JOURNAL OF COMPUTATIONAL PHYSICS 103, 16–42 (1992)

# Compact Finite Difference Schemes with Spectral-like Resolution

SANJIVA K. LELE*

*Center for Turbulence Research, NASA-Ames Research Center, MS 202A-1, Moffett Field, California 94035*

Received February 24, 1990; revised August 21, 1991

Finite difference schemes providing an improved representation of a range of scales (spectral-like resolution) in the evaluation of first, second, and higher order derivatives are presented and compared with well-known schemes. The schemes may be used on non-uniform meshes and a variety of boundary conditions may be imposed. Schemes are also presented for derivatives at mid-cell locations, for accurate interpolation and for spectral-like filtering. Applications to fluid mechanics problems are discussed. © 1992 Academic Press, Inc.

## 1. INTRODUCTION

Many physical phenomena possess a range of space and time scales, turbulent fluid flows being a common example. Direct numerical simulations of these processes require all the relevant scales to be properly represented in the numerical model. These requirements have led to the development of spectral methods [1–2]. Some examples of the direct simulation of turbulent flows by spectral methods may be found in [3–5]. The use of spectral methods is, however, limited to flows in simple domains and simple boundary conditions. These difficulties may be overcome by employing alternative numerical representations. For example, finite difference schemes or spectral (finite) element schemes may be used. Direct simulations of turbulent flows using these alternative schemes is relatively new. Rai and Moin [6, and references therein for earlier work] present simulations of a turbulent channel flow using a high-order, upwind-biased finite difference scheme. Work of Patera, Karniadakis, and their co-workers [7–9] illustrates the use of spectral element methods.

This paper presents finite difference schemes for use on problems with a range of spatial scales. Compared to the traditional finite difference approximations the schemes presented here provide a better representation of the shorter length scales. This feature brings them closer to the spectral methods, while the freedom in choosing the mesh geometry and the boundary conditions is maintained. The emphasis in this paper is on the resolution characteristics of the difference approximations rather than their formal accuracy (i.e., truncation error). By resolution characteristics we mean the accuracy with which the difference approximation represents the exact result over the full range of length scales that can be realized on a given mesh. This notion of resolution is quantified by means of a Fourier analysis of the differencing scheme. It is analogous to, but more general than, the notion of *intervals per wavelength* used by Swartz and Wendroff [10–13] and by Kreiss and Oliger [14] to compare the *resolving power* of different schemes. The notion of intervals per wavelength also uses Fourier analysis to quantify phase errors. For very small phase errors the number of intervals per wavelength needed by a differencing scheme is sensitive only to the behavior of the longest waves represented on a mesh. This is precisely the same information as obtained from the leading order truncation error (formal accuracy) of the scheme. It should be stressed that the quantitative importance of correctly resolving a particular range of length scales is dependent on the physical problem being solved as well as on the nature of results being sought from the numerical calculation.

The organization of the paper is as follows. Section 2 presents the basic schemes for approximating the first and second derivatives. Schemes for higher derivatives are described in Appendix A. Compact schemes on cell-centered mesh are discussed in Appendix B and the applications to interpolation and filtering are discussed in Appendix C. Section 3 presents analysis of the schemes, showing the associated dispersive errors and the anisotropy of the schemes in multi-dimensions. Comparisons with conventional finite difference schemes are made throughout these sections. This analysis leads to a definition of the resolving efficiency of the differencing schemes. Comments are also made on the aliasing errors encountered with nonlinear problems. Section 4 presents a treatment of boundaries in the derivative approximations. Assessment of the local boundary errors is presented. Its effect on the overall scheme is analyzed by means of numerical tests. An eigenvalue analysis of the complete scheme and the time-

* Present affiliation: Department of Mechanical Engineering and Department of Aeronautics and Astronautics, Stanford University.



0021-9991/92 $5.00

stepping restrictions for stability are also described in this section. General remaks on the application of the schemes are made in Section 5 and some example applications from fluid mechanics are presented.

## 2

### 2.1. *Approximation of First Derivative*

Given the values of a function on a set of nodes the finite difference approximation to the derivative of the function is expressed as a linear combination of the given function values. For simplicity consider a uniformly spaced mesh where the nodes are indexed by $i$. The independent variable at the nodes is $x_i = h(i-1)$ for $1 \leqslant i \leqslant N$ and the function values at the nodes $f_i = f(x_i)$ are given. The finite difference approximation $f'_i$ to the first derivative $(df/dx)(x_i)$ at the node $i$ depends on the function values at nodes near $i$. For second- and fourth-order central differences the approximation $f'_i$ depends on the sets $(f_{i-1}, f_{i+1})$ and $(f_{i-2}, f_{i-1}, f_{i+1}, f_{i+2})$, respectively. In the spectral methods, however, the value of $f'_i$ depends on all the nodal values. The Padé or compact finite difference schemes [15–19] mimic this global dependence. The schemes presented here are generalizations of the Padé scheme.

