order FD methods (Fornberg 1975, 1987, 1990a,b). The introduction to Section 3 lists some of the advantages this

latter approach offers. In the remaining sections, key properties/variations of PS methods are discussed (using whichever viewpoint is most illuminating for the issue being discussed).

To obtain a uniform style, this review has been written and illustrated by BF from jointly prepared material. As a consequence, readers looking for functional analysis or proofs of technicalities are hereby warned not to waste their time proceeding beyond this point.

For one of us (BF), the interest in PS methods goes back to his PhD in 1972 under the supervision of H.-O. Kreiss. For later interactions, we wish, in particular, to acknowledge discussions with L.N. Trefethen.

2. Introduction to spectral methods via orthogonal functions

Spectral methods are usually described in the way we first indicated – as expansions based on global functions. Given a differential equation with boundary conditions, the idea is to approximate a solution $u(x)$ by a finite sum $v(x) = \sum_{k=0}^{N} a_k \Phi_k(x)$ (in the case of a time-dependent problem, $u(x,t)$ approximated by $v(x,t)$ and $a_k(t)$). Two main questions arise:

1 from which function class to choose $\Phi_k(x)$, $k = 0, 1, \ldots$ and
2 how to determine the expansion coefficients a_k.

These are addressed in Sections 2.1 and 2.2. Section 2.3 introduces cardinal functions and differentiation matrices – important tools for both understanding and computation (to be discussed in greater generality in Section 3.4 following the FD-based introduction to PS methods).

Reviews of this 'classical' approach to spectral methods can be found in Gottlieb and Orszag (1977), Voigt *et al.* (1984), Boyd (1989), Mercier (1989) and Funaro (1992).

2.1. Function classes

Three requirements need to be met:

1 The approximations $\sum_{k=0}^{N} a_k \Phi_k(x)$ to $v(x)$ must converge rapidly (at least for reasonably smooth functions).
2 Given coefficients a_k, the determination of b_k such that

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{k=0}^{N} a_k \Phi_k(x)\right) = \sum_{k=0}^{N} b_k \Phi_k(x) \tag{2.1}$$

should be efficient.
3 It should be possible to convert rapidly between coefficients a_k, $k = 0, \ldots, N$ and the values for the sum $v(x_i)$ at some set of nodes x_i, $i = 0, \ldots, N$.

Periodic problems The choice here is easy – *trigonometric expansions* satisfy all the requirements. The first two are immediate; the third was satisfied in 1965 through the FFT algorithm.

Non-periodic problems In this case, *trigonometric expansions* fail on requirement 1 – an irregularity will arise where the periodicity is artificially imposed. In the case of a discontinuity, a 'Gibbs' phenomenon' will occur (see Section 2.3). The coefficients a_n then decrease only like $\mathcal{O}(1/N)$ as $N \to \infty$.

Truncated *Taylor expansions* $v(x) = \sum_{k=0}^{N} a_k x^k$ will also fail on requirement 1: convergence over $[-1, 1]$ requires extreme smoothness of $v(x)$, i.e. analyticity throughout the unit circle.

The function class that has proven, by far, the most successful is *orthogonal polynomials* of Jacobi type with Chebyshev and Legendre polynomials as the most important special cases (cf. Table 1). These polynomials arise in many contexts:

- Gaussian integration formulae achieve high accuracy by using zeros of orthogonal polynomials as nodes.
- Singular Sturm–Liouville eigensystems are well known to offer superb bases for approximation – the Jacobi polynomials are the only polynomials arising in this way.
- Truncated expansions in *Legendre polynomials* are optimal in the L^2-norm (for max-norm approximations of smooth functions, truncated *Chebyshev expansions* are particularly accurate).
- Interpolation at the Chebyshev nodes

$$x_k = -\cos(\pi k/N), \quad k = 0, 1, \ldots, N$$

give polynomials P_N^{CH} which are always within a very small factor of the optimal in max-norm approximation of any function $f(x)$:

$$\|f - P_N^{\mathrm{CH}}\| \le (1 + \Lambda_N^{\mathrm{CH}})\|f - P_N^{\mathrm{OPT}}\|.$$

Here Λ_N^{CH} is known as the Lebesgue constant of order N for Chebyshev interpolation. It depends only on N; properties of f affect $\|f - P_N^{\mathrm{OPT}}\|$ as described by Jackson's theorems (Cheney, 1966; Powell, 1981). Λ_N^{CH} is smaller than the constant for interpolation using Legendre expansions and far superior to the disastrous one for equi-spaced interpolation (more on this 'Runge phenomenon' in Section 3.3).

$$\Lambda_N^{\mathrm{CH}} = \mathcal{O}(\log N), \quad \Lambda_N^{\mathrm{LEG}} = \mathcal{O}(\sqrt{N}), \quad \Lambda_N^{\mathrm{EQI}} = \mathcal{O}(2^N/N \log N).$$

For references on the three Lebesgue constants, see Rivlin (1969, p. 90), Szegő (1959, p. 336) and Trefethen and Weideman (1991) respectively.

These points confirm that Jacobi polynomials satisfy requirement 1. Because of the first derivative recursions (and the lack of explicit x-dependence

Table 1. *Jacobi polynomials' fact sheet*

	LEGENDRE	CHEBYSHEV	JACOBI
Weight function $W(x)$	1	$\frac{1}{\sqrt{1-x^2}}$	$(1-x)^\alpha(1+x)^\beta$ $(\alpha=\beta=0$ Legendre) $\alpha>-1,\ \beta>-1$ $(\alpha=\beta=-\frac{1}{2}$ Chebyshev)
First few polynomials	1 x $\frac{3}{2}x^2-\frac{1}{2}$ $\frac{5}{2}x^3-\frac{3}{2}x$ $\frac{35}{8}x^4-\frac{15}{4}x^2+\frac{3}{8}$ $\frac{63}{8}x^5-\frac{35}{4}x^3+\frac{15}{8}x$	1 x $2x^2-1$ $4x^3-3x$ $8x^4-8x^2+1$ $16x^5-20x^3+5x$	1 $\frac{1}{2}(2+\alpha+\beta)x+\frac{1}{2}(\alpha-\beta)$ $\frac{1}{8}(2+\alpha+\beta)(4+\alpha+\beta)x^2+\frac{1}{4}(\alpha-\beta)(3+\alpha+\beta)x$ $+\frac{1}{8}[(\alpha-\beta)^2-(4+\alpha+\beta)]$ General n : $2^{-n}\sum_{k=0}^{n}\binom{n+\alpha}{k}\binom{n+\beta}{n-k}(x-1)^{n-k}(x+1)^k$
Orthogonality $\int_{-1}^{1}\Phi_m\Phi_n W\,\mathrm{d}x$	0 $: m\neq n$ $\frac{2}{2n+1}$ $: m=n$	0 $: m\neq n$ π $: m=n=0$ $\frac{\pi}{2}$ $: m=n>0$	0 $: m\neq n$ $\frac{2^{\alpha+\beta+1}\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{(2n+\alpha+\beta+1)n!\Gamma(n+\alpha+\beta+1)}$ $: m=n$
Three term recursion	$(n+1)L_{n+1}$ $-(2n+1)xL_n$ $+nL_{n-1}=0$	$T_{n+1}-2xT_n$ $+T_{n-1}=0$	$2(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta)P_{n+1}$ $-[(2n+\alpha+\beta+1)(\alpha^2-\beta^2)$ $+(2n+\alpha+\beta)(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)x]P_n$ $+2(n+\alpha)(n+\beta)(2n+\alpha+\beta+2)P_{n-1}=0$
Differential equation	$(1-x^2)L_n''-2xL_n'$ $+n(n+1)L_n=0$	$(1-x^2)T_n''-xT_n'$ $+n^2T_n=0$	$(1-x^2)P_n''+[(\beta-\alpha)-(\alpha+\beta+2)x]P_n'$ $+n(n+\alpha+\beta+1)P_n=0$
First derivative recursion	L_{n+1}' $=(2n+1)L_n+L_{n-1}'$	$\frac{1}{n+1}T_n'$ $=2T_n+\frac{1}{n-1}T_{n-1}'$	$2(2n+\alpha+\beta)(n+\alpha+\beta)(n+\alpha+\beta+1)P_{n+1}'$ $+2(\alpha-\beta)(n+\alpha+\beta)(2n+\alpha+\beta+1)P_n'$ $-2(n+\alpha)(n+\beta)(2n+\alpha+\beta+2)P_{n-1}'$ $=(n+\alpha+\beta)(2n+\alpha+\beta)(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)P_n$

in these, see Table 1), requirement 2 is met. In the case of Chebyshev polynomials, the relation sought in (2.1) becomes

$$\begin{bmatrix} 1 & 0 & -\frac{1}{2} & & & & \\ & \frac{1}{4} & 0 & -\frac{1}{4} & & & \\ & & \frac{1}{6} & 0 & -\frac{1}{6} & & \\ & & & \frac{1}{8} & 0 & -\frac{1}{8} & \\ & & & & \ddots & \ddots & \ddots \\ & & & & & \frac{1}{2N-4} & 0 \\ & & & & & & \frac{1}{2N-2} \end{bmatrix} \times \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_{N-2} \\ b_{N-1} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ \vdots \\ a_{N-1} \\ a_N \end{bmatrix}. \quad (2.2)$$

and similarly for Legendre and Jacobi expansions. Requirement 3 is clearly satisfied for the Chebyshev case if we choose $x_i = -\cos(\pi i/N)$, $i = 0, \ldots, N$ (cosine FFT) – as we shall see later, the additional cost in the other cases need not be prohibitive. For these reasons, Chebyshev (and to a lesser extent, Legendre) polynomials have become the almost universally preferred choice for nonperiodic spectral approximations.

The special case of Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ with $\alpha = \beta$ goes under its own name – Gegenbauer or Ultraspherical polynomials (again including Legendre and Chebyshev as special cases). For an overview of orthogonal polynomials, see, for example, Sansone (1959) and Szegő (1959).

Formulas similar to (2.1) are also available when (2.1) is generalized, for example, to

$$x^q \frac{\mathrm{d}^p}{\mathrm{d}x^p}\left(\sum_{k=0}^{N} a_k \phi_k(x)\right) = \sum_{k=0}^{N} b_k \phi_k(x), \quad p, q = 1, 2, \ldots$$

(see the Appendix in Gottlieb and Orszag (1977)). These are essential for Tau and Galerkin – but not for collocation (PS) approximations; see the next section and Appendix 1.

2.2. Techniques to determine the expansion coefficients

The three main techniques to determine expansion coefficients a_k are Tau, Galerkin and Collocation (PS). In all cases, we consider the residual $R_n(x)$ (or $R_n(x,t)$) when the expansion is substituted into the governing equation. We need to keep this residual small across the domain and also to satisfy the boundary conditions.

- In the Tau technique the a_k have to be selected so that the boundary conditions are satisfied and the residual has to be made orthogonal to as many of the basis functions as possible.
- For the Galerkin technique we combine the original basis functions into a new set in which all the functions satisfy the boundary conditions.

Then the residual is required to be orthogonal to as many of these new basis functions as possible.

- Finally the Collocation (PS) technique is similar to the Tau one: the a_k have to be selected so that the boundary conditions are satisfied but the residual is made zero at as many (suitably chosen) spatial points as possible.

Implementation details for a model problem are given in Appendix 1.

The Tau method was first used by Lanczos (1938). The Galerkin idea is central to FE methods. Spectral (global) versions of it have been in use since the mid-1950s. The FFT algorithm as well as contributions by Orszag (1969, 1970; on ways to deal with nonlinearities) have contributed to its current usage. The collocation approach was first used for PDEs with periodic solutions by Kreiss and Oliger (1972). It was referred to as the 'pseudospectral method' in Orszag (1972).

The collocation (PS) method can be viewed as a method for finding numerical approximations to derivatives at grid points. Then, in a FD-like manner, the governing equations are satisfied pointwise in physical space. This makes the PS method particularly easy to apply to equations with variable coefficients and nonlinearities, as these only give rise to products of numbers rather than to problems of determining the expansion coefficients for products of expansions.

The rest of this review article will focus entirely on the PS method.

2.3. Cardinal functions – example of a differentiation matrix

The concepts of Cardinal Functions (CFs) and of Differentiation Matrices (DMs) are both theoretically and numerically useful well beyond the realm of methods derived from orthogonal functions. Therefore, we postpone the main discussion of these until Section 4.3 (when our background is more general) and consider them here only in the case of the Fourier PS method.

The trigonometric polynomial which interpolates periodic data can be thought of as a weighted sum of CFs, each with the property of having unit value at one of the data points and zero at the rest. This is very similar to how Lagrange's interpolation formula works – the main difference being that, in this Fourier case, all the CFs are just translates of each other. For references to CFs, see E.T. Whittaker (1915), J.M. Whittaker (1927), Stenger (1981).

Assume for simplicity that we have an odd number $n = 2m + 1$ of grid points, at locations $x_i = i/(m + \frac{1}{2})$, $i = -m, \ldots, -1, 0, 1, \ldots, m$ in $[-1, 1]$. By inspection,

$$\Phi_m(x) = \frac{1}{m + \frac{1}{2}} \{ \tfrac{1}{2} + \cos \pi x + \cos 2\pi x + \ldots + \cos m\pi x \} = \frac{\sin(m + \frac{1}{2})\pi x}{(2m + 1) \sin \frac{1}{2}\pi x} \tag{2.3}$$

is of the right form and satisfies

$$\Phi_m(x_i) = \begin{cases} 1, i = 0 & [\pm(2m+1), \pm(4m+2), \\ & \text{etc., if periodically extended}] \\ 0 & \text{otherwise.} \end{cases}$$

Figure 1 displays the CF $\Phi_8(x)$ and illustrates how its translates add up to give the trigonometric interpolant to a step function, and the Gibbs' phenomenon.

From (2.3) it follows that

$$\frac{\mathrm{d}}{\mathrm{d}x}\Phi_m(x_i) = \begin{cases} 0, i = 0 & [\pm(2m+1), \pm(4m+2), \text{ etc.,} \\ & \text{if periodically extended}] \\ \dfrac{(-1)^i \pi}{2\sin[i\pi/(2m+1)]} & \text{otherwise.} \end{cases}$$

The PS derivative $v'(x_k)$ of a vector of data values $v(x_k), k = -m, \ldots, m$ can therefore be obtained as a matrix $\times$ vector product. The element at position (i, j) of this matrix is equal to $\mathrm{d}\Phi_m(x_{i-j})/\mathrm{d}x$. This matrix is the DM for the periodic PS method (first derivative, odd number of points).

3. Introduction to PS methods via finite differences

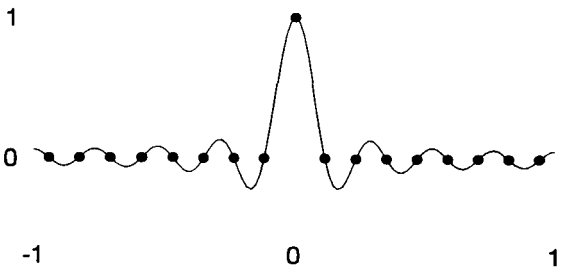
FD formulae of increasing orders of accuracy provide not only an alternative introduction to PS methods (for both the periodic and nonperiodic cases); they suggest generalizations and offer additional insights.

- Orthogonal polynomials/functions lead only to a small class of possible spectral methods – the FD viewpoint allows many generalizations. For example, all classical orthogonal polynomials cluster the nodes quadratically towards the ends of the interval – this is often, but not always best.
- An FD viewpoint offers a chance to explore 'intermediate' methods between low-order FD and PS. One might consider methods of not quite as high order as the Chebyshev spectral method and with nodes not clustered quite as densely – possibly trading some accuracy for stability and simplicity.
- Two separate ways to view any method always give more opportunities to understand/improve/analyse it.
- Many special enhancements have been designed for FD methods. Viewing PS methods as a special case of FD methods often makes it easier to carry over such ideas. Examples include staggered grids, upwind techniques, enhancements at boundaries, polar and spherical coordinates, etc. (discussed in Section 5).
- Comparisons between PS and FD methods can be made more consistent. PS methods represent limits of increasingly accurate FD methods

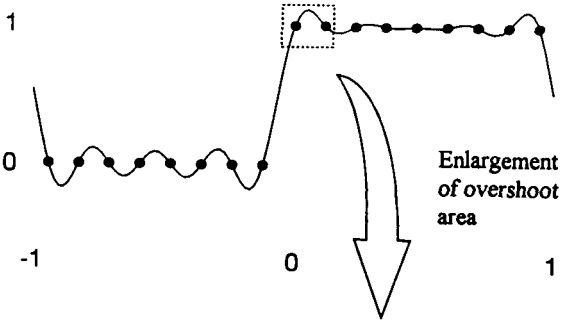
FOURIER CARDINAL FUNCTION (FCF)

$$\Phi_m(x) = \frac{\sin(m+\frac{1}{2})\pi x}{(2m+1)\sin\frac{\pi x}{2}}$$

shown for $m = 8$.



Sum of translates of FCFs = Fourier interpolation of a step function.



Equi-spaced Fourier interpolation →

Truncated Fourier series →

GIBBS' PHENOMENON

Overshoots (at each side of jump) approx. 14 % and 9 % resp. of its height (as $m \to \infty$).

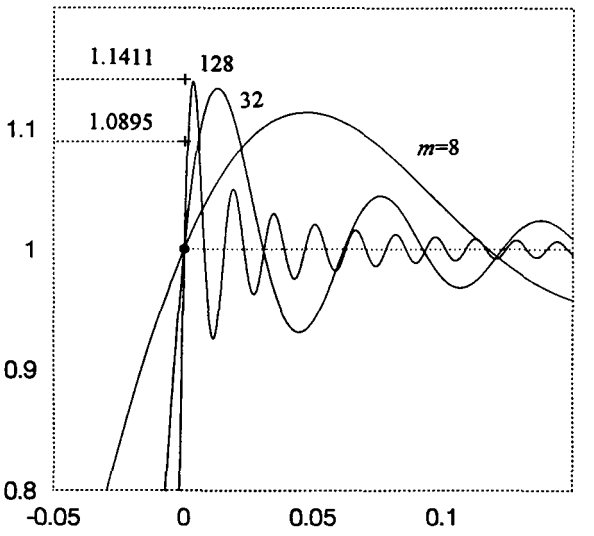


Fig. 1. Fourier CFs and Gibbs' phenomenon.

– the FD viewpoint provides a unifying framework in which to understand and interpret all these methods.

Sections 3.1 and 3.2 provide some general material on FD approximations, allowing us in Section 3.3 to discuss different types of node distribution. In Section 3.4, we derive the DM for the case that was considered in Section 2.3 (periodic problem, equi-spaced grid) and obtain an identical result – hence the methods are equivalent. In Section 3.5, this equivalence is seen to be very general.

3.1. Algorithm to find FD weights on arbitrarily spaced grids

Centred FD formulae for equi-spaced grids are readily available from tables and can be derived by symbolic manipulation of difference operators. For example, the centred approximations (at a grid point) to the first derivative are:

$$\begin{aligned} f'(x) &= [\quad -\tfrac{1}{2}f(x-h)+0f(x)+\tfrac{1}{2}f(x+h) \quad]/h+\mathcal{O}(h^2) \\ &= [\tfrac{1}{12}f(x-2h)-\tfrac{2}{3}f(x-h)+0f(x)+\tfrac{2}{3}f(x+h)-\tfrac{1}{12}f(x+2h)]/h+\mathcal{O}(h^4) \\ &\quad \vdots \\ &\quad \text{etc.}, \end{aligned}$$

exact for all polynomials of degrees $2, 4, \ldots$ resp. It is convenient to collect weights like these as is done in Table 2 (for the kth derivative, divide with h^k).

Another equi-spaced case of interest is one-sided stencils which are often necessary use at boundaries. Table 3 shows some weights in this case. The weights for the first derivative are the ones that arise in backward differentiation formulae for ODEs, see Lambert (1991).

To explore the general properties of FD schemes (and, more importantly, to use such schemes), it is desirable to have a simple algorithm for the more general problem:

Given: $x_0, x_1, \ldots, x_n$: grid points (nonrepeated, otherwise arbitrary)
ξ: point $x = \xi$ at which the approximations are wanted (may, but need not be at a grid point)
m: highest order of derivative of interest

Find: weights $c_{i,j}^k$ such that the approximations

$$\left.\frac{\mathrm{d}^k f}{\mathrm{d}x^k}\right|_{x=\xi} \approx \sum_{j=0}^{i} c_{i,j}^k f(x_j), \quad k = 0, 1, \ldots, m, \; i = k, k+1, \ldots, n$$

are all optimal.

A short and fast algorithm for this was discovered only recently (Fornberg, 1988b, in more detail 1992):

Table 1. *Weights for some centred FD formulae on an equi-spaced grid: D, order of derivative; A, order of accuracy.*

		Approximations at $x = 0$ (x = coordinates at nodes)								
D	A	−4	−3	−2	−1	0	1	2	3	4
0	∞					1				
1	2				$-\frac{1}{2}$	0	$\frac{1}{2}$			
	4			$\frac{1}{12}$	$-\frac{2}{3}$	0	$\frac{1}{2}$	$-\frac{1}{12}$		
	6		$-\frac{1}{60}$	$\frac{3}{20}$	$-\frac{3}{4}$	0	$\frac{3}{4}$	$-\frac{3}{20}$	$\frac{1}{60}$	
	8	$\frac{1}{280}$	$-\frac{4}{105}$	$\frac{1}{5}$	$-\frac{4}{5}$	0	$\frac{4}{5}$	$-\frac{1}{5}$	$\frac{4}{105}$	$-\frac{1}{280}$
2	2				1	-2	1			
	4			$-\frac{1}{12}$	$\frac{4}{3}$	$-\frac{5}{2}$	$\frac{4}{3}$	$-\frac{1}{12}$		
	6		$\frac{1}{90}$	$-\frac{3}{20}$	$\frac{3}{2}$	$-\frac{49}{18}$	$\frac{3}{2}$	$-\frac{3}{20}$	$\frac{1}{90}$	
	8	$-\frac{1}{560}$	$\frac{8}{315}$	$-\frac{1}{5}$	$\frac{8}{5}$	$-\frac{205}{72}$	$\frac{8}{5}$	$-\frac{1}{5}$	$\frac{8}{315}$	$-\frac{1}{560}$
3	2			$-\frac{1}{2}$	1	0	−1	$\frac{1}{2}$		
	4		$\frac{1}{8}$	−1	$\frac{13}{8}$	0	$-\frac{13}{8}$	1	$-\frac{1}{8}$	
	6	$-\frac{7}{240}$	$\frac{3}{10}$	$-\frac{169}{120}$	$\frac{61}{30}$	0	$-\frac{61}{30}$	$\frac{169}{120}$	$-\frac{3}{10}$	$\frac{7}{240}$
4	2			1	−4	6	−4	1		
	4		$-\frac{1}{6}$	2	$-\frac{13}{2}$	$\frac{28}{3}$	$-\frac{13}{2}$	2	$-\frac{1}{6}$	
	6	$\frac{7}{240}$	$-\frac{2}{5}$	$-\frac{122}{15}$	$\frac{91}{8}$	$-\frac{122}{15}$	$\frac{169}{60}$	$-\frac{2}{5}$	$\frac{7}{240}$	

$c^0_{0,0} := 1,\ \alpha := 1$
for $i := 1$ to n
    $\beta := 1$
    for $j := 0$ to $i-1$
        $\beta := \beta(x_i - x_j)$
        for $k := 0$ to $\min(i, m)$
            $c^k_{i,j} := ((x_i - \xi)c^k_{i-1,j} - kc^{k-1}_{i-1,j})/(x_i - x_j)$
    for $k := 0$ to $\min(i, m)$
        $c^k_{i,j} := \alpha(kc^{k-1}_{i-1,i-1} - (x_{i-1} - \xi)c^k_{i-1,i-1})/\beta$
    $\alpha := \beta$

Notes.

1 Any noninitialized quantity referred to is assumed to be equal to zero.
2 Only four operations are needed for each weight (to leading order; note that the subtractions $x_i - \xi$ and $x_i - x_j$ can be moved out of the innermost loop).

Table 2. *Weights for some one-sided FD formulae on an equi-spaced grid: D, order of derivative; A, order of accuracy.*

		Approximations at $x = 0$ (x = coordinates at nodes)								
D	A	0	1	2	3	4	5	6	7	8
0	∞	1								
1	1	-1	1							
	2	$-\frac{3}{2}$	2	$-\frac{1}{2}$						
	3	$-\frac{11}{6}$	3	$-\frac{3}{2}$	$\frac{1}{3}$					
	4	$-\frac{25}{12}$	4	-3	$\frac{4}{3}$	$-\frac{1}{4}$				
	5	$-\frac{137}{60}$	5	-5	$\frac{10}{3}$	$-\frac{5}{4}$	$\frac{1}{5}$			
	6	$-\frac{49}{20}$	6	$-\frac{15}{2}$	$\frac{20}{3}$	$-\frac{15}{4}$	$\frac{6}{5}$	$-\frac{1}{6}$		
	7	$-\frac{363}{140}$	7	$-\frac{21}{2}$	$\frac{35}{3}$	$-\frac{35}{4}$	$\frac{21}{5}$	$-\frac{7}{6}$	$\frac{1}{7}$	
	8	$-\frac{761}{280}$	8	-14	$\frac{56}{3}$	$-\frac{35}{2}$	$\frac{56}{5}$	$-\frac{14}{3}$	$\frac{8}{7}$	$-\frac{1}{8}$
2	1	1	-2	1						
	2	2	-5	4	-1					
	3	$\frac{35}{12}$	$-\frac{26}{3}$	$\frac{19}{2}$	$-\frac{14}{3}$	$\frac{11}{12}$				
	4	$\frac{15}{2}$	$-\frac{77}{6}$	$\frac{107}{6}$	-13	$\frac{61}{12}$	$-\frac{5}{6}$			
	5	$\frac{203}{45}$	$-\frac{87}{5}$	$\frac{117}{4}$	$-\frac{254}{9}$	$\frac{33}{2}$	$-\frac{27}{5}$	$\frac{137}{180}$		
	6	$\frac{469}{90}$	$-\frac{223}{10}$	$\frac{879}{20}$	$-\frac{949}{18}$	41	$-\frac{201}{10}$	$\frac{1019}{180}$	$-\frac{7}{10}$	
	7	$\frac{29531}{5040}$	$-\frac{962}{35}$	$\frac{621}{10}$	$-\frac{4006}{45}$	$\frac{691}{8}$	$-\frac{282}{5}$	$\frac{2143}{90}$	$-\frac{206}{35}$	$\frac{363}{560}$
3	1	-1	3	-3	1					
	2	$-\frac{5}{2}$	9	-12	7	$-\frac{3}{2}$				
	3	$-\frac{17}{4}$	$\frac{71}{4}$	$-\frac{59}{2}$	$\frac{49}{2}$	$-\frac{41}{4}$	$\frac{7}{4}$			
	4	$-\frac{49}{8}$	29	$-\frac{461}{8}$	62	$-\frac{307}{8}$	13	$-\frac{15}{8}$		
	5	$-\frac{967}{120}$	$\frac{638}{15}$	$-\frac{3929}{40}$	$\frac{389}{3}$	$-\frac{2545}{24}$	$\frac{268}{5}$	$-\frac{1849}{120}$	$\frac{29}{15}$	
	6	$-\frac{801}{80}$	$\frac{349}{6}$	$-\frac{18353}{120}$	$\frac{2391}{10}$	$-\frac{1457}{6}$	$\frac{4891}{30}$	$-\frac{561}{8}$	$\frac{527}{30}$	$-\frac{469}{240}$
4	1	1	-4	6	-4	1				
	2	3	-14	26	-24	11	-2			
	3	$\frac{35}{6}$	-31	$\frac{137}{2}$	$-\frac{242}{3}$	$\frac{107}{2}$	-19	$\frac{17}{6}$		
	4	$\frac{28}{3}$	$-\frac{111}{2}$	142	$-\frac{1219}{6}$	176	$-\frac{185}{2}$	$\frac{82}{3}$	$-\frac{7}{2}$	
	5	$\frac{1069}{80}$	$-\frac{1316}{15}$	$\frac{15289}{60}$	$-\frac{2144}{5}$	$\frac{10993}{24}$	$-\frac{4772}{15}$	$\frac{2803}{20}$	$-\frac{536}{15}$	$\frac{967}{240}$

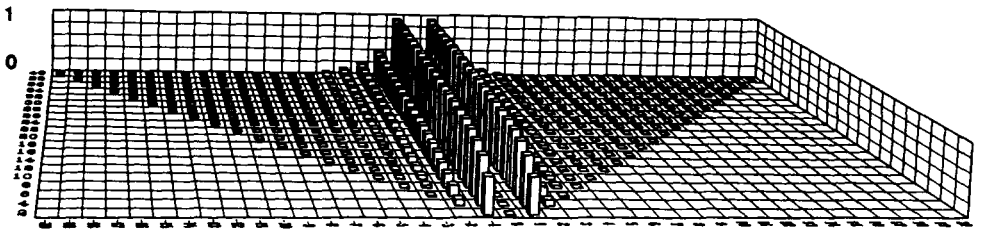


Fig. 2. Magnitude of weights for centred approximations to the first derivative on an equi-spaced grid (cf. Table 2).

3 The calculation of the weights is numerically stable (however, especially in the case of high derivatives, *applying* FD weights to a function may lead to severe cancellations and loss of significant digits).

4 The special case $m = 0$ offers the fastest way known to perform polynomial interpolation at a single point (in particular, significantly faster than the classical algorithms by Aitken and Neville).

5 If we are only interested in the weights for the stencils based on all the grid points $x_j, j = 0, 1, \ldots, n$ (and not in the lower-order stencils based on fewer points), we can omit the first of the two subscripts for c (i.e. it suffices to declare a two-dimensional array to hold $c_{0 \text{ to } n}^{0 \text{ to } m}$ – the 'overwriting' that will occur internally in the algorithm will be safe).

A Fortran code (with test driver) for this algorithm is given in Appendix 2.

3.2. Growth rates of FD weights on equi-spaced grids

Figures 2 and 3 illustrate how the magnitudes of the weights for the first derivative grow with increasing orders of accuracy (cf. Tables 2 and 3).

In the centred case, approximations of increasing orders of accuracy converge to a limit method of formally infinite-order of accuracy. For the first derivative ($m = 1$), this can be seen directly from the closed form expression for the weights $c_{p,j}^1$ (p (even) = order of accuracy, $j = x$-position of weight):

$$c_{p,j}^1 = \begin{cases} \dfrac{(-1)^{j+1}(\frac{1}{2}p)!^2}{j(\frac{1}{2}p+j)!(\frac{1}{2}p-j)!} & j = \pm 1, \pm 2, \ldots, \pm\tfrac{1}{2}p \\ 0 & j = 0. \end{cases}$$

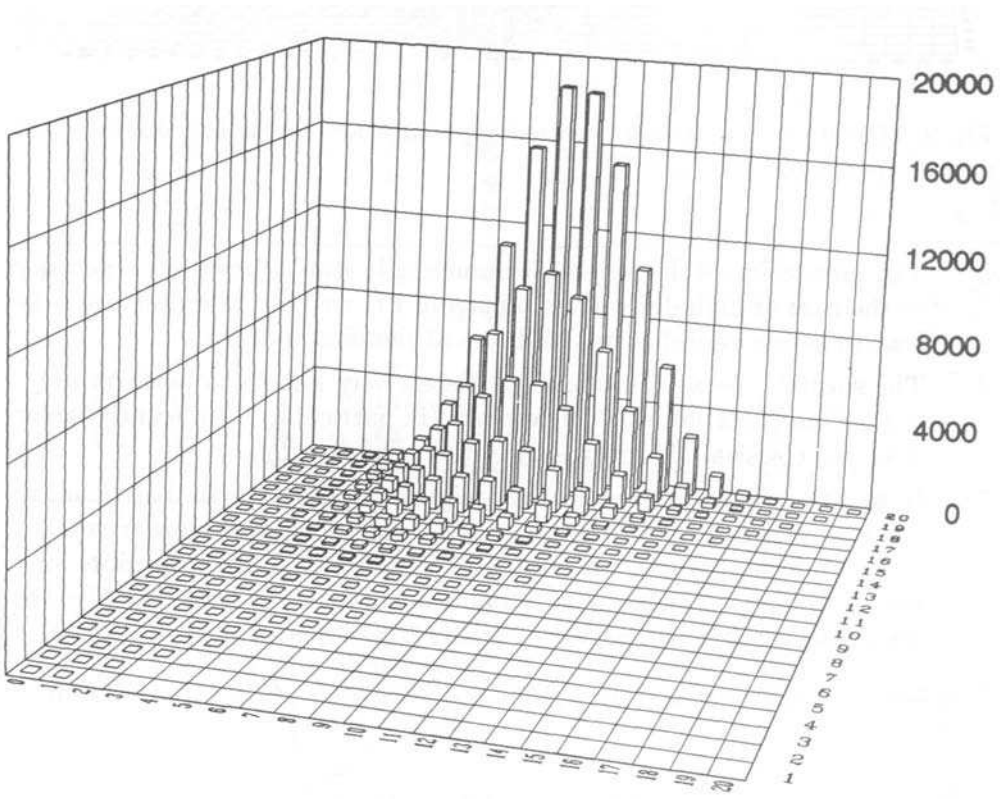


Fig. 3. Magnitude of weights for one-sided approximations to the first derivative on an equi-spaced grid (cf. Table 3).

Clearly, the limit for $p \to \infty$ exists and

$$c^1_{\infty,j} = \begin{cases} \dfrac{(-1)^{j+1}}{j} & j = \pm 1, \pm 2, \ldots \\ 0 & j = 0. \end{cases} \tag{3.1}$$

Beyond the second derivative, for which

$$c^2_{p,j} = 2c^1_{p,j}/j, \quad j = \pm 1, \pm 2, \ldots, \pm \tfrac{1}{2}p, \quad c^2_{p,0} = -2\sum_{i=1}^{p/2} 1/i^2,$$

closed form expressions for weights become very complicated. However, that does not affect the ease with which they can be calculated (using the algorithm in Section 3.1) or the existence of simple limits. For the second derivative, the limit becomes

$$c^2_{\infty,j} = \begin{cases} \dfrac{2(-1)^{j+1}}{j^2} & j = \pm 1, \pm 2, \ldots \\ -\frac{1}{3}\pi^2 & j = 0. \end{cases}$$

For higher derivatives the decay rates alternate between $\mathcal{O}(1/j)$ and $\mathcal{O}(1/j^2)$ for odd and even derivatives respectively (exact formulae for

$$c^m_{\infty,j}, \quad m = 1, 2, \ldots$$

are given in Fornberg (1990a)).

The situation is very different for one-sided approximations. The closed form expression for the first derivative is

$$c^1_{p,j} = \begin{cases} \dfrac{(-1)^{j+1}}{j} \begin{pmatrix} p \\ j \end{pmatrix} & j = 1, 2, \ldots, p \\ -\sum_{i=1}^{p} 1/i & j = 0. \end{cases}$$

The magnitudes of the weights form (nearly) a Gaussian distribution, which becomes increasingly peaked at the centre of the stencil while growing in height exponentially with p ($\sim \pi^{-1/2} p^{-3/2} 2^{p+3/2}$). For higher derivatives, the general character and growth rates remain similar. Partly one-sided approximations initially grow more slowly but will also ultimately diverge exponentially. In the case of the first derivative (just described), the asymptotic rate is multiplied by a factor of $s!/p^s$ if the derivative is evaluated s steps in from the boundary.

3.3. Generalized node distributions

The previous discussion about the size of the weights in centred versus one-sided FD formulae suggests, for a nonperiodic problem, that the nodes need to be concentrated towards the ends of the interval (to offset the loss of accuracy which would otherwise arise because of the very large weights).

Let $\mu(x)$ denote how the *density* of a node distribution varies over $[-1, 1]$ (i.e. the distance between adjacent nodes decreases like $1/n\mu(x)$ as n increases. To prevent changes in $\mu(x)$ affecting the total number of nodes that fit into $[-1, 1]$, we require $\int_{-1}^{1} \mu(x)\,dx = 1$. A one-parameter family

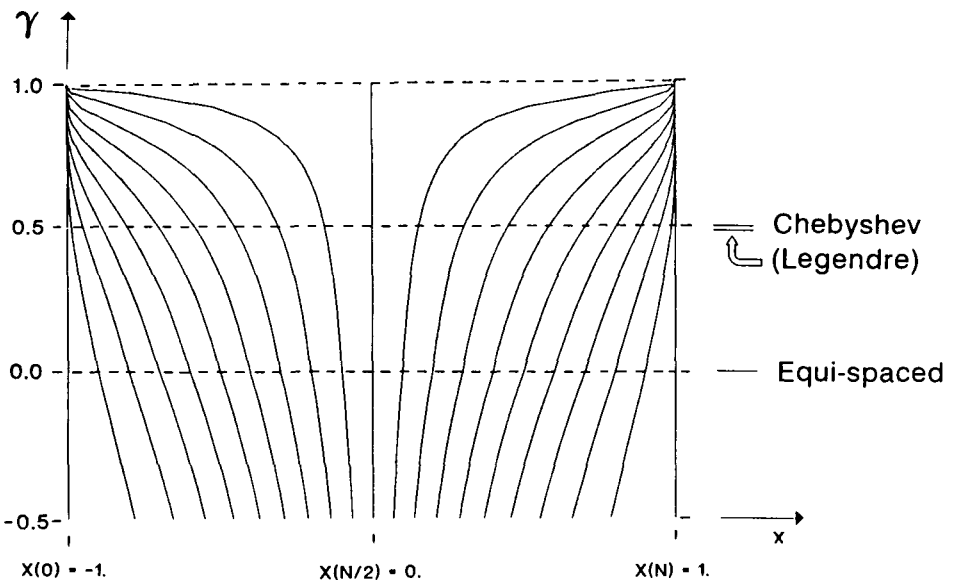


Fig. 4. Distribution of nodes corresponding to density function $\mu_\gamma(x)$, shown for $\gamma \in [-0.5, 1]$ ($n = 20$). The Legendre distribution (of extrema) is not obtained exactly for any γ, but $\gamma \approx 0.4785$ gives the closest (least-squares) fit in this case of $n = 20$. This difference (to 0.5 – it vanishes as $n \to \infty$) is illustrated at the right edge of the figure.

of density functions is outlined below, followed by two special cases that it incorporates:

Density function	Node locations $x_j,\ j = 0, 1, \ldots, n$	Comments
$\mu_\gamma(x) = c_\gamma/(1-x^2)^\gamma$	$j/n = \int_{-1}^{x_j} \mu_\gamma(x)\,\mathrm{d}x$	$\gamma < 1;\ c_\gamma = \dfrac{\Gamma(\frac{3}{2}-\gamma)}{\pi^{1/2}\Gamma(1-\gamma)}$
$\mu_0(x) \equiv \frac{1}{2}$	$x_j = -1 + 2j/n$	$\gamma = 0$; equi-spaced
$\mu_{1/2}(x) = 1/\pi\sqrt{1-x^2}$	$x_j = -\cos(\pi j/n)$	$\gamma = \frac{1}{2}$; Chebyshev

Figure 4 shows how the nodes move as a function of γ. One key question is whether quadratic clustering (the case $\gamma = 0.5$) is necessary. Several general arguments suggest this in the limit of $n \to \infty$ (with the FD approach to PS methods, the effects of other clusterings can be explored in special cases – see Section 5):

All classical orthogonal polynomials feature quadratic node clustering at the ends. Changing α and β in the weight function $(1-x)^\alpha(1-x)^\beta$ for Jacobi polynomials will still leave the nodes quadratically clustered

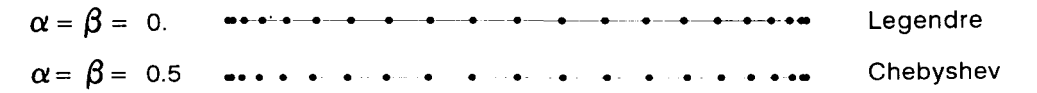


Fig. 5. Difference between the location of extrema for Legendre and Chebyshev polynomials ($n = 20$).

(this follows, for example, from their differential equation, see Table 1). Figure 5 compares the nodes (extrema) for Legendre and Chebyshev polynomials of order 20 (corresponding to $\alpha = \beta = 0$ and $\alpha = \beta = \frac{1}{2}$) - there is hardly any noticeable difference. Figure 4 also illustrates this, i.e. how small an effect changing $\alpha = \beta$ has compared with changes in γ.

To get the least possible interpolation error, the nodes must cluster as in the Chebyshev case. A heuristic argument for this goes as follows: let $p_n(x)$ be the interpolation polynomial of degree n to $f(x)$ on $[-1, 1]$. The remainder term is

$$f(x) - p_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi) \prod_{j=0}^{n} (x - x_j)$$

for some $\xi \in [-1, 1]$. The only part that can be controlled by re-positioning the nodes x_j is the product. Since the highest order term is $1 \cdot x^{n+1}$, the question becomes: which polynomial of that form stays smallest over $[-1, 1]$? This is a well known property of Chebyshev polynomials.

With any other type of clustering, convergence will require f to be analytic in some domain away from the interval $[-1, 1]$. In a complex $z = x + iy$ plane, a Taylor series converges in the *largest circle* around the expansion point that is free of singularities. This result generalizes to interpolating polynomials (when the nodes are distributed over an interval rather than all lumped at one point) as follows (Krylov (1962), Ch. 12, Markushevich (1967); general results on polynomial interpolation can further be found in Walsh (1960), Davis (1975), Gaier (1987) etc.).

Given a node density $\mu(x)$ (on $[-1, 1]$), form the potential function

$$\phi(z) = -\int_{-1}^{1} \mu(x) \log |z - x| \, dx \quad + \text{constant.} \tag{3.2}$$

Then $p_n(z)$ converges to $f(z)$ inside the *largest equi-potential curve* that does not enclose any singularity of $f(z)$ (and diverges outside it).

Figure 6(a) shows (as a thick straight line above the x-axis, at centre of the figure) the graph of $\mu_0(x) \equiv \frac{1}{2}$ and a matching equi-spaced set of nodes on the x-interval $[-1, 1]$ (barely visible). The heavy contour line on the

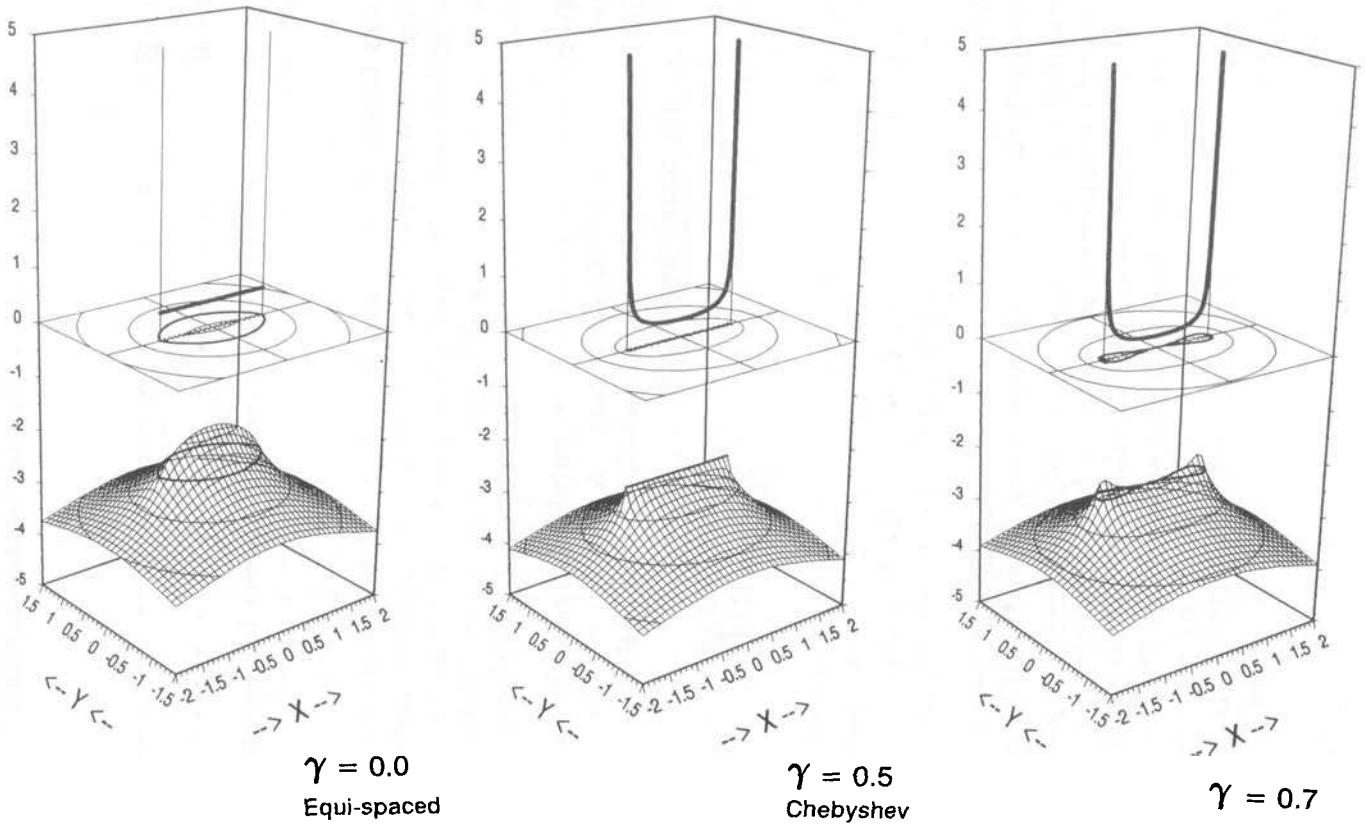


Fig. 6. Node density functions $\mu_\gamma(x)$ and their corresponding logarithmic potentials $\phi_\gamma(x, y)$. The heavy contour lines mark the regions which must be free from singularities for convergence to occur (over $[-1, 1]$) as the number of nodes $n \to \infty$. Lower contour lines (vertically separated by 0.5) correspond to geometric convergence rates α^n, $\alpha = e^{-0.5} \approx 0.607$, $e^{-1.0} \approx 0.368$, $e^{-1.5} \approx 0.223$ etc.

potential surface

$$\phi(z) = \tfrac{1}{2}\mathrm{Re}[(1-z)\log(1-z) - (-1-z)\log(-1-z)] + C$$

surrounds the smallest domain that includes $[-1,1]$ and is bounded by an equipotential curve. The function $f(z)$ must therefore be analytic everywhere within this domain for convergence to occur on $[-1,1]$. Any singularity within this domain restricts convergence to a still smaller equi-potential region, leading to the 'Runge phenomenon' – divergence near the ends of the interval.