These generalizations are derived by writing approximations of the form:

$$\beta f'_{i-2} + \alpha f'_{i-1} + f'_i + \alpha f'_{i+1} + \beta f'_{i+2} = c\frac{f_{i+3} - f_{i-3}}{6h} + b\frac{f_{i+2} - f_{i-2}}{4h} + a\frac{f_{i+1} - f_{i-1}}{2h}. \tag{2.1}$$

The relations between the coefficients $a, b, c$ and $\alpha, \beta$ are derived by matching the Taylor series coefficients of various orders. The first unmatched coefficient determines the formal truncation error of the approximation (2.1). These constraints are:

$$a + b + c = 1 + 2\alpha + 2\beta \quad \text{(second order)} \tag{2.1.1}$$

$$a + 2^2 b + 3^2 c = 2\frac{3!}{2!}(\alpha + 2^2\beta) \quad \text{(fourth order)} \tag{2.1.2}$$

$$a + 2^4 b + 3^4 c = 2\frac{5!}{4!}(\alpha + 2^4\beta) \quad \text{(sixth order)} \tag{2.1.3}$$

$$a + 2^6 b + 3^6 c = 2\frac{7!}{6!}(\alpha + 2^6\beta) \quad \text{(eighth order)} \tag{2.1.4}$$

$$a + 2^8 b + 3^8 c = 2\frac{9!}{8!}(\alpha + 2^8\beta) \quad \text{(tenth order).} \tag{2.1.5}$$

If the dependent variables are periodic in $x$, then the system of relations (2.1) written for each node can be solved together as a linear system of equations for the unknown derivative values. This linear system is a cyclic pentadiagonal (tridiagonal) when $\beta$ is nonzero (zero). The general non-periodic case requires additional relations appropriate for the near boundary nodes. These are described in Sections 4.1 and 4.2. The resulting linear system is amenable to efficient numerical solution.

The relation (2.1), along with a mathematically defined mapping between a non-uniform physical mesh and a uniform computational mesh, provides derivatives on a non-uniform mesh. It is also possible to derive relations analogous to (2.1) for a non-uniform mesh directly (e.g., relations corresponding to the traditional Padé scheme were derived in [19–21]). We now consider the various special cases of (2.1). In the discussion below at least the first two of the constraints (2.1.1)–(2.1.5) are imposed. Thus all the schemes described have at least a fourth-order formal accuracy.

In Section 3.1 an analysis of the dispersive errors of schemes (2.1) is presented. This analysis shows the improved representation of the shorter length scales (i.e., spectral-like resolution) of the schemes presented here. The analysis also leads to schemes with very small dispersive errors (almost spectral). These are also presented in Section 3.1. In the present section we proceed in the traditional way to classify the differencing schemes generated by (2.1) in terms of the formal truncation error and the computational stencil required.

The general relation (2.1) with (2.1.1), (2.1.2) can be regarded as a three-parameter family of fourth-order schemes. If the schemes are restricted to $\beta = 0$ a variety of tridiagonal systems are obtained. For $\beta \neq 0$ pentadiagonal schemes are generated. If the additional constraint of sixth-order formal accuracy is imposed, a two-parameter family of sixth-order pentadiagonal schemes is obtained. These may be further specialized into a one-parameter family of eighth-order pentadiagonal schemes or a single tenth-order scheme.

First the tridiagonal schemes are described. These are generated by $\beta = 0$. If a further choice of $c = 0$ is made, a one-parameter ($\alpha$) family of fourth-order tridiagonal schemes is obtained. For these schemes

$$\beta = 0, \quad a = \tfrac{2}{3}(\alpha + 2), \quad b = \tfrac{1}{3}(4\alpha - 1), \quad c = 0. \tag{2.1.6}$$

The truncation error on the r.h.s. of (2.1) (unless stated otherwise the term truncation error will be used in this sense from here on) for this scheme and for other schemes to be described below are listed in Table I. The stencil sizes indicated in the table are the maximum stencil sizes needed within a class of schemes.

As $\alpha \to 0$ this family merges into the well-known fourth-order central difference scheme. Similarly for $\alpha = \frac{1}{4}$ the classical Padé scheme is recovered. Furthermore, for $\alpha = \frac{1}{3}$