In the Jacobi polynomial case

$$\mu(x) = 1/(\pi\sqrt{1-x^2}).$$

Equation (3.2) can then again be evaluated in closed form:

$$\phi(z) = -|\log|z + \sqrt{z^2-1}|| + C$$

(like the previous formula, correct for all complex values of z when selecting the conventional branches for the logarithm and square root functions).

Figure 6(b) shows how the potential surface $\phi(z)$ forms a perfectly flat ridge of the potential surface along $[-1,1]$. This clearly looks optimal – the only possibility that convergence to $f(z)$ on $[-1,1]$ does not require $f(z)$ to be analytic anywhere off the interval $[-1,1]$.

Finally, Figure 6(c) shows what happens when the nodes cluster still more densely towards the ends ($\gamma = 0.7$) – convergence again requires analyticity outside the interval.

3.4. Example of a differentiation matrix

We consider the same situation as in Section 2.3 – the first derivative approximated on a periodic, equi-spaced grid. Instead of using trigonometric interpolation, we employ the limiting FD formula (3.1) – of infinite width, but possible to apply since periodic data can be repeated indefinitely. Assuming $N = 2M+1$ and adjusting the weights for a mesh spacing of $h = 2/N$ (rather than $h = 1$ as in (3.1)), we get

$$c^1_{\infty,j} = \begin{cases} \dfrac{N}{2}\dfrac{(-1)^{j+1}}{j} & j \neq 0 \\ 0 & j = 0. \end{cases}$$

As Figure 7 illustrates, period-wide sections of the stencil can be added together to create an equivalent stencil covering only one period of the data. Its weights become

$$d^1_{\infty,j} = \begin{cases} \dfrac{N}{2}(-1)^{j+1}\displaystyle\sum_{k=-\infty}^{\infty}\frac{(-1)^k}{j+Nk} = \frac{1}{2}\sum_{k=-\infty}^{\infty}\frac{(-1)^k}{k+(j/N)} & \begin{matrix} j=\pm1,\pm2,\ldots,\pm M \\ (\pm(M+1),\ldots,\pm(N-1)) \end{matrix} \\ 0 & j = 0. \end{cases}$$

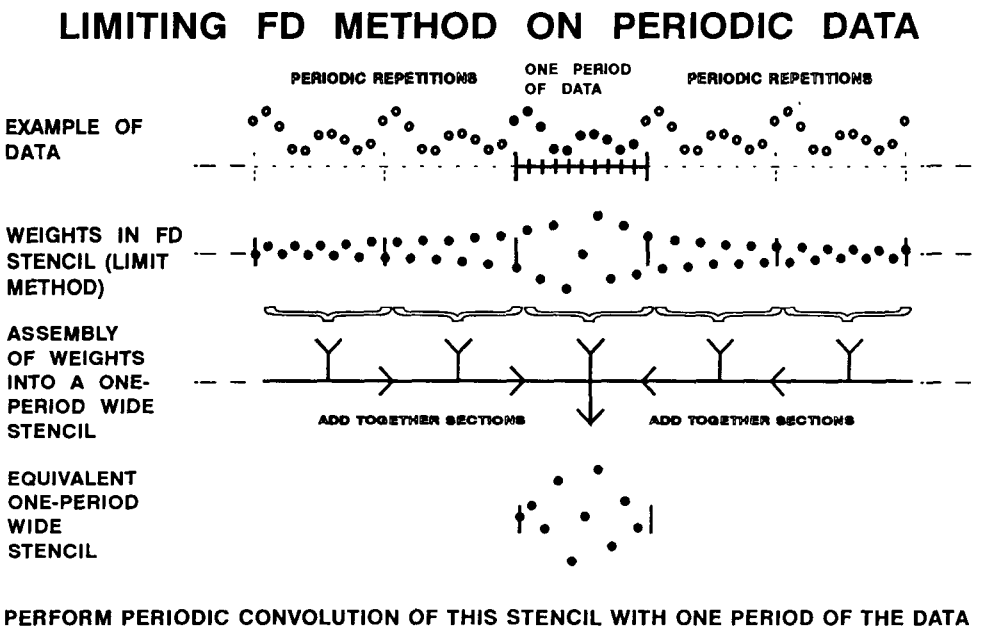


Fig. 6. Application of the limiting FD method to periodic data.

The DM is cyclic: its i, jth element is $d^1_{\infty,i-j}$. Noting the identity

$$\sum_{k=-\infty}^{\infty} \frac{(-1)^k}{k+x} = \frac{\pi}{\sin \pi x},$$

we get

$$D_{i,j} = \begin{cases} \dfrac{\pi(-1)^{i-j}}{2\sin(\pi(i-j)/N)} & i \neq j \\ 0 & i = j. \end{cases} \tag{3.3}$$

3.5. Equivalence of PS methods and limits of FD methods

Periodic case The DMs derived in Sections 2.3 and 3.4 are identical – hence the two methods are equivalent. With only little additional effort this can be shown to generalize to derivatives of any order, to even numbers of points, to 'staggered grids' (a topic discussed in Section 5.3) etc.; for details, see Fornberg (1990a).

Nonperiodic case No periodic data extensions are now available. The

order of accuracy for the approximations corresponds to the number of grid points (rather than being formally infinite). The PS method now turns out to be equivalent to using the FD approximations whose stencils extend over all the grid points. This can be seen as follows.

PS approach Consider data given at $n+1$ points on $[-1, 1]$ (distributed, for example, according to the zeros or extrema of some orthogonal polynomial – as is customary in PS collocation; however, their distribution is irrelevant for our present argument.) By means of expansion in these polynomials, the PS method provides the exact derivative of the interpolation polynomial going through the data at these points.

FD approach With no periodic data extensions, we can, at best, consider FD stencils which are as wide as the grid is wide. To approximate derivatives at the grid points, the FD weights have to be calculated separately for each point. Every one of these approximations will be exact for any nth-degree polynomial. In particular, they are all exact for the interpolating polynomial.

For any given data and distribution of (distinct) nodes, the interpolating polynomial of minimal degree is unique. Since both approaches give the exact results for this polynomial, they will always give the same results – hence, the approaches are equivalent.

4. Key properties of PS approximations

In the previous sections, we have repeatedly referred to the exponential convergence rate of spectral methods for analytic functions. This is discussed in more detail in Section 4.1. When functions are not smooth, PS theory is much less clear. An approximation can appear very good in one error norm and, at the same time, very bad in another. As illustrated in Section 4.2, PS performance can still be very impressive – this is exploited in the major PS applications. The concluding Sections 4.3–4.5 deal with implementation issues; primarily differentiation matrices and their influence on time stepping procedures.

4.1. Convergence of PS methods for smooth functions

Nonperiodic case Polynomial interpolation of smooth functions based on the Chebyshev nodes (as well as expansions in Chebyshev polynomials) are well known to provide approximations with nearly uniform accuracy over $[-1, 1]$ whereas interpolation based on equi-spaced points can diverge near the ends (the 'Runge phenomenon'). The potential functions described in Section 3.3 provide a general tool for addressing issues like these – when and with what rates convergence will occur.

In the Chebyshev case ($\gamma = 0.5$), the relationship between the potential

contours and the convergence rates becomes particularly simple. For a convergence rate α^N, $\alpha \in (0,1)$, on $x \in [-1,1]$, the nearest singularity is located on the ellipse

$$\frac{x^2}{(\frac{1}{2}(\alpha + 1/\alpha))^2} + \frac{y^2}{(\frac{1}{2}(\alpha - 1/\alpha))^2} = 1, \tag{4.1}$$

(with foci at ± 1). Conversely, given the location of the nearest singularity, this equation gives the convergence factor α. The derivation of (4.1) requires no potential considerations – it follows from the form of Lagrange's interpolation formula and requires no potential considerations – it follows from the form of Lagrange's interpolation formula and noting that

$$T_n(x) = \tfrac{1}{2}(z^n + 1/z^n),$$

where

$$\tfrac{1}{2}(z + 1/z) = x.$$

If we here consider x and z as complex variables, the ellipses (4.1) in the x-plane correspond to circles in the z-plane, centred at the origin and with radii $1/\alpha$.

The fundamental result about exponential convergence rates given in the caption to Figure 6 is valid for any node distribution functions – not just those defined through the parameter γ.

To use these analytic results to illustrate how convergence depends on the smoothness of a function $f(x)$ (f and x now real), let us consider the two-parameter class of functions

$$f_{\xi,\eta}(x) = \frac{1}{1 + ((x - \xi)/\eta)^2}.$$

Graphically, these functions have a 'hump' of unit height, centred at $x = \xi$, with widths (and radius of curvature at the tips) proportional to η. The equi-spaced PS method will be just borderline converging/diverging when the closest singularity of $f_{\xi,\eta}(x)$ (it has only two singularities – they are located at $x = \xi + i\eta$) falls on the equipotential curve passing through $x = \pm 1$ (drawn bold in Figure 6(a)). Figure 8(a) shows these most peaked functions$f_{\xi,\eta}(x)$ for different values of $\xi \in [-1,1]$. The equi-spaced PS method is clearly much better able to resolve high curvatures near the ends of $[-1,1]$ than in the interior.

This issue was also discussed by Solomonoff and Turkel (1989). They quote the closed form expression for $\phi_0(x,y)$, but select an inappropriate branch for an arctan which arises, leading to some flawed results.

However, it is also equally clear that (in this form) the equi-spaced non-periodic PS method is quite useless – no matter how many points are employed, it will not converge over $[-1,1]$ for functions any less smooth than those shown in Figure 8(a).

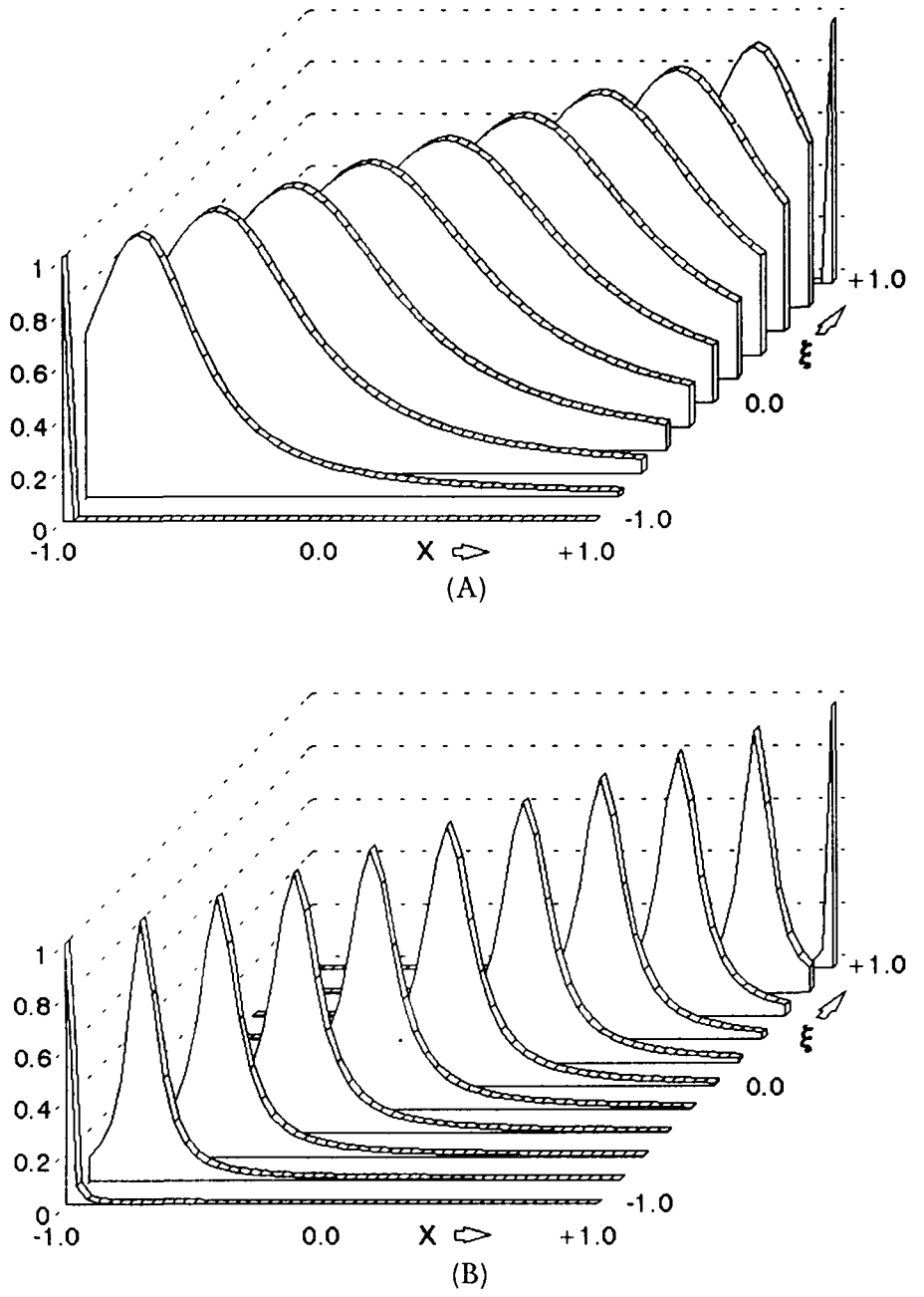


Fig. 8. Functions $f_{\xi,\eta}(x) = 1/\{1 + [(x - \xi)/\eta]^2\}$ with minimal η (i.e. maximal curvature at the tip) such that for the equi-spaced nonperiodic PS method, there is borderline convergence/divergence as $n \to \infty$ (a), and for the Chebyshev PS method there is convergence as α^n for $\alpha = 0.9$ (b).

Interpolation at the Chebyshev nodes will work for any ξ and any $\eta > 0$. However, if we require a 'reasonable' convergence rate, say $\alpha = 0.9$ (i.e. approximately a factor of 10^{-6} for every 130 node points), the situation is again somewhat similar, see Figure 8(b). Once more, the highest resolution is obtained near the boundaries. This has been exploited frequently (for example to resolve boundary layers in fluid mechanics). Note, however, that this is *not* an immediate consequence of the grid being finer there (this effect was no less prominent in the case of equi-spaced grids).

Periodic case For the periodic PS method (on $[-1, 1]$), the formula corresponding to (4.1) becomes

$$y = \pm\frac{2}{\pi}\ln\alpha \tag{4.2}$$

(related to the fact that a Fourier series converges in the widest horizontal strip around the x-axis that is free from singularities).

General discussion The relative resolution ability of different methods is sometimes expressed in the number of points needed per wavelength. For a Fourier expansion, this number is 2 (Kreiss and Oliger, 1972). For nonperiodic PS methods, it is π in the Chebyshev case (Gottlieb and Orszag, 1977) and 6 in the equi-spaced case (Weideman and Trefethen, 1988).

Figure 9 compares the curves given by (4.1) and (4.2) for $\alpha = 0.5$ and $\alpha = 0.9$. The ratio $2/\pi$ between points per wavelength for the periodic and Chebyshev PS methods follows from the ratio of the y-axis intercepts (values at $x = 0$) of these curves as $\alpha \to 1$. (Again, this is *not* a direct consequence of the fact that the Chebyshev grid happens to be $2/\pi$ as dense as the equi-spaced one at this location.)

In every instance of PS methods applied to functions analytic in some neighbourhood of $[-1, 1]$, the convergence takes the form $\mathcal{O}(\alpha^n)$ (discussed so far only for interpolation but clearly true as well – with the same α – for approximations to any derivative). This rate distinguishes spectral methods from FD and FE methods (where the rate for a pth-order method would be $\mathcal{O}(1/n^p)$; polynomial rather than exponential convergence). Whether or not there happens to be any classical family of orthogonal polynomials associated with the PS method is quite irrelevant.

We can only indicate one starting point here for analysis comparing FD against PS methods. Assume periodicity and $N + 1$ grid points spaced $h = 2/N$ apart within the period $[-1, 1]$. The range of Fourier modes $e^{i\omega x}$ that can be represented on this grid is $-\omega_{\max} \le \omega \le \omega_{\max}$, where $\omega_{\max} = \pi/h$.

Any mode ω outside $[-\omega_{\max}, \omega_{\max}]$ will, on the grid, appear equivalent to a mode within this range – an 'aliasing' error. How much is present of the different modes depends on the regularity of the function we approximate.

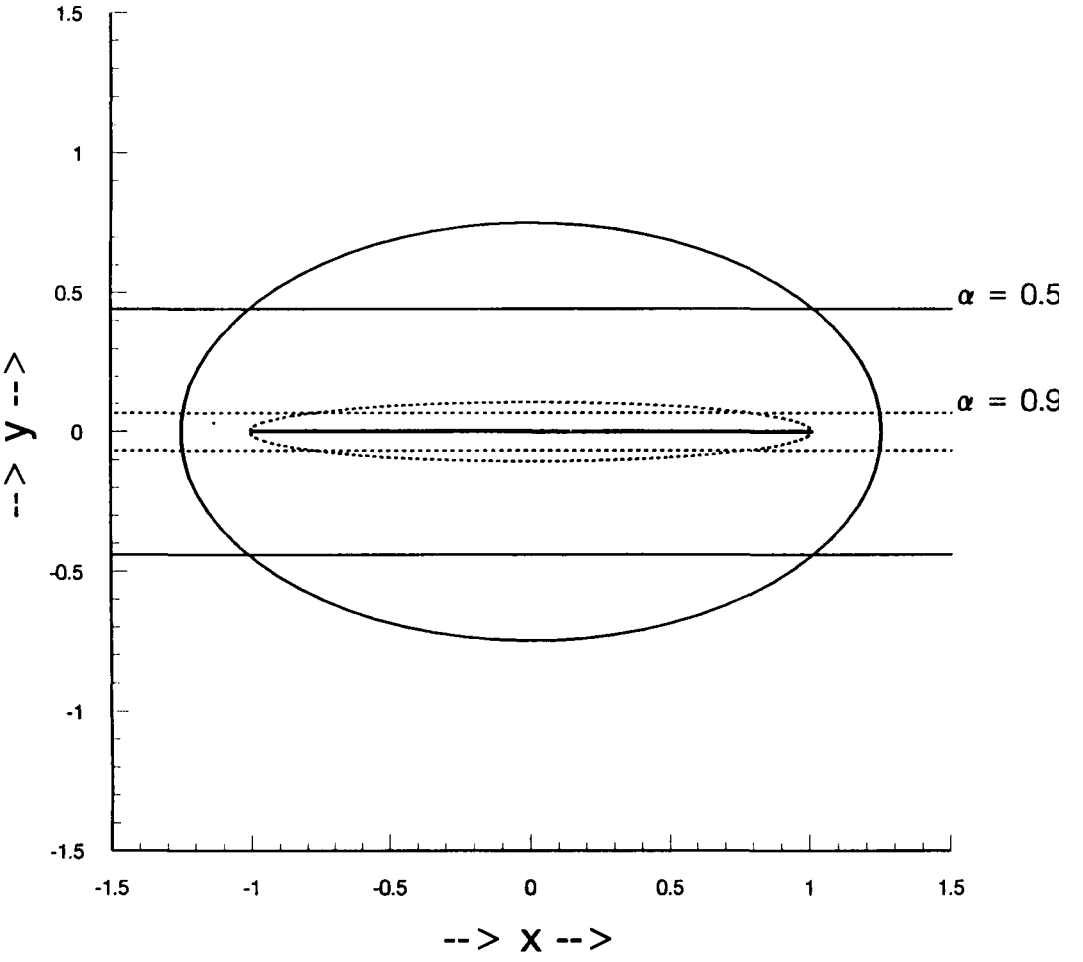


Fig. 9. Comparisons of domains in complex x-plane which need to be free of singularities to obtain convergence rates α^n, $\alpha = 0.5$ and $\alpha = 0.9$. Chebyshev method: ellipses with foci 1; equation (4.1). Periodic PS method: horizontal strips around the x-axis; equation (4.2).

Suppose we want to approximate $\mathrm{d}/\mathrm{d}x$. For a mode $\mathrm{e}^{\mathrm{i}\omega x}$, the exact answer should be

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{\mathrm{i}\omega x} = \mathrm{i}\omega\mathrm{e}^{\mathrm{i}\omega x}.$$

With FD2, we get

$$D_2\mathrm{e}^{\mathrm{i}\omega x} = \frac{\mathrm{e}^{\mathrm{i}\omega(x+h)} - \mathrm{e}^{\mathrm{i}\omega(x-h)}}{2h} = \mathrm{i}\frac{\sin\omega h}{h}\mathrm{e}^{\mathrm{i}\omega x} = \mathrm{i}f(2,\omega,h)\mathrm{e}^{\mathrm{i}\omega x}. \qquad (4.3)$$

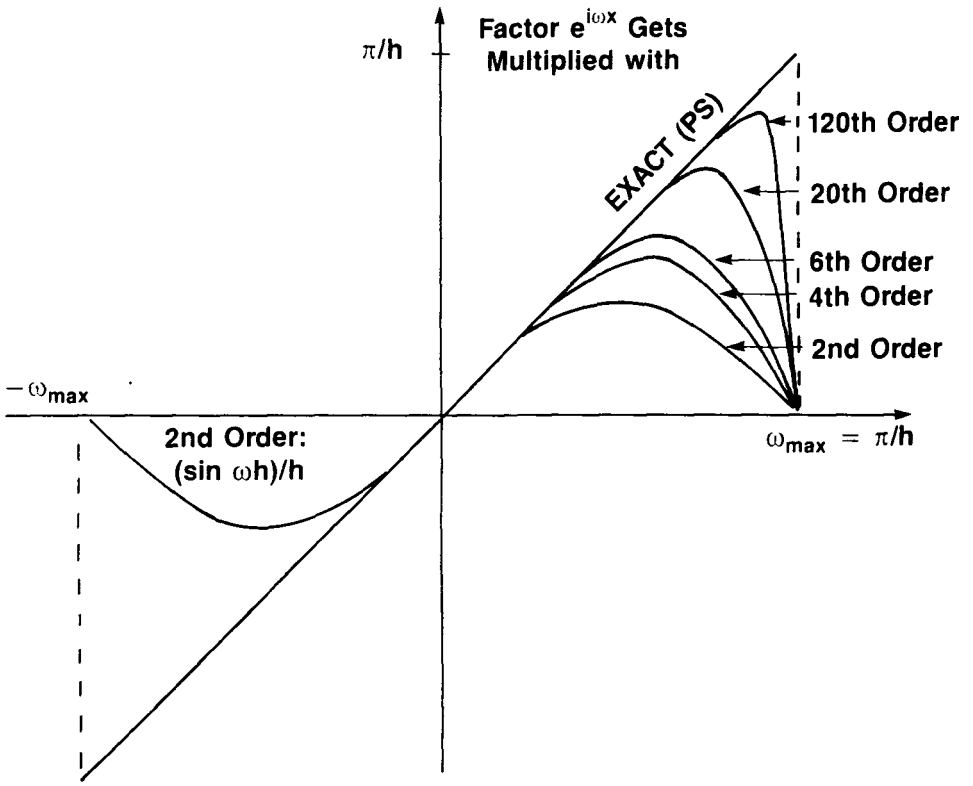


Fig. 10. Multiplicative factors $f(p,\omega,h)$ arising when the pth-order FD approximation for $\mathrm{d}/\mathrm{d}x$ is applied to $\mathrm{e}^{\mathrm{i}\omega x}$.

For the centred pth-order FD scheme ($p = 2, 4, 6, \ldots$), we get similarly

$$f(p,\omega,h) = \left\{ \frac{\sin\omega h}{h} \sum_{k=0}^{\frac{1}{2}p-1} \frac{(k!)^2}{(2k+1)!} 2^{2k} (\sin\tfrac{1}{2}\omega h)^{2k} \right\}. \tag{4.4}$$

Figure 10 compares $f(p,\omega,h)$ to the exact result ω. For $p = 2$, only a fraction of the Fourier modes present are treated even nearly correctly.

As $p \to \infty$, convergence is seen to occur as in a Taylor expansion – the number of correct derivatives at the origin is the same as the order of the FD scheme. It can make sense to give up some of the (unnecessarily high) accuracy for low ω (i.e. for long waves) in exchange for keeping the accuracy within some uniform tolerance over a wider ω range (or over a specific narrow frequency band relevant to a particular application). Such types of compact FD schemes can be very effective, for example in three-

dimensional seismic modelling, see Holberg (1987), Mittet *et al.* (1988) and Kindelan *et al.* (1990).

Solomonoff (1994) presents still another approach to generate FD schemes that are optimized in application-specific ways (i.e. rather than being designed to be exact for polynomials of as high orders as possible). He notes that such schemes can be made less vulnerable to the Runge phenomenon.

Lele (1992) displays figures similar to Figure 10, also including various compact schemes that attain high orders by means of including additional unknowns at the grid points (for example the values of derivatives as well as function values).

In the $p = \infty$ limit (the periodic PS method), the only errors are 'aliasing' errors. Due to variable coefficients and/or nonlinearities, high Fourier modes outside the range $[-\omega_{\max}, \omega_{\max}]$ are generated and then possibly mistreated. One approach to controlling such errors is to apply weak damping (dissipation). However, as we will see, anybody who views 'aliasing' only as a source of errors is missing out on one of the most important (and intriguing) strengths of PS methods.

4.2. Convergence of PS methods for nonsmooth functions

The PS method sometimes performs very well even in the cases of nonsmooth functions. Several of the major PS applications depend on this (turbulence modelling, weather forecasting, seismic modelling etc.).

As an illustration, let us consider the one-dimensional acoustic wave equation

$$\begin{cases} u_t = v_x \\ v_t = c^2(x) u_x \end{cases} \quad \text{where } c(x) = \begin{cases} 1 & -1 < x < 0, \\ \frac{1}{2} & 0 < x < 1, \end{cases} \tag{4.5}$$

periodic outside $[-1, 1]$.

We will be using grids such that 0 and ± 1 fall half-way between grid points, thus saving us from having to decide on the values of $c(x)$ at these locations.

In each of the intervals $[-1, 0]$ and $[0, 1]$, equation (4.5) supports solutions travelling to the right ($\Rightarrow$) and to the left ($\Leftarrow$) of the following forms:

$$\text{In } [-1,0] \begin{cases} u(x,t) = -v(x,t) & \Rightarrow \\ u(x,t) = +v(x,t) & \Leftarrow \end{cases} \quad \text{in } [0,1] \begin{cases} u(x,t) = -2v(x,t) & \Rightarrow \\ u(x,t) = +2v(x,t) & \Leftarrow . \end{cases}$$

We choose as initial condition

$$u(x,0) = 2v(x,0) = \exp(-1600(x - \tfrac{1}{4})^2).$$

Figure 11 shows the time evolution for $u(x,t)$ – the one for $v(x,t)$ is qualitatively similar. After the pulses have hit the interfaces at $x = 0$ and $x = 1$ numerous times – on each occasion generating two outgoing pulses (trans-

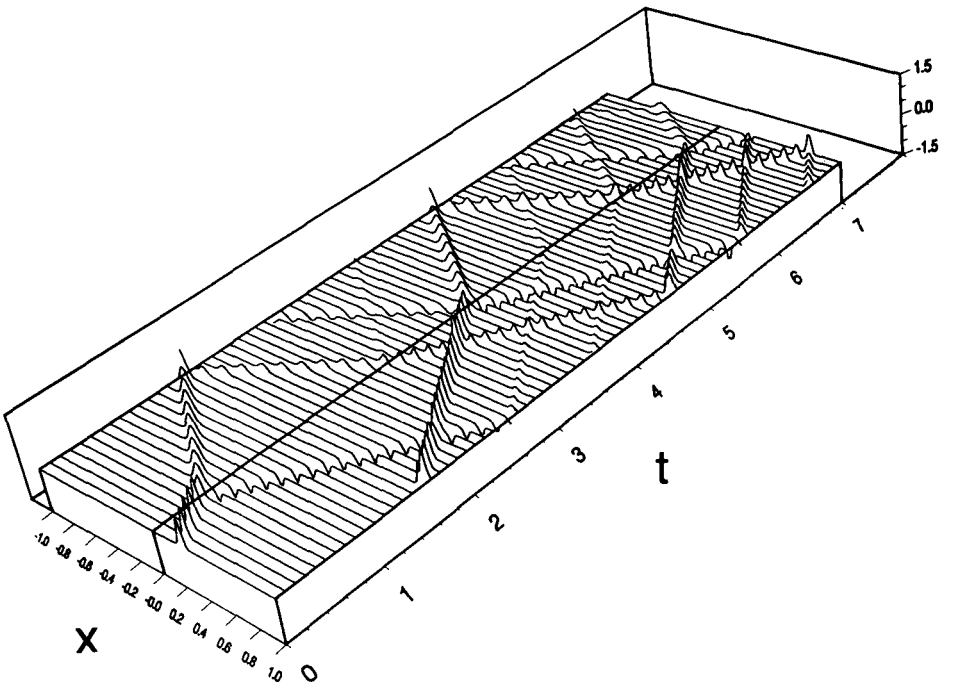


Fig. 11. Time evolution of $u(x,t)$ solving equation (4.5).

mitted and reflected) – the analytic solution at $t = 7$ consists of just three pulses.

The periodic second- and fourth-order FD and the PS methods give at $t = 7$ the results shown in Figure 12.

No numerical smoothing has been applied in any of these cases. The time integration was performed with leap-frog (centred second-order FD in time) with a sufficiently small time step that the errors which are seen are all due to the spatial discretizations. Many other ODE solvers could have been used equally well (such as Runge–Kutta, Adams Bashforth etc.).

We can note the following points.

- Already with $N = 64$, the PS method has retained considerable accuracy (in spite of the initial pulse being only about two grid points wide). For the higher values of N, the performance of the PS method is nearly flawless, and far superior to that of the FD methods.
- One might have expected the PS method to develop Gibbs-type oscillations. It is instead the FD methods which develop problems with the high modes. The generally used remedy against spurious high-frequency oscillations is to apply some viscous damping – preferably as little as possible to avoid smearing out the pulses themselves. This

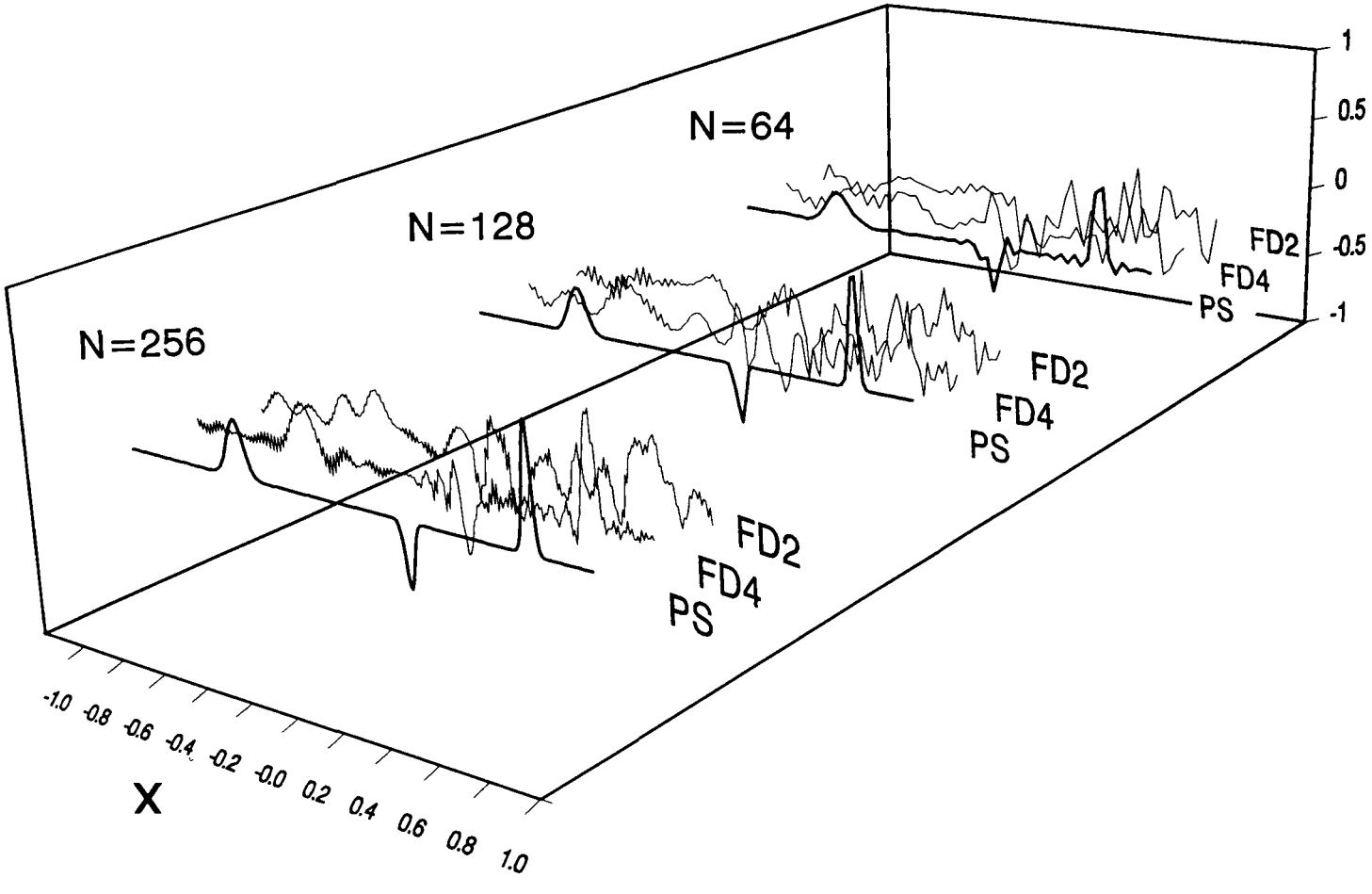


Fig. 12. Comparison of numerical solutions for $u(x,t)$ at time $t = 7$ using different methods and grid densities.

example shows that, for the PS method, often very little suffices (to use none – as in this example – is unnecessarily risky).

The difference between the methods lies not so much in the size of the local errors, as in how these accumulate or cancel over time. In the case of the PS method, relatively large errors cancel systematically.

In Fornberg (1987, 1988a), many similar tests were carried out for the two-dimensional elastic wave equation (a system of five first-order equations supporting both pressure and shear waves; see Figures 17 and 24) – with very similar results.

If Gibbs oscillations do arise in PS calculations, spectral pointwise accuracy can still sometimes be recovered in smooth regions by postprocessing. The paper 'Don't suppress the wiggles – they're trying to tell you something' (Greshko and Lee, 1981) discusses this and points out that viscous damping during a calculation (even when applied with good intent) can lead to an irretrievable loss of information. Although the convergence for step solutions is bad in most common norms (Majda *et al.*, 1978), Abarbanel *et al.* (1986) note this need not be the case for certain 'negative Sobolev norms'. Spectral accuracy of 'moments' provide information that can be used to restore Gibbs-affected solutions.

For many nonlinear equations, the discontinuties that arise are not of 'shock'-type but rather like contact discontinuities which are quite passively transported around in a linear fashion (for example the case in direct simulations of turbulence and – to a lesser extent – in weather forecasting). In such cases, it may suffice to add very little viscosity and rely on the method's ability to handle linear situations.

The problem with more severe nonlinearities is primarily that they can introduce couplings, disrupting the delicate cancellation process on which the PS method for nonsmooth functions is dependent. One idea is to add some more viscosity (just enough to be able to exploit the PS method's power in the case of smooth solutions but not so much that the solution itself gets too severely affected). Spectrally accurate solutions can sometimes be obtained in this way even for shock problems. Different versions of this idea have been proposed. One is the 'Spectral Viscosity Method' (Tadmor, 1989). Further discussions on this and similar methods can be found in Cai *et al.* (1989, 1992), Tadmor (1990, 1993) and Maday *et al.* (1993).

4.3. Differentiation matrices

For both computational and theoretical purposes, it is often convenient to collect all the weights for the approximations at the grid points in a 'differentiation matrix' (DM; cf. Sections 2.3 and 3.4). Finding the derivatives of a vector of data becomes a matrix × vector multiplication.

The relative efficiencies of straightforward matrix × vector multiplications

($\mathcal{O}(n^2)$ operations) versus FFT-based Chebyshev recursions ($\mathcal{O}(n \log n)$ operations) have been compared many times. The estimates for the point of break-even ranges at least from $n = 16$ (Canuto *et al.*, 1988) to n around 100 (Taylor *et al.*, 1984) – the point can be higher still for vector and parallel machines. Furthermore, Solomonoff (1992) shows how a restructuring of the matrix × vector multiplication for DMs can nearly double the speed of this approach.

The 'Fast Multipole Method' achieves an $\mathcal{O}(n \log n)$ operation count for arbitrary node distributions (Boyd, 1992). However, the proportionality constant is *much* higher than for FFTs – the approach appears not to be competitive in the present context.

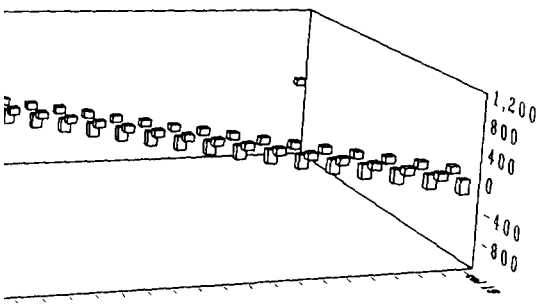
Beylkin *et al.* (1991) describe a wavelet approach for converting between finite Legendre and Chebyshev expansions – but again, with a large constant in the $\mathcal{O}(n \log n)$.

Figures 13(a)–(d) illustrate what the DMs look like for the second derivative in the case of $n = 20$ (i.e. 21 grid points if both ends are included, 20 within the period for periodic problems). Figure 13(a) shows the (periodic) stencil $[1 \ -2 \ 1]/h^2$ and (b) the periodic PS matrix; (c) and (d) show the nonperiodic equi-spaced and Chebyshev matrices ($\gamma = 0$ and 0.5 respectively). Large elements are seen in the top and bottom rows (corresponding to approximations near the boundaries).

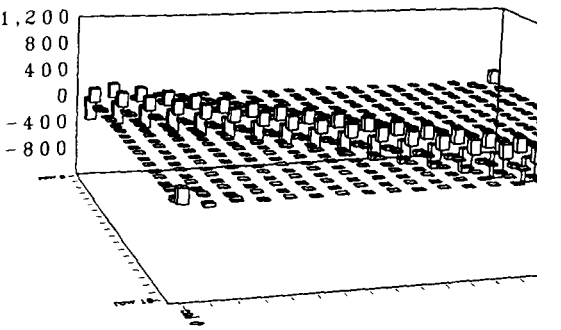
For nonperiodic problems, the algorithm in Section 3.1 (code in Appendix B) can be used to generate DMs very conveniently. In the PS case:

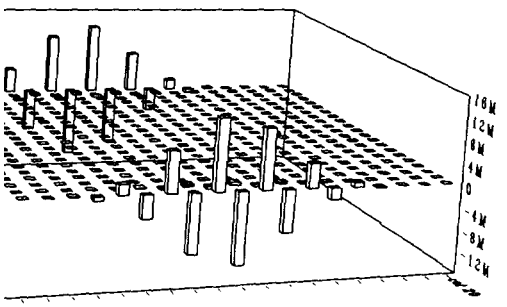
```
                                                    Specify size of grid
      PARAMETER (N=...  , M=...  )                  and highest derivative
      DIMENSION X(0:N),C(0:N,0:N,0:M),DM(0:N,0:N,M)
      DO 10 I=0,N
10      X(I) = ....                                 Specify the grid points
        DO 20 I=0,N
          CALL WEIGHTS (X(I),X(0),N,M,C)
          DO 20 L=1,M
            DO 20 J=0,N
20            DM(I,J,L) = C(J,N,L)                  DM(*,*,L)
                                                    contains now the DMs
                                                    for the Lth deriv.,
.....                                               L=1,2,...,M.
```

The computer time taken generating DMs is seldom critical. However, if this has to be done many times, the code above should not be used (since it fails to exploit the fact that all the separate calls to WEIGHTS are based on the same grid – some intermediate quantities need not be recalculated repeatedly).

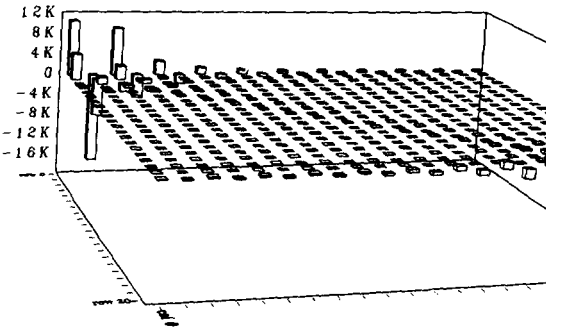


a





c



d

Differentiation matrices for approximations to d^2/dx^2, $n = 20$ (in parentheses, maximum magnitud
ent): (a) second-order FD, periodic (200.); (b) PS, periodic (331.); (c) PS, non-periodic, equi-spa
$\times 10^6$); (d) PS, non-periodic, Chebyshev (17×10^3).

With

$$a_k = \prod_{\substack{i=0 \\ i\neq k}}^{n} (x_k - x_i), \quad F_k(x) = \frac{1}{a^k} \prod_{\substack{i=0 \\ i\neq k}}^{n} (x - x_i),$$

Lagrange's interpolation formula becomes

$$p_n(x) = \sum_{k=0}^{n} f(x_k) F_k(x).$$

Relatively straightforward manipulation of this (Nielsen, 1956, pp. 150–154) allows the elements of D^1 to be computed in, to leading order, only four operations per element:

$$D^1_{jk} = \begin{cases} \dfrac{a_j}{a_k(x_j - x_k)} & j \neq k \\ \displaystyle\sum_{\substack{i=0 \\ i\neq k}}^{n} \frac{1}{x_k - x_i} & j = k. \end{cases}$$

DMs for higher derivatives can, in this case, be obtained as matrix powers of D^1. A much less costly recursion is also available – D^p can be obtained from D^{p-1} in only five operations per element ($p = 2, 3, \ldots$) (Huang and Sloan, 1993, Welfert, 1993). Welfert also notes:

- The PS literature contains many instances of authors assuming the relation $D^p = (D^1)^p$ when it does not hold (for example it fails for the periodic PS method if the number of points is even). A sufficient condition for this relation is presented.
- Closed form expressions for D^1_{jk} and D^2_{jk} become particularly simple for many cases of orthogonal polynomials. A comprehensive list has been collected.

Rounding error propagations within different methods for calculating Chebyshev DMs are discussed by Breuer and Everson (1992).

4.4. Eigenvalues of differentiation matrices

A major difficulty with nonperiodic PS methods is that their DMs tend to have very large spurious eigenvalues (in addition to their physically relevant ones). This adversely affects time stepping techniques (to be discussed in Section 4.5). Many of the special techniques in Section 5 are designed to (partially) overcome this. In order to provide a background for these discussions, we will describe some typical eigenvalue (EV) distributions.

Example 1 Periodic PS; advection equation $u_t = u_x$.

The DM is derived explicitly both in Sections 2.3 and 3.4 (cf. equation (3.3)). The DM is anti-symmetric; its eigenvalues are equi-spaced on the

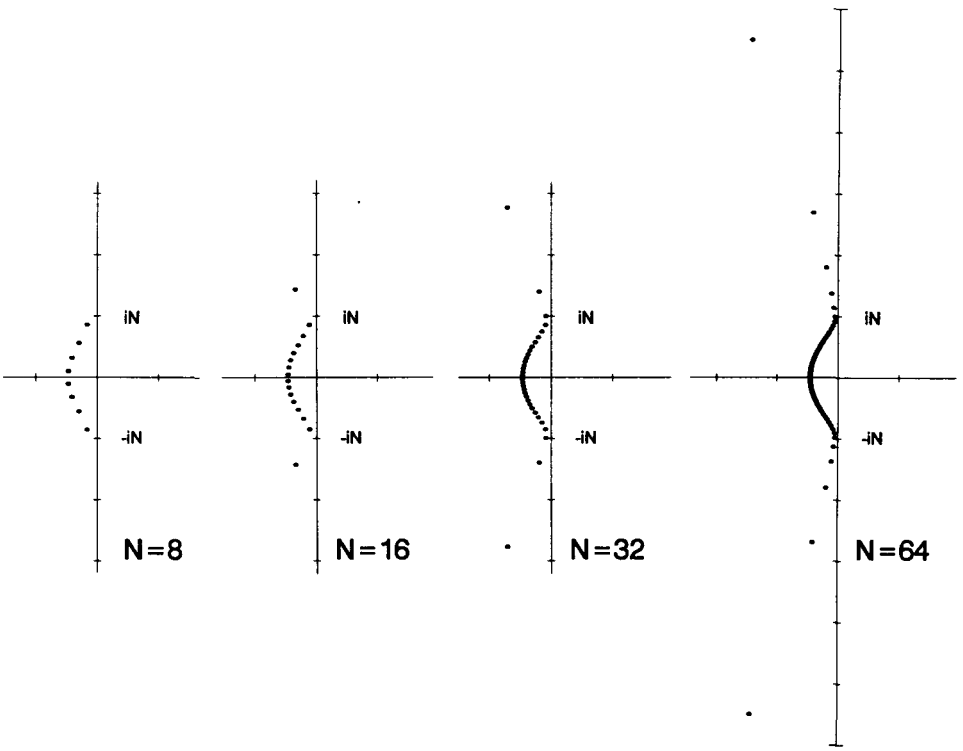


Fig. 14. Eigenvalues of Chebyshev DMs for an advection equation.

imaginary axis between $-N(\pi/2)\,\mathrm{i}$ and $+N(\pi/2)\,\mathrm{i}$ (when N is odd – very minor differences for N even).

Example 2 Chebyshev PS; advection equation $u_t = u_x$, $u(1) = 0$.

Figure 14 shows the EVs for $N = 8$, 16, 32 and 64. Although most of the EVs converge to a curve in the left half-plane between $-\mathrm{i}N$ and $+\mathrm{i}N$, a few spurious 'outliers' diverge at rates proportional to N.

Trefethen and Trummer (1987) note that the small (physical) eigenvalues for large values of N exhibit a very large sensitivity to rounding errors (however, still leaving them distributed along very distinct paths).

When the grid points are instead distributed as the zeros of Legendre polynomials $L_N(x)$, Dubiner (1987) noted that the spurious outliers were absent in this model problem. However, the EV sensitivity remains large (and EVs alone fail to fully describe stability issues as the DMs are highly nonnormal matrices). It is questionable whether use of Legendre polynomials offers any practical advantage (Trefethen, 1988).

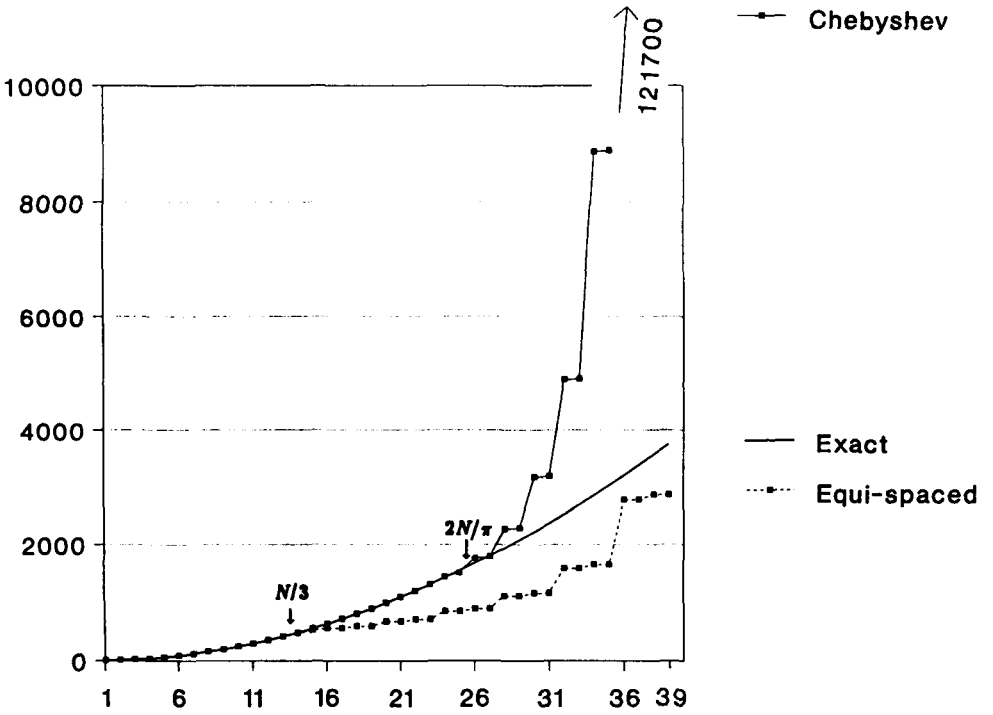


Fig. 15. Magnitudes of eigenvalues of PS DMs (for equi-spaced and Chebyshev grids) compared to analytic eigenvalues; $N = 40$.

Example 3 Chebyshev and equi-spaced nonperiodic PS; heat equation $u_t = u_{xx}$, $u(\pm 1) = 0$.

The continuous problem $u_{xx} = \lambda x$, $u(\pm 1) = 0$ has the EV $\lambda_k = -(k\pi/2)^2$, $k = 1, 2, \ldots$. The EVs of the Chebyshev DM for u_{xx} are all real and negative (Gottlieb and Lustman, 1983). In the case of $N = 20$, this DM is the matrix in Figure 13(d) with the first and last rows and columns removed. Figure 15 compares the magnitude of the EVs with those for the equi-spaced nonperiodic PS method (cf. Figure 13(c); in this case, many higher EVs are complex) and the exact ones. The portions $2/\pi$ and $1/3$ of the EVs are spectrally accurate in the two cases (cf. the numbers of points per wavelength 2, 6 and π mentioned in Section 4.1).

PS methods can also be devised for infinite domains. For eigenvalues of 'Hermite' and 'rational spectral' PS DMs, see Weideman (1992). Other such cases include Laguerre eigenvalues (Funaro, 1992) and sinc eigenvalues (Stenger, 1981).

4.5. Time-stepping methods and stability conditions

Stability (meaning that the numerical solution remains bounded up to a fixed time T as time and space steps Δt and $\Delta x \to 0$) is essential because of the Lax Equivalence Theorem which can be stated as follows.

Lax Equivalence Theorem For a well-posed linear problem, a consistent approximation converges if and only if it is stable.

Stability analysis for spectral methods is simple in only one case – the Fourier PS method for a periodic constant coefficient problem. The DMs are cyclic with known eigenvalues. Assuming again for simplicity $n = 2m+1$, the eigenfunctions (in one dimension) are $e^{i\omega x}$, $\omega = -m, \ldots, m$ with eigenvalues $\partial/\partial x \leftrightarrow i\pi\omega$, $\partial^2/\partial x^2 \leftrightarrow -\pi^2\omega^2$ etc., $(\omega = -m, \ldots, m)$. Had we instead used second-order FD in space, we would have obtained $\partial/\partial x \leftrightarrow i\sin(\pi\omega\Delta x)/\Delta x$ (cf. (4.3)), $\partial^2/\partial x^2 \leftrightarrow -4(\sin(\pi\omega\Delta x))^2/\Delta x^2$ etc. The stability restrictions when time stepping (of the forms $\Delta t/\Delta x <$ constant and $\Delta t/\Delta x^2 <$ constant respectively) therefore have constants $1/\pi$ and $4/\pi^2$ times those that arise with second-order FD methods (i.e. they are hardly any more severe).

For higher-order FD methods, similar ratios can be read from the maximum values of the curves in Figure 10 and their generalization to higher derivatives (the equivalents of (4.4) are given in Fornberg, 1990a).

Stability conditions like $\Delta t/\Delta x <$ constant and $\Delta t/\Delta x^2 <$ constant are normally not restrictive in connection with spectral methods. With, say, a fourth-order Runge–Kutta method in time and a better-than-eighth-order PS method in space, $\Delta t/\Delta x^2 <$ constant is needed anyway to make the temporal accuracy match the spatial one. However, conditions on $\Delta t/\Delta x^p$, $p > 2$ will arise in many nonperiodic PS cases. One of the main issues in designing (nonperiodic) spectral methods is to circumvent these.

For more realistic problems, several complications arise (both for FD and for PS methods):

- For variable coefficients or nonlinearities, stability for all problems with 'frozen' coefficients is neither necessary nor sufficient for stability (Kreiss, 1962, Richtmyer and Morton, 1967).
- With boundaries present, local mode analysis for the interior needs to be complemented by 'GKS' analysis at the boundaries (Kreiss, 1968, 1970, Gustafson, Kreiss and Sundström, 1972); for simplified versions of this, see Trefethen (1983), Goldberg and Tadmor (1985)).

Additional problems are more specific to PS methods:

- One needs to distinguish between 'Lax stability' (fixed T and $\Delta t \to 0$) and 'eigenvalue stability' (fixed Δx and Δt as $T \to \infty$). For the highly nonnormal DMs that arise from nonperiodic PS methods, large growths

can initially arise if the norms are large even if all eigenvalues fall within (or on) the unit circle (cf. the Kreiss matrix theorem, in Richtmyer and Morton (1967)).

- PS methods can be unstable even when the corresponding FD methods of increasing orders are *all* stable. Tadmor (1987) addresses this phenomenon in connection with a linear model equation $u_t = c(x)u_x$. Reddy and Trefethen (1990) use 'pseudospectra' to provide further insight into this and similar phenomena.

'Energy methods' provide a powerful general tool for PS analysis, e.g. Gottlieb *et al.* (1980, 1987, 1991). However, due to their technical complexity, we restrict the discussion here to eigenvalue stability. Although limited, it can still provide useful guidelines for selecting time integrators.

The most common time-stepping approach is the 'Method of Lines' (MOL) which amounts to discretizing in space only and then applying a 'packaged' ODE solver (based, for example, on Runge–Kutta or backwards differentiation) to the resulting system of ODEs.

This approach allows the ODE solver to be developed and analysed separately from the spatial discretization method. The user need not be concerned with many tedious issues like starting techniques for multi-step methods, time step and order adjustments etc.

For constant coefficients, the MOL gives rise to a system of ODEs

$$\begin{bmatrix} u \end{bmatrix}_t = \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} u \end{bmatrix} + \begin{bmatrix} f \end{bmatrix},$$

where A is the differentiation matrix for the approximation (which we now assume to be diagonalizable). For a first cut at assessing what kind of ODE package to select, the stability regions of the time integrator have to be compared with the eigenvalues of A.

Stability regions An ODE solver is *stable* for Δt and (complex) λ if the numerical solution to $u_t = \lambda u$ does not grow with t. It is called *A-stable* if it is stable for all λ in the negative half-plane (the ideal situation – this matches the same property of the analytic solution $u(t) = \mathrm{e}^{\lambda t}$). A-stability can seldom be achieved for methods of high accuracies. (cf. the 'Dahlquist barriers' (Dahlquist, 1956, 1985)). Instead, for most methods we find that $\lambda\Delta t$ must lie within some smaller domain than the full left half-plane. If any eigenvalues λ of A happen to be large in magnitude, this restricts Δt.

Example Forward Euler:

$$v(t + \Delta t) = v(t) + \Delta t \lambda v(t) \Rightarrow$$
$$v(t + k\Delta t) = (1 + \lambda\Delta t)^k v(t) \Rightarrow$$

Stable if $|1 + \lambda\Delta t| \leq 1 \Rightarrow$ stability region (values of $\lambda\Delta t$ giving no growth) is a circle with radius 1 centred at -1.

However, very low accuracy and no stability coverage along the imaginary axis makes this method very unattractive.

Commonly used ODE solvers represent compromises between low operation counts, high accuracies and large stability domains. They include many Runge–Kutta (RK) schemes, Adams-type methods and, for 'stiff' problems (with some eigenvalues far away in the left half-plane) BDF methods. For discussions on ODE solvers, see, for example, Gear (1971), Shampine and Gordon (1975), Lambert (1991), Hairer *et al.* (1987), Hairer and Wanner (1991). Many stability domains are illustrated in Sand and Østerby (1979).

5. PS variations/enhancements

Up to this point, we have described 'basic' PS implementations. However, many variations are possible, offering advantages in different respects. In this section, we discuss a few of these.

5.1. Use of additional information from the governing equations

This idea (like most others) is best described through the use of examples. To use it the problem must first be manipulated analytically (for example by repeated differentiation) to provide more information than is immediately available from its original formulation.

Example 1 Exploiting additional derivative information at the boundaries for the eigenvalue problem $u_{xx} = \lambda u$, $u(\pm 1) = 0$.

Since $u(\pm 1) = 0$, clearly also $u''(\pm 1) = 0$ and $u''''(\pm 1) = 0$ (for this example, we ignore that this pattern continues indefinitely and that u becomes periodic). The information on u'' and u'''' can be exploited in different ways:

A: Reduce the largest spurious EVs (cf. Figure 15).
Each 'extra' boundary condition has a corresponding (one-sided) difference stencil. From each row of the DM (as shown in Figure 13(c) and (d)), we can subtract any multiple of these stencils. These multiples can be choosen to minimize the sum of the squares of the elements of the resulting DM. As shown in Fornberg (1990b), this procedure much improves the conditioning of the DM without any degradation in its spectral accuracy:

$\Rightarrow$ Using only $u''(\pm 1) = 0$ and $u''''(\pm) = 0$ reduces the largest matrix elements of the DMs shown in Figures 13(c) and (d) (edge elements not included) to about 500 – down from about 500 000 and 7000 respectively!).

$\Rightarrow$ Each time an 'extra' boundary condition is applied to a Chebyshev-type approximation, the largest (remaining) spurious eigenvalue gets changed to *exactly* zero.

B: Increase the accuracy of the computed EVs.
Given k extra relations (for example information at the boundaries), we can proceed as follows:

1 Introduce k new fictitious grid points (anywhere – inside or outside the domain).
2 Find the $N+1$ stencils for u_{xx} which are accurate at the original grid point locations $x_0, x_1, \ldots, x_N$ respectively, but which extend also over the fictitious points.
3 Find the k stencils which express the k extra pieces of information (again extending over all grid points – original and fictitious).
4 Add/subtract multiples of these k stencils from the ones calculated in Step 2 – so that the weights at all the fictitious points are eliminated.

The resulting stencils for u_{xx} become exact for all polynomials of degree $N+k$ which take prescribed values at $x_0, x_1, \ldots, x_N$ *and* satisfy the k extra relations. They are therefore more accurate, but still no more costly to apply than straightforward PS approximations also extending over $x_0, x_1, \ldots, x_N$ (these would be exact only for polynomials up to degree N).

The idea of introducing points outside a boundary and then eliminating them again using boundary conditions is often used with FD methods. The PS case is remarkable in that the location of these (temporary) points turns out to have no influence at all on the final result (apart from rounding errors). In this PS case, they offer a very convenient way of generating stencils satisfying 'side conditions' without the need for any additional analytical devices.

Another boundary situation for which large improvements can be achieved is the treatment of the origin in polar coordinates. This is discussed further in Section 5.6. The example below (from Fornberg (1994)) demonstrates this idea in a case of axial symmetry.

Example 2 Exploiting symmetry at a boundary.

Bessel's equation arises from Poisson's equation in the case of axial symmetry. Consider its eigenvalue problem

$$u'' + \frac{1}{r}u' - \frac{n^2}{r^2}u = -\lambda u, \quad n = 0, 1, \ldots, \ u(0) \text{ bounded}, \ u(1) = 0.$$

The exact eigenvalues $\lambda_{n,k}$, $k = 1, 2, \ldots$ satisfy $J_n(\sqrt{\lambda_{n,k}}) = 0$.

We compare two approximation methods:

1 Note that

$$\begin{cases} n = 0 & u'(0) = 0 \\ n \neq 0 & u(0) = 0 \end{cases}$$

and use straightforward Chebyshev approximation on [0,1]. Grid points are located at $r_k = (1 - \cos k\pi/N)/2$, $k = 0, 1, \ldots, N$.

2 Note that

$$\begin{cases} n \text{ even} & u(r) \text{ even} \\ n \text{ odd} & u(r) \text{ odd.} \end{cases}$$

Consider FD stencils extending over $[-1, 1]$, but use symmetry to reduce actual calculations to within [0,1]. Grid points are located at $r_k = \sin k\pi/2N$, $k = 0, 1, \ldots, N$ (i.e. no clustering at $r = 0$).

Figure 16 compares the accuracies in the numerical EVs obtained through these two methods for $n = 7$, $k = 1$ ($\lambda_{7,1} \approx 122.9$; cf. Gottlieb and Orszag (1977, pp. 152–153)). The values for method 1 are taken from Huang and Sloan (1993a).

Further variations of boundary implementations are discussed by Canuto and Quarteroni (1987) and Funaro and Gottlieb (1988).

Morals When using PS methods, always consider whether the FD viewpoint can offer any advantages (in accuracy, simplicity, flexibility, etc.)

The fundamental reason for clustering grid points at the ends of an interval is to compensate for the large error terms in one-sided approximations. The more information we can exploit at boundaries, the less we should cluster.

5.2. Use of different PS approximations for different terms in an equation

When a single variable appears more than once in an equation, it is normally approximated in a similar way at each instance. However, Huang and Sloan (1993b,c) note two situations when it is better to use different types of PS approximation.

Example 1 Solve the singular perturbation problem

$$\epsilon u'' + u' = 1, \quad x \in [-1, 1], \quad 1 \gg \epsilon > 0.$$

Straightforward centred second-order FD approximations for both u' and u'' give an oscillatory solution with $\mathcal{O}(1)$ errors across $[-1, 1]$ for any N when $\epsilon < 1/N$. Approximating u_x by the one-sided FD stencil $[u(x+h) - u(x)]/h$ reduces the errors to $\mathcal{O}(1/N)$. When using Chebyshev PS approximations, we can similarly approximate u' with stencils based on all grid points *but* the one at $x = -1$. In the limit of $\epsilon \to 0$, this gives spectral accuracy across $[-1, 1]$ (rather than $\mathcal{O}(1)$ errors).

For another idea to solve this problem with a PS method, see Eisen and Heinrichs (1992).

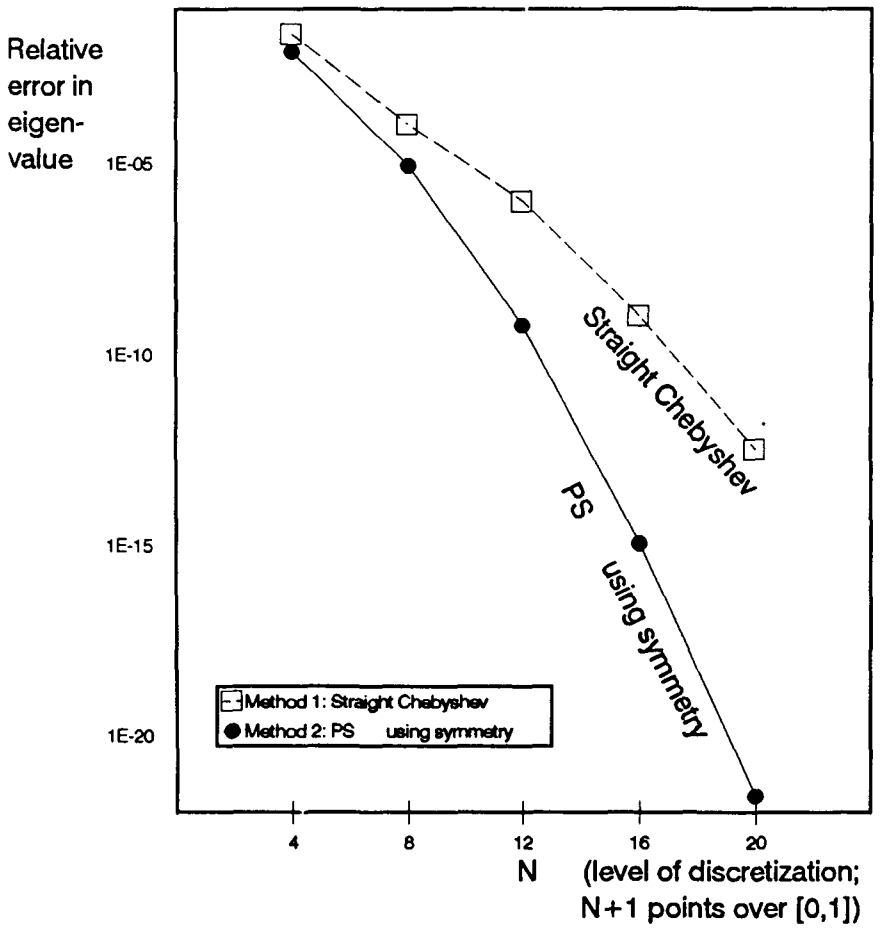


Fig. 16. Errors in eigenvalue $\lambda_{7,1}$ for Bessel's equation when approximated using two different strategies.

Example 2 Solve the eigenvalue problem $u'''' + 4u''' = \lambda u''$, $x \in [-1, 1]$, $u(\pm 1) = u'(\pm 1) = 0$.

Problems similar to this arise, for example, in linearized stability analysis in fluid mechanics. Spurious EVs denote in this case EVs appearing incorrectly in the right half-plane, suggesting physical instabilities that do not exist.

If we approximate all derivatives of u on a Chebyshev grid, incorporating $u'(\pm 1) = 0$ as in Example 1, Part B, Section 5.1, we will get spurious EVs. However, ignoring $u'(\pm 1) = 0$ when approximating u'' overcomes this.

In both these examples, variable coefficients would have added no complications (as is usually the case for PS methods – in sharp contrast to spectral Galerkin or Tau methods).

5.3. Staggered grids

When using an FD method, it is customary to compute values for each unknown at each grid point. Figure 17 illustrates an alternative, which can be employed quite frequently. Even for equations with only one unknown variable, a similar staggering can sometimes be used effectively from time level to time level. The idea is to gain accuracy – derivative approximations at 'half-way' points are often much more accurate than at grid points.

For the first derivative:

		Second-order accuracy	Leading error terms	Ratio of error
Reg.	$f'(x) =$	$\{\frac{1}{2}f(x-h)$		
		$+\frac{1}{2}f(x+h)\}/h$	$+\frac{1}{6}h^2 f'''(x)\ldots$	
Stag.	$f'(x) =$	$\{f(x-\frac{1}{2}h)$		
		$+\frac{1}{2}f(x+\frac{1}{2}h)\}/h$	$+\frac{1}{24}h^2 f'''(x)\ldots$	$\frac{1}{4} = 0.25$
		Fourth-order accuracy		
Reg.	$f'(x) =$	$\{\frac{1}{12}f(x-2h)$		
		$-\frac{2}{3}f(x-h)$		
		$+\frac{2}{3}f(x+h)$		
		$-\frac{1}{12}f(x+2h)\}/h$	$-\frac{1}{30}h^4 f^v(x)+\ldots$	
Stag.	$f'(x) =$	$\{\frac{1}{24}f(x-\frac{3}{2}h)$		
		$-\frac{9}{8}f(x-\frac{1}{2}h)$		
		$+\frac{9}{8}f(x+\frac{1}{2}h)$		
		$-\frac{1}{24}f(x+\frac{3}{2}h)\}/h$	$-\frac{3}{640}h^4 f^v(x)+\ldots$	$\frac{9}{64} \approx 0.141$

For approximations of order p, the ratio of error terms turns out to be

$$\left\{\frac{p!}{2^p\{(\frac{1}{2}p)!\}^2}\right\}^2 \approx \frac{2}{\pi p}.$$

Since the periodic PS method can be viewed as the limit of $p \to \infty$, this suggests that the idea of staggering would also be advantageous in that case.

Another suggestive argument follows from comparing the weights in the stencils. Figure 18(a) shows the magnitudes of the weights for increasingly accurate approximations to the first derivative as is also displayed in the right half of Figure 2. In the limit, they become $(-1)^\nu/\nu$, $\nu = 1, 2, \ldots$. For the staggered approximations, the limit is much more local in nature: $(-1)^{\nu+1/2}/\pi\nu^2$, $\nu = \frac{1}{2}, \frac{3}{2}, \frac{5}{3}, \ldots$. Since the derivative is a local property of a function, a more compact approximation makes more sense than one relying on extensive cancellation of distant contributions.

For nonperiodic problems, staggering can be achieved, for example, by using grids based on Chebyshev extrema (as usual) and Chebyshev zeros.

Staggering turns out to be advantageous for odd derivatives (first, third,

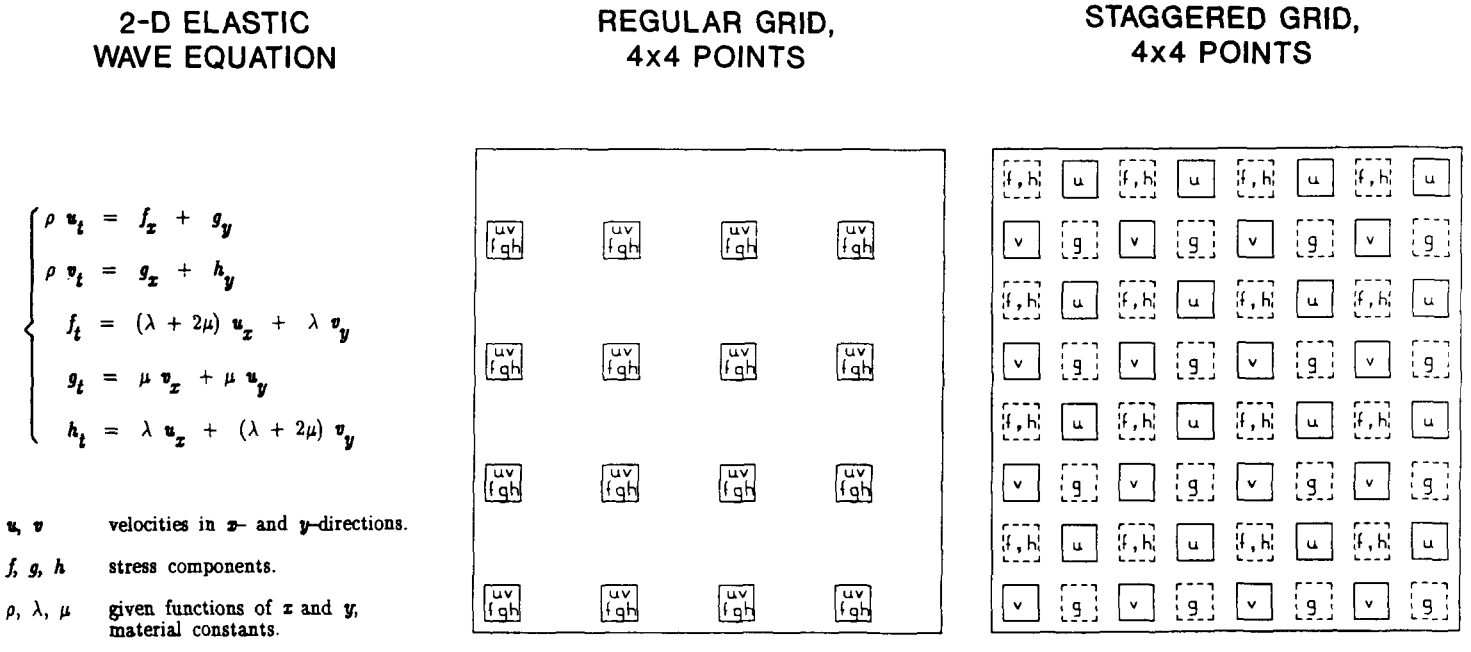


Fig. 17. Example of staggered grid arrangement.

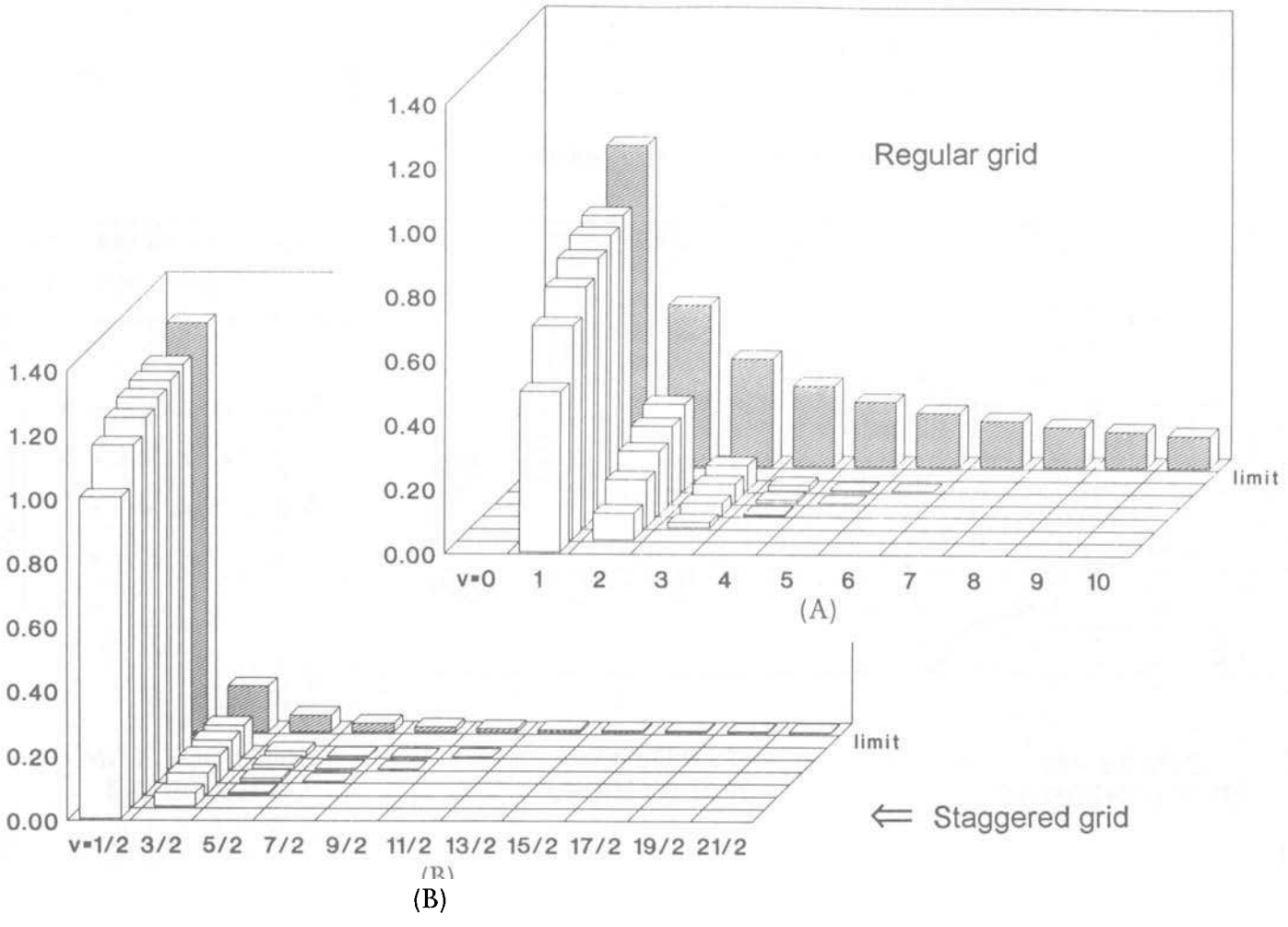


Fig. 18. Magnitudes of weights for increasingly accurate approximations to the first derivative on regular and staggered grids (right halves of stencils displayed).

etc.) whereas regular grids are better for even derivatives. For analysis and references on grid staggering in connection with high-order methods, see Fornberg (1990a).

5.4. Preconditioning

The large spurious eigenvalues in many PS DMs can make explicit time stepping methods very costly (forcing the use of extremely small time steps). Implicit methods often have unbounded stability domains, but they require the solution of a full linear system every time step.

The idea of preconditioning works in any number of dimensions, but is easiest to illustrate in one dimension.

Example Preconditioning for a Chebyshev PS solution of a two-point boundary value problem $u''(x) = f$, where $u(\pm 1)$ and f are given.

Chebyshev PS discretization (viewed as a FD method) gives rise to a linear system $Cu = f$, where C is the Chebyshev DM for $\mathrm{d}^2/\mathrm{d}x^2$ (as shown in Figure 13(d), but with the edge rows and columns removed). C is neither symmetric nor diagonally dominant – standard iterative techniques will not converge.

Let F be the second-order FD DM based on the same Chebyshev grid (with elements obtained using the algorithm in Section 3.1). Since F is tri-diagonal, it is easily inverted (in higher dimensions, one can, for example, use alternating direction arrangements for the tri-diagonal matrices). The system to be solved can be written as $[F^{-1}C]u = g$, where $g = F^{-1}f$. The matrix $F^{-1}C$ is illustrated in Figure 19. Any standard iterative technique can be used to rapidly obtain the (spectrally accurate) vector u.

For matrices of this form (near-symmetric, diagonally dominant), convergence of some methods improve if the ratio of largest-to-smallest eigenvalues is lowered. For $F^{-1}C$, $\lambda_{\max}/\lambda_{\min} \to \frac{1}{4}\pi^2$ as $N \to \infty$ (Haldenwang *et al.*, 1984). Using higher-order FD methods, this ratio can be lower still (Phillips *et al.*, 1986) but savings may be off-set by a higher cost of applying F^{-1}. In contrast, we can note that for the (nonsymmetric, nondiagonally dominant) matrix C, $\lambda_{\max}/\lambda_{\min}$ grows like $\mathcal{O}(N^4)$.

For odd derivatives, FD preconditioning normally requires the use of staggered grids. This is discussed for Chebyshev methods by Hussaini and Zang (1984) and Funaro (1987). Mulholland and Sloan (1992) considers FD preconditioners for staggered approximations to $\partial^3/\partial x^3$ (applicable, for example, to the Korteweg–de Vries equation).

FE-based preconditioners have been discussed by Canuto and Quarteroni (1985), Deville and Mund (1985) and Canuto and Pietra (1987). General references on preconditioning include Canuto *et al.* (1988) and Boyd (1989).

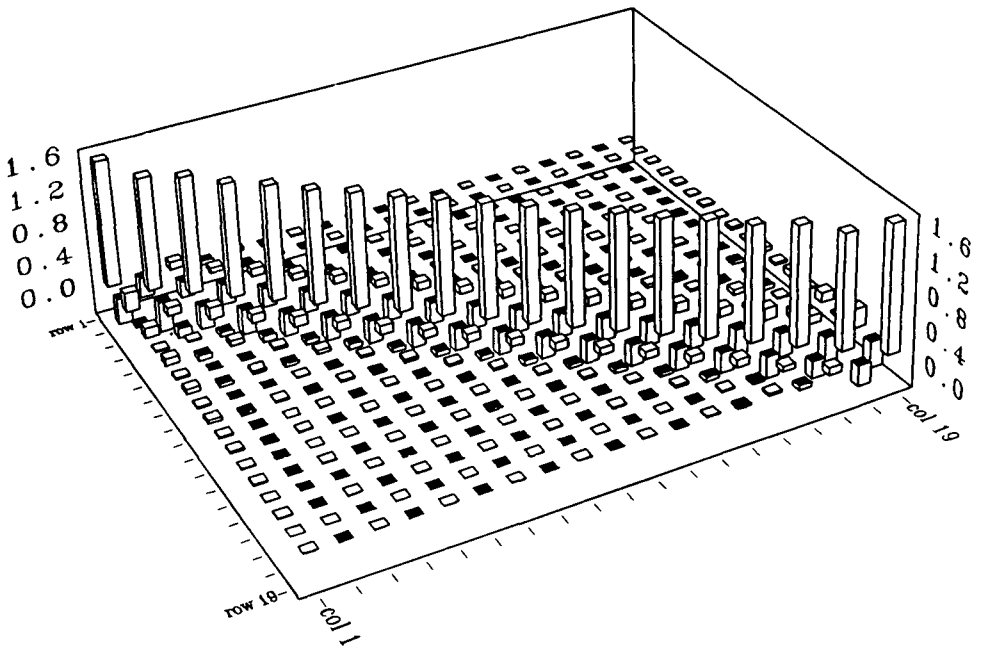


Fig. 19. Display of the matrix $F^{-1}C$ resulting from second-order FD preconditioning of the Chebyshev DM C for $\mathrm{d}^2/\mathrm{d}x^2$ (with $n = 20$).

5.5. Improved conditioning through change of variable

All the approximations to the derivatives that we have considered so far (for nonperiodic problems) have been based on differentiating interpolating polynomials. The difficulties at boundaries have been linked to large weights in one-sided stencils – in turn a consequence of the rapid growth of high-degree polynomials at increasing distances from the origin.

Kosloff and Tal-Ezer (1993) proposed to change first the independent variable x into y through $x = \arcsin(\alpha y)/\arcsin(\alpha)$ (both x and $y \in [-1, 1]$, the parameter $\alpha \in [0, 1]$). In the governing equations, $\partial/\partial x$ needs then to be replaced by

$$\frac{\arcsin(\alpha)}{\alpha}\sqrt{1-(\alpha y)^2}\,\frac{\partial}{\partial y}$$

(and similarly for higher derivatives). Applying a standard Chebyshev PS method in the y variable corresponds, in the x variable, to working with nonpolynomial basis functions.

In the limit of $\alpha \to 0$, y is equal to x, and we have the regular Chebyshev PS method. As $\alpha \to 1$, the x grid approaches uniform spacing. Close to this limit, the Chebyshev polynomials (in the y variable) have, in the x variable,

become stretched to resemble trigonometric functions. This reduces the spurious EVs. In Example 2 of Section 4.4, they decrease from $\mathcal{O}(N^2)$ to $\mathcal{O}(N)$. Figure 20(a) shows how they move in this case when α increases from 0 (as in Figure 14) to 0.9. Figure 20(b) shows the effect this has on the accuracy of different Fourier modes.

Compared to the standard Chebyshev PS method, this procedure offers reduced spurious EVs, better conditioned DMs (much closer to being normal matrices) and a wider range of accurately treated Fourier modes in exchange for less accuracy for the lowest Fourier modes.

If we had just moved the grid points towards equi-spaced locations without the accompanying change of variable, we would suffer all the problems of the Runge phenomenon (cf. Figures 3 and 8(a)) – disastrous growth of condition number and an inability to approximate anything but extremely smooth functions.

The properties of this method are similar to those of medium-to-high order FD schemes – no conclusive efficiency comparisons have yet been carried out.

The idea of changing variable to improve the Chebyshev PS method was proposed earlier by Bayliss *et al.* (1989) for quite a different purpose – to achieve additional grid clustering at interior locations where extra resolution might be needed (see also Bayliss and Turkel (1992)).

5.6. PS methods in polar and spherical coordinates

Separation of variables for the Laplacian operator in three-dimensional polar coordinates leads to a class of functions called 'spherical harmonics' – combinations of trigonometric and Gegenbauer polynomials. These offer a complicated (but workable) base for spectral methods. For an overview of this approach, see, e.g., Boyd (1989, Ch. 15).

An alternative is outlined below, for simplicity first for polar coordinates in the plane. Instead of generating PS methods from some set of basis functions, we start from the basic FD premise that derivatives in different spatial directions can be approximated entirely separately from each other. In each direction, we thus use the most appropriate FD scheme of maximal order (typically one-dimensional Fourier or Chebyshev PS approximations).

Polar coordinates A polar coordinate system on the unit circle can be obtained through

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases} \qquad 0 \le r \le 1, \ -\pi \le \theta \le \pi.$$

At $r = 0$, all θ positions collapse into one physical grid point – therefore requiring only one governing equation. At this location, one can use a Cartesian x–y-based FD stencil (free of the singularities that might have been introduced by the polar coordinate formulation).

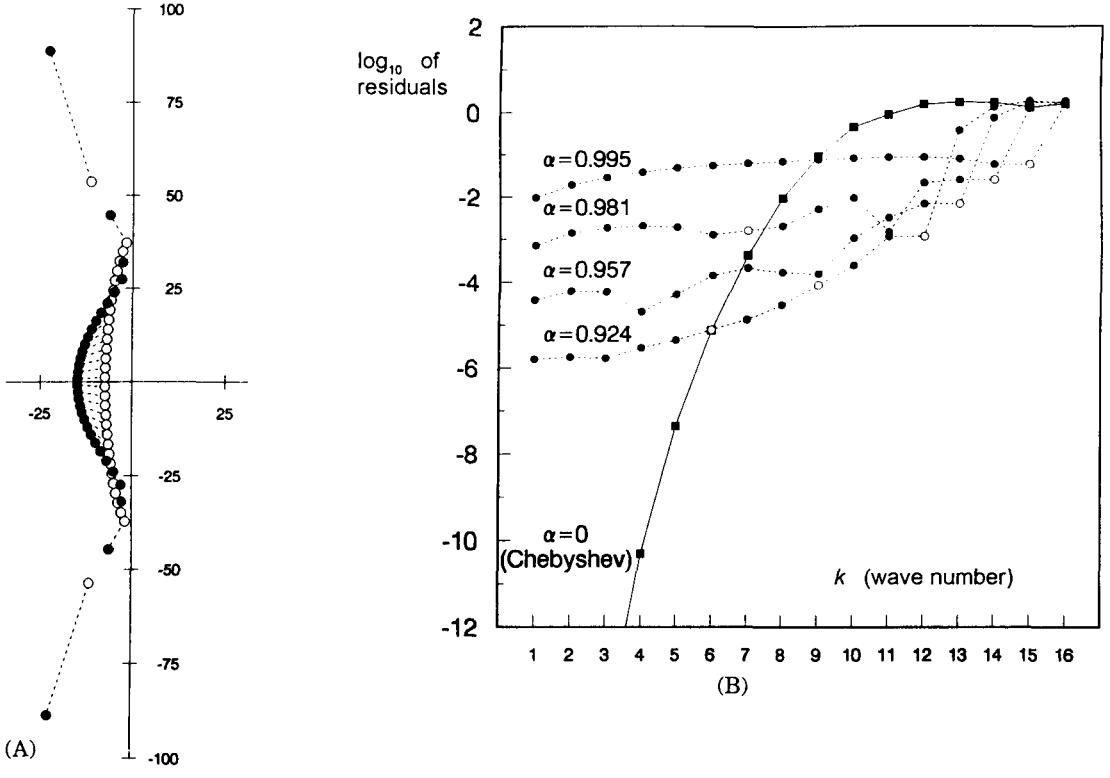


Fig. 20. Effect on EVs and accuracy from changes in mapping parameter α ($N = 32$). (a) Changes in EVs when α is increased from 0 (Chebyshev case, cf. Figure 14) to 0.9. (b) Residuals when first derivative approximations are applied to $u_k(x) = \cos(k\pi x)$, $k = 1, 2, \ldots, 16$. The values of $\alpha = \cos(\pi j/32)$, $j = 1,\ 2,\ 3$ and 4 correspond to requiring 'accuracy' up to node $16 - j$. The approximations become exact in certain cases – indicated by open circles (their vertical positions are artificial and serve only to make the chart easier to read). Numerical data are taken from Kosloff and Tal-Ezer (1993).

One might be tempted to proceed by using Fourier PS in θ and Chebyshev PS in r. A much better alternative is to consider $-1 \leq r \leq 1$, $0 \leq \theta \leq \pi$ (instead of $0 \leq r \leq 1$, $-\pi \leq \theta \leq \pi$). There is then no longer any reason to refine the grid in the r direction near $r = 0$:

- saves grid points;
- higher-order accuracy of PS stencils in r directions (since they extend over twice as many grid points – cf. Example 2 in Section 5.1);
- less severe two-dimensional point clustering near the origin;
- high degree of smoothing in the θ direction for small r values provides favourable CFL (Courant–Friedrichs–Levy) stability conditions without damaging overall accuracy;
- Fourier PS available as before in the θ direction.

Spherical coordinates We consider a surface φ, θ – grid as shown in Figure 21. The dotted arrows indicate how periodicity can be implemented in both φ and θ. The observation for two-dimensional coordinates carry over;

- for θ near $\pm\pi$, polar stability can be enhanced by smoothing in the φ direction;
- an r direction can again be added as in two-dimensions (with $-1 \leq r \leq 1$ and halving the angular domain in case $r = 0$ is in the region of interest).

This PS method has been tested for convective flow in different directions over the surface of a sphere (Fornberg, 1994). Its performance turns out to be entirely unharmed by the presence of polar singularities – the accuracy is as high as is typical for one-dimensional periodic problems.

6. Comparisons of computational cost – FD versus PS methods

High-order FD and the PS methods are particularly advantageous in cases of high smoothness of solution (but note again the discussion in Section 4.2), stringent error requirement, long time integrations and more than one space dimension.

Since the PS methods for periodic and nonperiodic problems are quite different, they will be discussed separately.

6.1. Periodic problems

To obtain more precise insights into how the formal order of a method affects the accuracy, we consider the model problem $\partial u/\partial t + \partial u/\partial x = 0$ on $[-1, 1]$, integrated in time from 0 to 2 (the time it takes the analytical solution $u(x, t)$ to move once across the period). The data in Figure 10 can be recast

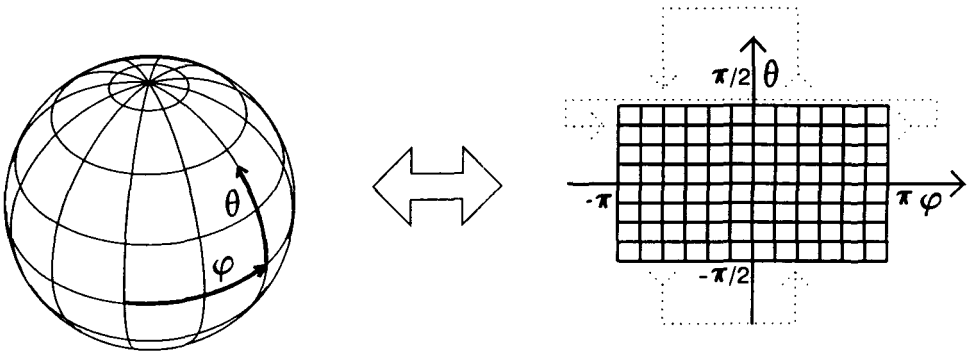


Fig. 20. Grid arrangement and periodicities for spherical coordinates.

into Figure 22 (for details, see Fornberg, 1987). The following explains how to interpret this figure.

Using an accurate time integrator, Fourier modes in the numerical solution of this equation will develop phase but not amplitude errors. If a phase error is π, that mode will have the wrong sign and will not add any accuracy to a Fourier expansion. Here we consider (somewhat arbitrarily) a mode to be 'accurate' if its phase error is less than $\pi/4$.

A second-order FD method with $N_{GP} = 500$ (i.e. 500 grid points in the spatial direction) is seen to give the same accuracy (have the same horizontal position in the figure) as a fourth-order FD method with $N_G \approx 160$ and a PS method with $N_G \approx 32$. The numbers on the axes indicate: horizontally (approximately location '16' in the example above), modes up to $\sin(16\pi x)$, $\cos(16\pi x)$ are 'accurate' at the end of the integration; vertically, in the different cases the number of grid points needed per wavelength.

The governing equations for the test case shown in Figures 23(a) and (b) are those in Figure 17 (here using a 'regular' grid). A sharp 'pressure wave' pulse is sent down through an elastic medium which carries both pressure and shear waves with lower velocities near the centre of the domain. Following focusing and the subsequent development of a cusp-shaped wave front, Figure 23(b) shows second- and fourth-order FD and PS results. In the three cases, comparable accuracies are obtained on grids of densities 512×512, 128×128 and 32×32 respectively (in quite good agreement with the discussion just above). With the PS method implemented by FFTs, the relative computer times scale as 20:2:1. In three dimensions, these numbers would become 300:8:1. For memory requirements, the differences become even larger: 256:16:1 in two dimensions and 4096:64:1 in three dimensions.

The largest cost–benefits from high-order methods arise in two and three

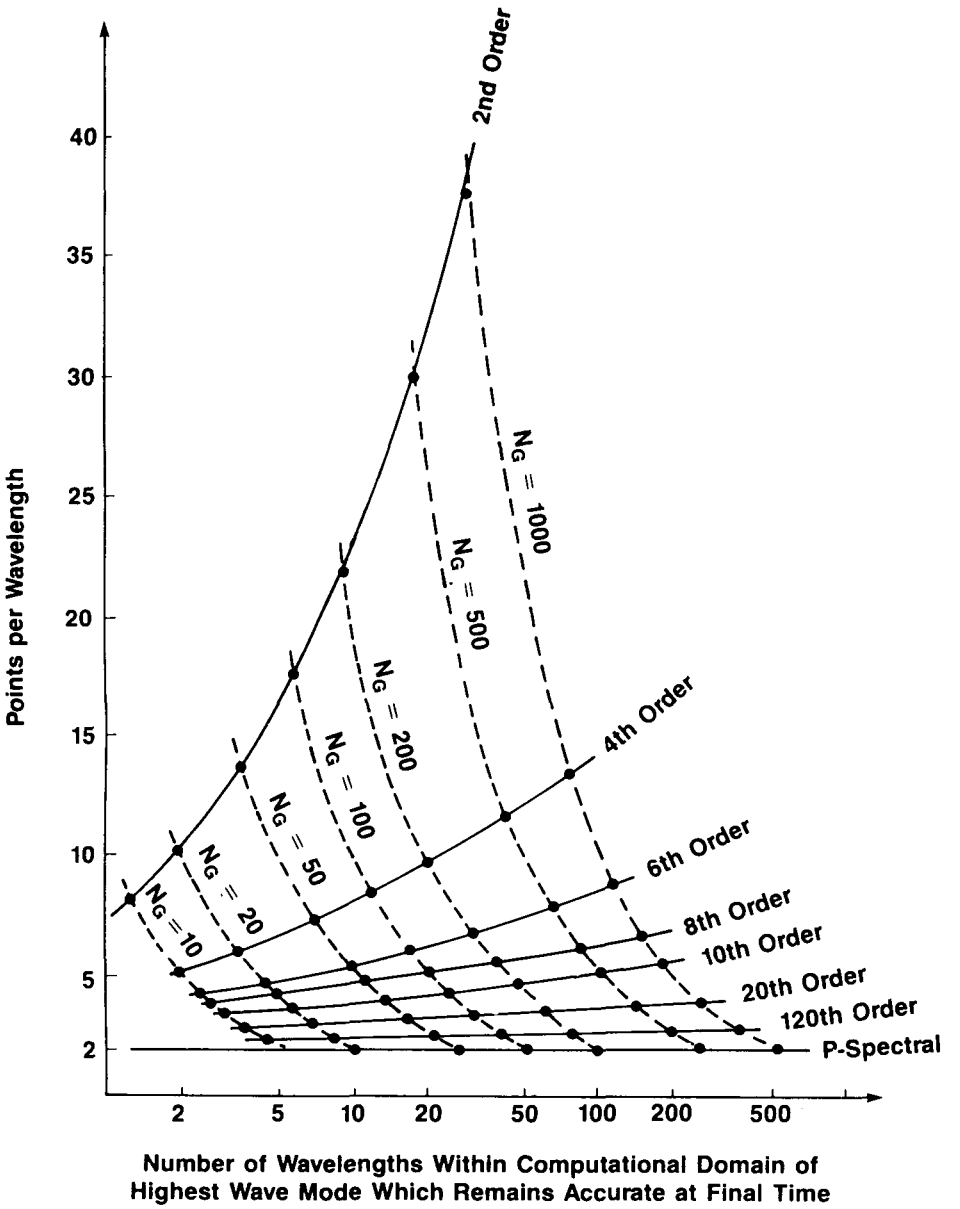
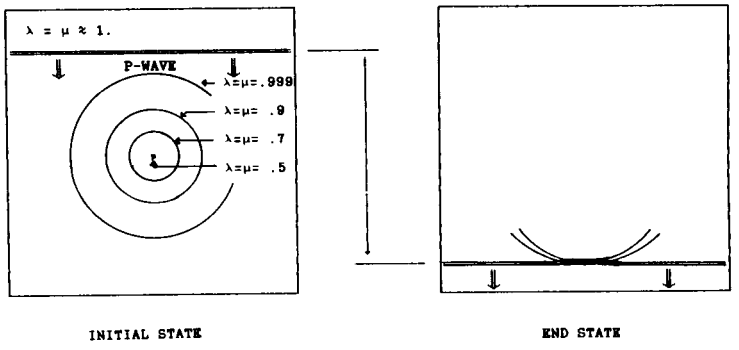
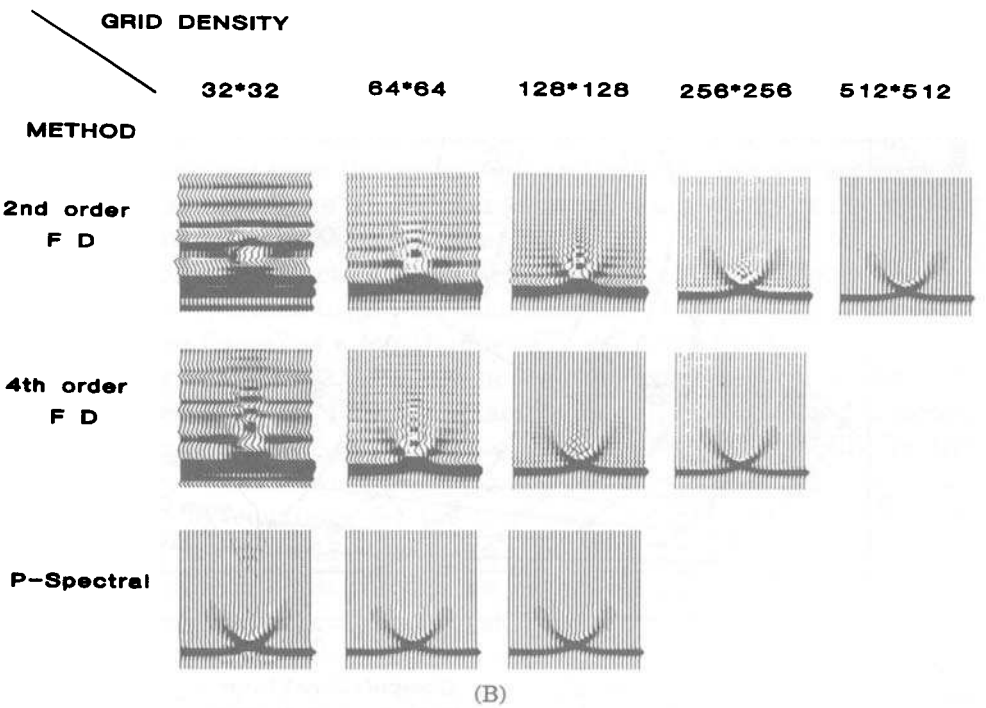


Fig. 22. Relations between grid densities and obtained accuracies when applying different methods to a model problem.



(A)



(B)

Fig. 23. (a) Contour curves for the variable medium and schematic illustrations of the initial and end states of the test runs. (b) Numerical results for the test problem (variable f displayed). Comparison between different methods and grid densities.

dimensions. Hou and Kreiss (1993) note that in one dimension and with near-singular solutions (for example with thin internal layers to resolve), fourth- and sixth-order FD methods sometimes match (or even exceed) the PS method in efficiency.

6.2. Nonperiodic problems

In this case it is more difficult to provide any single (and still simple) test example that is general enough to be meaningful. Passing from periodic to nonperiodic problems, PS methods encounter many more problems than (low-order) FD methods:

FD: Some more care is needed in stability analysis.

PS: Grid clustering is necessary near edges. This leads to

- conditioning and stability problems (especially notable when time stepping),
- need for preconditioners,
- the prevalence of spurious EVs, especially for high derivatives (cf. Merryfield and Shizgal (1993), on the KdV equation – in sharp contrast to a very favourable situation for periodic PS methods (Fornberg and Whitham, 1978),
- reduced ability to resolve Fourier modes (need π versus 2 points per wavelength).

Formal order of accuracy is the same as the number of grid points – not infinite as in the periodic case (but the significance of this 'philosophical' difference is unclear).

Performance in nonsmooth cases is less well known.

The many successes of Chebyshev-type PS methods in a wide range of applications prove that the added complications are often outweighed by the advantage of exponential accuracy.

One convenient way to keep many options open when developing application codes is to write a FD code of variable order of accuracy on a grid with variable density (using the algorithm in Section 3.1 and Appendix 2). By simply changing parameter values, one can then explore (and exploit) the full range of methods from low-order FD on a uniform grid to Chebyshev (Legendre etc.) and other PS methods. (Obviously, it is also desirable to structure codes so that time stepping methods (if present) are easily interchangeable.) The optimal selections may well turn out to depend not only on the problem type but also on the solution regimes that are studied, the accuracy that is desired etc.

Appendix A. Implementation of Tau, Galerkin and Collocation (PS) for a 'toy' problem

We consider the following model problem

$$u_{xx} + u_x - 2u + 2 = 0, \quad -1 \le x \le 1,$$

$$u(-1) = u(1) = 0,$$

and approximate the exact solution

$$u(x) = 1 - \frac{\sinh(2)e^x + \sinh(1)e^{-2x}}{\sinh(3)}$$

by

$$v(x) = \sum_{k=0}^{4} a_k T_k(x).$$

From (2.1) and (2.2) it follows that the residual

$$R(x) = v_{xx} + v_x - 2v + 2 = \sum_{k=0}^{4} A_k T_k(x) \tag{A.1}$$

satisfies

$$\begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 4 & 3 & 32 \\ 0 & -2 & 4 & 24 & 8 \\ 0 & 0 & -2 & 6 & 48 \\ 0 & 0 & 0 & -2 & 8 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{A.2}$$

The matrix is obtained as

$$\begin{bmatrix} 0 & 0 & 4 & 0 & 32 \\ & 0 & 0 & 24 & 0 \\ & & 0 & 0 & 48 \\ & & & 0 & 0 \\ & & & & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 & 3 & 0 \\ & 0 & 4 & 0 & 8 \\ & & 0 & 6 & 0 \\ & & & 0 & 8 \\ & & & & 0 \end{bmatrix} - 2 \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix}$$

corresponding to v_{xx}, v_x and $-2v$ respectively. In general, $\partial^p v/\partial x^p$, the matrix becomes A^p where A is the inverse of the matrix in (2.1) with all zero first column and last row added (a consequence of the shift in the indices between the two column vectors in (2.2)). This procedure to find the elements of A generalizes immediately to Jacobi ploynomials. Closed form expressions for $A^p_{i,j}$ (the element at row i, column j, $0 \le i, j \le n$) turn out to be very simple for both Legendre and Chebyshev expansions. In the Chebyshev case:

$$A_{i,j} = \begin{cases} 1/c_i \times 2j & j > i, i+j \text{ odd}, \\ 0 & \text{otherwise}, \end{cases}$$

$$A^2_{i,j} = \begin{cases} 1/c_i \times (j-i)j(j+i), & j > i, i+j \text{ even}, \\ 0, & \text{otherwise}, \end{cases}$$

where

$$c_i = \begin{cases} 2, & i = 0 \text{ odd}, \\ 1, & i > 0, \end{cases}$$

and

$$A^{p+2}_{i,j} = A^p_{i,j} \times (j-i-p)(j+i+p)(j-i+p)(j+i-p)/[16p(p+1)], \quad p \geq 1$$

(can be shown using the theorem in Karageorghis (1988); for Legendre polynomials, see Phillips (1988)).

Enforcing the boundary conditions $v(-1) = v(1) = 0$ leads to

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{A.3}$$

Ideally, we would like to get $A_i = 0$, $i = 0, 1, \ldots, 4$ while still satisfying (A.3). However, this would mean satisfying seven relations with only five free parameters a_i, $i = 0, 1, \ldots, 4$. The three spectral methods differ in how they approximate this overdetermined system.

Tau Require $R(x)$ (A.1) to be orthogonal to $T_k(x)$, $k = 0, 1, 2$:

$$\int_{-1}^{1} \frac{R(x)T_k(x)}{\sqrt{1-x^2}}\,\mathrm{d}x = 0 \quad \Rightarrow \quad \begin{aligned} A_0 &= 0, \\ A_1 &= 0, \\ A_2 &= 0. \end{aligned}$$

The top three lines of (A.2) together with (A.3) give

$$[a_0, \ldots, a_4] = [0.2724, -0.0444, -0.2562, 0.0444, -0.0162].$$

Galerkin Create from $T_0, \ldots, T_4$ three basis functions Φ_2, Φ_3, Φ_4 which satsfy both boundary conditions:

$$\begin{aligned} \Phi_2(x) &= T_2(x) - T_0(x), \\ \Phi_3(x) &= T_3(x) - T_1(x), \\ \Phi_4(x) &= T_4(x) - T_0(x). \end{aligned}$$

Then $v(x) = \sum_{k=2}^{4} c_k \Phi_k(x)$ which is equal to $\sum_{k=0}^{4} a_k T_k(x)$ constrained

by (A.3). Require $R(x)$ to be orthogonal to $\Phi_k(x)$, $k = 2,3,4$:

$$\int_{-1}^{1} \frac{R(x)\Phi_k(x)}{\sqrt{1-x^2}}\,dx = 0 \quad \Rightarrow \begin{bmatrix} 2 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 2 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Together with (A.2) and (A.3):

$$[a_0, \ldots, a_4] = [0.2741, -0.0370, -0.2593, 0.0370, -0.0148].$$

Collocation (PS) Force $R(x_i) = 0$ at $x_i = \cos(i\pi/4)$, $i = 1,2,3$:

$$\begin{bmatrix} 1 & 1/\sqrt{2} & 0 & -1/\sqrt{2} & -1 \\ 1 & 0 & -1 & 0 & 1 \\ 1 & -1/\sqrt{2} & 0 & 1/\sqrt{2} & -1 \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The section of the discrete cosine transform matrix has the entries $T_k(x_i) = \cos ki\pi/4$, $k = 0, \ldots, 4$, $i = 1,2,3$. Together with (A.2) and (A.3):

$$[a_0, \ldots, a_4] = [0.2473, -0.0371, -0.2600, 0.0143].$$

In exact arithmetic,

$$[a_0, \ldots, a_4] = [\tfrac{48}{175}, -\tfrac{13}{350}, -\tfrac{13}{50}, \tfrac{13}{350}, -\tfrac{1}{70}]$$

and the values at the node locations $x_i, i = 0, \ldots, 4$ become

$$0, \tfrac{101}{350}, +\tfrac{13}{350}\sqrt{2}, \tfrac{13}{25}, \tfrac{101}{350}, -\tfrac{13}{350}\sqrt{2}, 0].$$

This description of the PS approach followed the style of those for Tau and Galerkin, but gave no indication why the PS approach is more flexible than the other two in cases of variable coefficients and nonlinearities. We therefore describe the PS method again, this time in terms of nodal values rather than expansion coefficients. If ν_i denote the approximations at the nodes $x_i = \cos i\pi/4$, $i = 0, 1, \ldots, 4$, the first and second derivatives of the interpolating polynomial become

$$\begin{bmatrix} \nu_{x\,0} \\ \nu_{x\,1} \\ \nu_{x\,2} \\ \nu_{x\,3} \\ \nu_{x\,4} \end{bmatrix} = \begin{bmatrix} -\frac{11}{2} & 4+2\sqrt{2} & -2 & 4-2\sqrt{2} & -\frac{1}{2} \\ -1-\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & \sqrt{2} & -\frac{1}{2}\sqrt{2} & 1-\frac{1}{2}\sqrt{2} \\ \frac{1}{2} & -\sqrt{2} & 0 & \sqrt{2} & -\frac{1}{2} \\ -1+\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & -\sqrt{2} & -\frac{1}{2}\sqrt{2} & 1+\frac{1}{2}\sqrt{2} \\ \frac{1}{2} & -4+2\sqrt{2} & 2 & -4-2\sqrt{2} & \frac{11}{2} \end{bmatrix} \begin{bmatrix} \nu_0 \\ \nu_1 \\ \nu_2 \\ \nu_3 \\ \nu_4 \end{bmatrix} \tag{A.4}$$

and

$$\begin{bmatrix} \nu_{xx\,0} \\ \nu_{xx\,1} \\ \nu_{xx\,2} \\ \nu_{xx\,3} \\ \nu_{xx\,4} \end{bmatrix} = \begin{bmatrix} 17 & -20-6\sqrt{2} & 18 & -20+6\sqrt{2} & 5 \\ 5+3\sqrt{2} & -14 & 6 & -2 & 5-3\sqrt{2} \\ -1 & 4 & -6 & 4 & -1 \\ 5-3\sqrt{2} & -2 & 6 & -14 & 5+3\sqrt{2} \\ 5 & -20+6\sqrt{2} & 18 & -20-6\sqrt{2} & 17 \end{bmatrix} \begin{bmatrix} \nu_0 \\ \nu_1 \\ \nu_2 \\ \nu_3 \\ \nu_4 \end{bmatrix} \tag{A.5}$$

These matrices are examples of 'differentiation matrices' – discussed in numerous places in this review. Section 4.3 describes how their elements can be obtained very conveniently.

Enforcing

$$R(x) = \nu_{xx} + \nu_x - 2\nu + 2 = 0$$

at the node points x_k, $k = 1, 2, 3$ and the boundary conditions $\nu_0 = \nu_4 = 0$ lead to

$$\begin{bmatrix} -16+\frac{1}{2}\sqrt{2} & 6+\sqrt{2} & -2-\frac{1}{2}\sqrt{2} \\ 4-\sqrt{2} & -8 & 4+\sqrt{2} \\ -2+\frac{1}{2}\sqrt{2} & 6-\sqrt{2} & -16-\frac{1}{2}\sqrt{2} \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \\ -2 \end{bmatrix} \tag{A.6}$$

with the same solution as before:

$$\begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix} = \begin{bmatrix} \frac{101}{350}+\frac{13}{350}\sqrt{2} \\ \frac{13}{25} \\ \frac{101}{350}-\frac{13}{350}\sqrt{2} \end{bmatrix} \tag{A.7}$$

Had variable coefficients been present, their values at the nodes would have been used to multiply the rows of (A.4), (A.5) etc., when assembling (A.6).

Figure A.1 shows how the accuracy increases with n – in all three cases featuring an exponential rate of convergence. For comparison, curves for second- and fourth-order FD approximations (on equi-spaced grids) are also included.

Appendix B. Fortran code and test driver for algorithm to find weights in FD formulae

```
      SUBROUTINE WEIGHTS (XI,X,N,M,C)
C     INPUT PARAMETERS:
C       XI  POINT AT WHICH THE APPROXIMATIONS ARE TO BE ACCURATE
C       X   X-COORDINATES FOR THE GRID POINTS, ARRAY DIMENSIONED X(0:N)
C       N   THE GRID POINTS ARE AT X(0),X(1), ...  X(N) (I.E. N+1 IN ALL)
C       M   HIGHEST ORDER OF DERIVATIVE TO BE APPROXIMATED
C
```

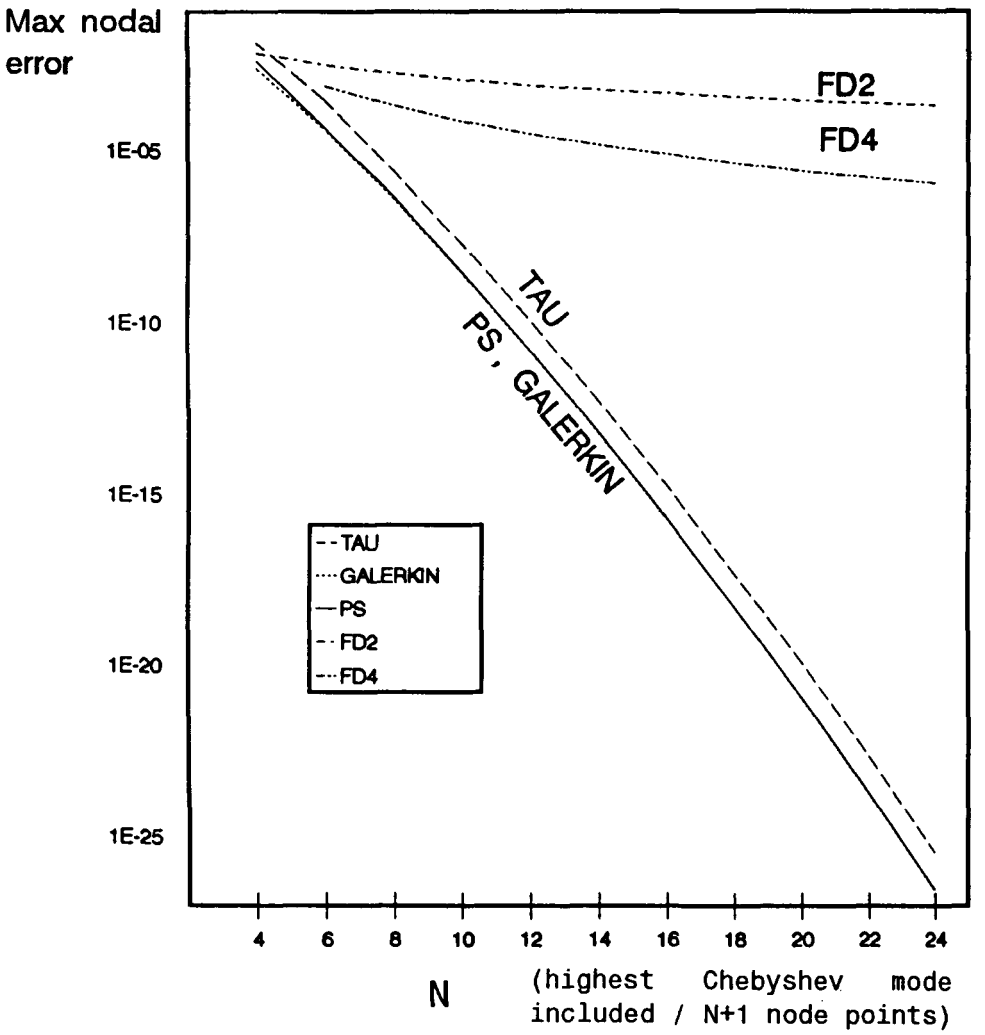


Fig. A.1. Maximum nodal errors for different methods when applied to 'toy problem' in Appendix 1 – comparison between three spectral implementations and equispaced FD methods of second and fourth order.

```
C    OUTPUT PARAMETERS:
C      C  WEIGHTS, ARRAY DIMENSIONED C(0:N,0:N,0:M).
C         ON RETURN, THE ELEMENT C(K,J,I) CONTAINS THE WEIGHT TO BE
C         APPLIED AT X(K) WHEN THE I:TH DERIVATIVE IS APPROXIMATED
C         BY A STENCIL EXTENDING OVER X(0),X(1), ...  , X(J).
C
C
     DIMENSION X(0:N),C(0:N,0:N,0:M)
     C(0,0,0) = 1.
     C1       = 1.
     C4       = X(0)-XI
     DO 40 J=1,N
       MN = MIN(J,M)
       C2 = 1.
       C5 = C4
       C4 = X(J)-XI
       DO 20 K=0,J-1
          C3 = X(J)-X(K)
          C2 = C2*C3
          IF (J.LE.M) C(K,J-1,J)=0.
          C(K,J,0) = C4*C(K,J-1,0)/C3
          DO 10 I=1,MN
10           C(K,J,I) = (C4*C(K,J-1,I)-I*C(K,J-1,I-1))/C3
20        CONTINUE
       C(J,J,0) = -C1*C5*C(J-1,J-1,0)/C2
       DO 30 I=1,MN
30        C(J,J,I) = C1*(I*C(J-1,J-1,I-1)-C5*C(J-1,J-1,I))/C2
40     C1 = C2
     RETURN
     END
```

Note If N is very large, the calculation of the variable C2 might cause overflow (or underflow). For example, in generating extensions of Tables 2 and 3, this problem arises when N! exceeds the largest possible number (i.e. N > 34 in typical 32 bit precision with 3×10^{38} as the largest number; N > 965 in CRAY single precision (64 bit word length, 15 bit exponent, largest number approximately 10^{2465})). In such cases, scaling of C1 and C2 (only used in forming the ratio C1/C2) should be added to the code.

The following test program will print out all the entries in Table 3 (including a table of coefficients for the zeroth derivative – interpolation weights – trivial here since approximations are requested at a grid point).

```
PROGRAM TEST
PARAMETER (M=4,N=8)
DIMENSION X(0:N),C(0:N,0:N,0:M)
```

```
      DO 10 I=0,N
10      X(I) = I
      CALL WEIGHTS (0.,X,N,M,C)
      DO 30 I=0,M
        DO 20 J=I,N
20        WRITE (6,40) (C(K,J,I),K=0,J)
30      WRITE (6,*)
40    FORMAT (1X,9F8.3)
      STOP
      END
```

All the data in Table 2 can similarly be obtained from a single call to SUBROUTINE WEIGHTS by initializing

$$\texttt{X(0:8) to /0, -1,1,-2,2,-3,3,-4,4/}$$

(and ignore every second line of the output).

REFERENCES

S. Abarbanel, D. Gottlieb and E. Tadmor (1986), 'Spectral methods for discontinuous problems', in *Numerical Methods for Fluid Dynamics II* (K.W. Morton and M.J. Baines, eds), Clarendon Press (Oxford), 129–153.

A. Bayliss, D. Gottlieb, B.J. Matkowsky and M. Minkoff (1989), 'An adaptive pseudospectral method for reaction diffusion problems', *J. Comput. Phys.* **81**, 421–443.

A. Bayliss and E. Turkel (1992), 'Mappings and accuracy for Chebyshev pseudospectral approximations', *J. Comput. Phys.* **101**, 349–359.

G. Beylkin, R. Coifman and V. Rokhlin (1991), 'Fast wavelet transforms and numerical algorithms', *Comm. Pure Appl. Math.* **44**, 141–183.

J.P. Boyd (1989), *Chebyshev and Fourier Spectral Methods*, Springer (New York)

J.P. Boyd (1992), 'Multipole expansions and pseudospectral cardinal functions: a new generalization of the fast Fourier transform', *J. Comput. Phys.* **103**, 184–186.

K.S. Breuer and R.M. Everson (1992), 'On the errors incurred calculating derivatives using Chebyshev polynomials', *J. Comput. Phys.* **99**, 56–67.

W. Cai, D. Gottlieb and C.-W. Shu (1989), 'Essentially nonoscillatory spectral Fourier methods for shock wave calculation', *Math. Comput.* **52**, 389–410.

W. Cai, D. Gottlieb and C.W. Shu (1992), 'One-sided filters for spectral Fourier approximation of discontinuous functions', *SIAM J. Numer. Anal.* **29**, 905–916.

C. Canuto, M.Y. Hussaini, A. Quarteroni and T. Zang (1988), *Spectral Methods in Fluid Dynamics*, Springer (New York).

C. Canuto and P. Pietra (1987), 'Boundary and interface conditions with a FE preconditioner for spectral methods'. Report No. 553, I.A.N., Pavia University, Italy.

C. Canuto and A. Quarteroni (1985), 'Preconditioned minimal residual methods for Chebyshev spectral calculations', *J. Comput. Phys.* **60**, 315–337.

C. Canuto and A. Quarteroni (1987), 'On the boundary treatment in spectral method for hyperbolic systems', *J. Comput. Phys.* **71**, 100–110.

E.W. Cheney (1966), *Introduction to Approximation Theory*, McGraw-Hill (New York).

J.W. Cooley and J.W. Tukey (1965), 'An algorithm for the machine calculation of complex Fourier series', *Math. Comput.* **19**, 297–301.

G. Dahlquist (1956), 'Convergence and stability in the numerical integration of ordinary differential equations', *Math. Scand.* **4**, 33–53.

G. Dahlquist (1985), '33 years of numerical instability, part I', *BIT* **25**, 188–204.

P.J. Davis (1975), *Interpolation and Approximation*, Dover (New York).

P.J. Davis and P. Rabinowitz (1984), *Methods of Numerical Integration*, 2nd Edn, Academic Press (London, New York).

M. Deville and E. Mund (1985), 'Chebyshev PS solution of second-order elliptic equations with finite element preconditioning', *J. Comput. Phys.* **60**, 517–553.

M. Dubiner (1987), 'Asymptotic analysis of spectral methods', *J. Sci. Comput.* **2**, 3–31.

H. Eisen and W. Heinrichs (1992), 'A new method of stabilization for singular perturbation problems with spectral methods', *SIAM J. Numer. Anal.* **29**, 107–122.

B. Fornberg (1975), 'On a Fourier method for the integration of hyperbolic equations', *SIAM J. Numer. Anal.* **12**, 509–528.

B. Fornberg (1987), 'The pseudospectral method: comparisons with finite differences for the elastic wave equation', *Geophysics* **52**, 483–501.

B. Fornberg (1988a), 'The pseudospectral method: accurate representation of interfaces in elastic wave calculations', *Geophysics* **53**, 625–637.

B. Fornberg (1988b), 'Generation of finite difference formulae on arbitrarily spaced grids', *Math. Comput.* **51**, 699–706.

B. Fornberg (1990a), 'High order finite differences and the pseudospectral method on staggered grids', *SIAM J. Num. Anal.* **27**,904–918.

B. Fornberg (1990b), 'An improved pseudospectral method for initial-boundary value problems', *J. Comput. Phys.* **91**, 381–397.

B. Fornberg (1992), 'Fast generation of weights in finite difference formulas', in *Recent Developments in Numerical Methods and Software for ODEs/DAEs/PDEs* (G.D. Byrne and W.E. Schiesser, eds), World Scientific (Singapore), 97–123.

B. Fornberg (1994), 'A pseudospectral approach for polar and spherical geometries', *SIAM J. Sci. Comput.*, submitted.

B. Fornberg and G.B. Whitham (1978), 'A numerical and theoretical study of certain nonlinear wave phenomena', *Phil. Trans. Roy. Soc. London* A **289**, 373–404.

J.B.J. Fourier (1822), *Théorie analytique de la chaleur*, Paris.

D. Funaro (1987), 'A preconditioned matrix for the Chebyshev differencing operator', *SIAM J. Numer. Anal.* **24**, 1024–1031.

D. Funaro (1992), *Polynomial Approximation of Differential Equations*, Lecture Notes in Physics 8, Springer (Berlin).

D. Funaro and D. Gottlieb (1988), 'A new method of imposing boundary conditions in pseudospectral approximations of hyperbolic equations', *Math. Comput.* **51**, 599–613.

D. Gaier (1987), *Lectures on Complex Approximation*, Birkhäuser (Boston).

C.W. Gear (1971), *Numerical Solution of Ordinary and Partial Differential Equations*, Prentice Hall (Englewood Cliffs, NJ).

M. Goldberg and E. Tadmor (1985), 'Convenient stability criteria for difference approximations of hyperbolic initial-boundary value problems', *Math. Comput.* **44**, 361–377.

D. Gottlieb and L. Lustman (1983), 'The spectrum of the Chebyshev collocation operator for the heat equation', *SIAM J. Numer. Anal.* **20**, 909–921.

D. Gottlieb, L. Lustman and E. Tadmor (1987), 'Stability analysis of spectral methods for hyperbolic initial-boundary value problems', *SIAM J. Num. Anal.* **24**, 241–258.

D. Gottlieb and S.A. Orszag (1977), *Numerical Analysis of Spectral Methods*, SIAM (Philadelphia).

D. Gottlieb and E. Tadmor (1991), 'The CFL condition for spectral approximation to hyperbolic BVPs', *Math. Comput.* **56**, 565–588.

D. Gottlieb and E. Turkel (1980), 'On time discretization for spectral methods', *Stud. Appl. Math.* **63**, 67–86.

P.M. Greshko and R.L. Lee (1981), 'Don't suppress the wiggles – they're trying to tell you something', *Computers and Fluids* **9**, 223–253.

B. Gustafsson, H.-O. Kreiss and A. Sundstrom (1972), 'Stability theory of difference approximations for mixed initial-boundary value problems II', *Math. Comput.* **26**, 649–685.

E. Hairer, S.P. Nørsett and G. Wanner (1987), *Solving Ordinary Differential Equations I - Nonstiff Problems*, Springer (Berlin).

E. Hairer and G. Wanner (1991), *Solving Ordinary Differential Equations II - Stiff and Differential-Algebraic Problems*, Springer (Berlin).

P. Haldenwang, G. Labrosse, S. Abboudi and M. Deville (1984), 'Chebyshev 3-D spectral and 2-D pseudospectral solvers for the Helmholtz equation', *J. Comput. Phys.* **55**, 115–128.

O. Holberg (1987), 'Computational aspects of the choice of operator and sampling interval for numerical differentiation in large-scale simulation of wave phenomena', *Geophys. Prospecting* **35**, 629–655.

Y.-C. C. Hou and H.-O. Kreiss (1993), 'Comparison of finite difference and the pseudo-spectral approximations for hyperbolic equations', to be published.

W. Huang and D.M. Sloan (1993a), 'Pole condition for singular problems: the pseudospectral approximation', *J. Comput. Phys.* **107**, 254–261.

W. Huang and D.M. Sloan (1993b), 'A new pseudospectral method with upwind features', *IMA. J. Numer. Anal.* **13**, 413–430.

W. Huang and D.M. Sloan (1993c), 'The pseudospectral method for solving differential eigenvalue problems', *J. Comput. Phys.*, to appear.

M.Y. Hussaini and T.A. Zang (1984), 'Iterative spectral methods and spectral solution to compressible flows', in *Spectral Methods for PDEs* (R. Voigt, D. Gottlieb and M.Y. Hussaini, eds), SIAM (Philadelphia).

A. Karageorghis (1988), 'A note on the Chebyshev coefficients of the general order derivative of an infinitely differentiable function', *J. Comput. Appl. Math.* **21**, 129–132.

M. Kindelan, A. Kamel and P. Sguazzero (1990), 'On the construction and efficiency of staggered numerical differentiators for the wave equation', *Geophysics* **55**, 107–110.

D. Kosloff and H. Tal-Ezer (1993), 'A modified Chebyshev pseudospectral method with an $\mathcal{O}(N^{-1})$ time step restriction', *J. Comput. Phys.* **104**, 457-469.

H.-O. Kreiss (1962), 'Über die Stabilitätsdefinition für Differenzengleichungen die partielle Differentialgleichungen approximieren', *Nordisk Tidskr. Informationsbehandling* **2**, 153–181.

H.-O. Kreiss (1968), 'Stability theory for difference approximations of mixed initial boundary value problems I', *Math. Comput.* **22**, 703–714.

H.-O. Kreiss (1970), 'Initial boundary value problems for hyperbolic systems', *Comm. Pure Appl. Math.* **23**, 277–288.

H.-O. Kreiss and J. Oliger (1972), 'Comparison of accurate methods for the integration of hyperbolic equations', *Tellus* **XXIV**, 199–215.

V.I. Krylov (1962), *Approximate Calculation of Integrals*, Macmillan (New York).

J.D. Lambert (1991), *Numerical Methods for Ordinary Differential Systems: The Initial Value Problem*, Wiley (Chichester, UK).

C. Lanczos (1938), 'Trigonometric interpolation of empirical and analytical functions', *J. Math. Phys.* **17**, 123–199.

S.K. Lele (1992), 'Compact finite difference schemes with spectral-like resolution', *J. Comput. Phys.* **103**, 16–42.

Y. Maday, S.M.O. Kaber and E. Tadmor (1993), 'Legendre PS viscosity methods for nonlinear conservation laws', *SIAM J. Numer. Anal.* **30**, 321–342.

A. Majda, J. McDonough and S. Osher (1978), 'The Fourier method for non-smooth initial data', *Math. Comput.* **32**, 1041–1081.

A.I. Markushevich (1967), *Theory of Functions of a Complex Variable*, Vol. III (Transl. by R.A. Silverman), Prentice Hall (New York).

B. Mercier (1989), *An Introduction to the Numerical Analysis of Spectral Methods*, Springer (Berlin).

W.J. Merryfield and B. Shizgal (1993), 'Properties of collocation third-derivative operators', *J. Comput. Phys.* **105**, 182–185.

R. Mittet, O. Holberg, B. Arntsen and L. Amundsen (1988), 'Fast finite difference modeling of 3-D elastic wave equation', *Society of Exploration Geophysics Expanded Abstracts* **I**, 1308–1311.

L.S. Mulholland and D.M. Sloan (1992), 'The role of preconditioning in the solution of evolutionary PDEs by implicit Fourier PS methods', *J. Comput. Appl. Math.* **42**, 157–174.

K.L. Nielsen (1956), *Methods in Numerical Analysis*, Macmillan (New York).

S.A. Orszag (1969), 'Numerical methods for the simulation of turbulence', *Phys. Fluids Suppl. II*, **12**, 250–257.

S.A. Orszag (1970), 'Transform method for calculation of vector coupled sums: Application to the spectral form of the vorticity equation', *J. Atmos. Sci.* **27**, 890–895.

S.A. Orszag (1972), 'Comparison of pseudospectral and spectral approximations', *Stud. Appl. Math.* **51**, 253–259.

T.N. Phillips (1988), 'On the Legendre coefficients of a general-order derivative of an infinitely differentiable function', *IMA J. Numer. Anal.* **8**, 455–459.

T.N. Phillips, T.A. Zang and M.Y. Hussaini (1986), 'Preconditioners for the spectral multigrid method', *IMA J. Numer. Anal.* **6**, 273–292.

M.J.D. Powell (1981), *Approximation Theory and Methods*, Cambridge University Press (Cambridge).

S.C. Reddy and L.N. Trefethen (1990), 'Lax-stability of fully discrete spectral methods via stability regions and pseudo-eigenvalues', *Comput. Meth. Appl. Mech. Engrg* **80**, 147–164.

R.D. Richtmyer and K.W. Morton (1967), *Difference Methods for Initial-value Problems*, 2nd Edn, Wiley (New York).

T.J. Rivlin (1969), *An Introduction to the Approximation of Functions*, Dover (New York).

J. Sand and O. Østerby (1979), 'Regions of absolute stability', Report DAIMI PP-102, Computer Science Department, Aarhus University, Denmark.

G. Sansone (1959), *Orthogonal Functions*, Interscience (New York).

L.F. Shampine and M.K. Gordon (1975), *Computer Solution of Ordinary Differential Equations*, W.H. Freeman (San Francisco).

A. Solomonoff (1992), 'A fast algorithm for spectral differentiation', *J. Comput. Phys.* **98**, 174–177.

A. Solomonoff (1994), 'Bayes finite difference schemes', *SIAM J. Num. Anal.*, submitted.

A. Solomonoff and E. Turkel (1989), 'Global properties of pseudospectral methods', *J. Comput. Phys.* **81**, 239–276.

F. Stenger (1981), 'Numerical methods based on Whittaker cardinal, or sinc functions', *SIAM Review* **23**, 165–224.

G. Szegő (1959), *Orthogonal Polynomials*, American Mathematical Society (Washington, D.C.).

E. Tadmor (1987), 'Stability analysis of finite-difference, pseudospectral and Fourier–Galerkin approximations for time-dependent problems', *SIAM Review* **29**, 525–555.

E. Tadmor (1989), 'Convergence of spectral methods for nonlinear conservation laws', *SIAM J. Num. Anal.* **26**, 30–44.

E. Tadmor (1990), 'Shock capturing by the spectral viscosity method', *Comput. Meth. Appl. Mech. Engrg* **80**, 197–208.

E. Tadmor (1993),'Superviscosity and spectral approximations of nonlinear conservation laws', in *Numerical Methods for Fluid Dynamics IV* (M.J. Baines and K.W. Morton, eds), Oxford University Press (Oxford), 69–82.

T.D. Taylor, R.S. Hirsh and N.M. Nadworny (1984), 'Comparison of FFT, direct inversion and conjugate gradient methods for use in pseudospectral methods', *Comput. Fluids* **12**, 1–9.

L.N. Trefethen (1983), 'Group velocity interpretation of the stability theory of Gustafsson, Kreiss and Sundström', *J. Comput. Phys.* **49**, 199–217.

L.N. Trefethen (1988), 'Lax-stability vs. eigenvalue stability of spectral methods', in *Numerical Methods for Fluid Dynamics III* (K.W. Morton and M.J. Baines, eds), Clarendon Press (Oxford), 237–253.

L.N. Trefethen and M.R. Trummer (1987), 'An instability phenomenon in spectral methods', *SIAM J. Numer. Anal.* **24**, 1008–1023.

L.N. Trefethen and J.A.C. Weideman (1991), 'Two results on polynomial interpolation in equally spaced points', *J. Approx. Theory* **65**, 247–260.

R.G. Voigt, D. Gottlieb and M.Y. Hussani (eds) (1984), *Spectral Methods for Partial Differential Equations*, SIAM (Philadelphia).

J.L. Walsh (1960), *Interpolation and Approximation by Rational Functions in the Complex Domain*, Colloquium Publications of the Amer. Math. Soc., 20, 3rd Edn.

J.A.C. Weideman (1992), 'The eigenvalues of Hermite and rational spectral DMs', *Numer. Math.* **61**, 409–432.

J.A.C. Weideman and L.N. Trefethen (1988), 'The eigenvalues of second-order spectral differentiation matrices', *SIAM J. Numer. Anal.* **25**, 1279–1298.

B.D. Welfert (1993), 'A remark on pseudospectral differentiation matrices', submitted to *SIAM J. Num. Anal.*

E.T. Whittaker (1915), 'On the functions which are represented by the expansions of the interpolation theory', *Proc. Roy. Soc. Edinburgh* **35**, 181–194.

J.M. Whittaker (1927), 'On the cardinal function of interpolation theory', *Proc. Edinburgh Math. Soc. Ser. 1*, **2**, 41–46.