**TABLE I**

Truncation Error for the First Derivative Schemes

| Scheme | Max. l.h.s. stencil size | Max. r.h.s. stencil size | Truncation error in (2.1) |
|---|---|---|---|
| (2.1.6) | 3 | 5 | $\frac{4}{5!}(3\alpha-1)\,h^4 f^{(5)}$ |
| (2.1.7) | 3 | 5 | $\frac{4}{7!}\,h^6 f^{(7)}$ |
| (2.1.8) | 3 | 7 | $\frac{12}{7!}(-8\alpha+3)\,h^6 f^{(7)}$ |
| (2.1.8) & $\alpha=\frac{3}{8}$ | 3 | 7 | $\frac{-36}{9!}\,h^8 f^{(9)}$ |
| (2.1.9) | 5 | 7 | $\frac{4}{5}(-1+3\alpha-12\beta+10c)\,h^4 f^{(5)}$ |
| (2.1.10) | 5 | 7 | $\frac{12}{7!}(3-8\alpha+20\beta)\,h^6 f^{(7)}$ |
| (2.1.11) | 5 | 5 | $\frac{4}{7!}(9\alpha-4)\,h^6 f^{(7)}$ |
| (2.1.12) | 5 | 5 | $-\frac{16}{9!}\,h^8 f^{(9)}$ |
| (2.1.13) | 5 | 7 | $\frac{144}{9!}(2\alpha-1)\,h^8 f^{(9)}$ |
| (2.1.14) | 5 | 7 | $\frac{144}{11!}\,h^{10} f^{(11)}$ |

the leading order truncation error coefficient vanishes and the scheme is formally sixth-order accurate. Its coefficients are

$$\alpha=\tfrac{1}{3},\quad \beta=0,\quad a=\tfrac{14}{9},\quad b=\tfrac{1}{9},\quad c=0. \qquad (2.1.7)$$

The specific tridiagonal schemes obtained for $\alpha=\frac{1}{4}$ and $\alpha=\frac{1}{3}$ were given by Collatz [22, p. 538].

With $\beta=0$ and $c\neq 0$ the family of schemes (2.1.6) is extended to a two-parameter family of fourth-order tridiagonal schemes. Contained within these is a one-parameter family of sixth-order tridiagonal schemes. For this (sixth-order) family

$$\beta=0,\quad a=\tfrac{1}{6}(\alpha+9),\quad b=\tfrac{1}{15}(32\alpha-9),\quad c=\tfrac{1}{10}(-3\alpha+1). \qquad (2.1.8)$$

The sixth-order tridiagonal scheme (2.1.7) is a member of this family (with $c=0$, $\alpha=\frac{1}{3}$). This sixth-order family can be further specialized into an eighth-order scheme by choosing $\alpha=\frac{3}{8}$. This is the tridiagonal scheme ($\beta=0$) with the highest formal accuracy within (2.1).

Pentadiagonal schemes are generated with $\beta\neq 0$. In general this fourth-order three-parameter ($\alpha$, $\beta$, and $c$) family is given by

$$a=\tfrac{1}{3}(4+2\alpha-16\beta+5c),\quad b=\tfrac{1}{3}(-1+4\alpha+22\beta-8c). \qquad (2.1.9)$$

Schemes of sixth-order formal accuracy contain two parameters $\alpha$ and $\beta$. They are given by

$$a=\tfrac{1}{6}(9+\alpha-20\beta),\quad b=\tfrac{1}{15}(-9+32\alpha+62\beta),\quad c=\tfrac{1}{10}(1-3\alpha+12\beta). \qquad (2.1.10)$$

The tridiagonal sixth-order family of (2.1.8) is a subclass within (2.1.10). Another subclass is obtained with $\beta\neq 0$ and $c=0$. This sixth-order pentadiagonal family has

$$\beta=\tfrac{1}{12}(-1+3\alpha),\quad a=\tfrac{2}{9}(8-3\alpha),\quad b=\tfrac{1}{18}(-17+57\alpha),\quad c=0. \qquad (2.1.11)$$

This family limits to the sixth-order tridiagonal scheme (2.1.7) as $\beta\to 0$ or $\alpha=\frac{1}{3}$. The leading truncation error coefficient for (2.1.11) vanishes for $\alpha=\frac{4}{9}$ yielding an eighth-order scheme. This eighth-order scheme has

$$\alpha=\tfrac{4}{9},\quad \beta=\tfrac{1}{36},\quad a=\tfrac{40}{27},\quad b=\tfrac{25}{54},\quad c=0. \qquad (2.1.12)$$

This particular scheme is also given by Collatz [22, p. 538] and analyzed by Swartz and Wendroff [10–13].

By choosing $\beta=\frac{1}{20}(-3+8\alpha)$ in (2.1.10) a one-parameter family of eighth-order pentadiagonal schemes is generated. This eighth-order family has

$$\beta=\tfrac{1}{20}(-3+8\alpha),\quad a=\tfrac{1}{6}(12-7\alpha),\quad b=\tfrac{1}{150}(568\alpha-183),\quad c=\tfrac{1}{50}(9\alpha-4). \qquad (2.1.13)$$

The specific eighth-order schemes obtained earlier viz., (a) scheme (2.1.8) with $\alpha=\frac{3}{8}$ and (b) scheme (2.1.12), belong to this one-parameter family.

By choosing $\alpha=\frac{1}{2}$ in (2.1.13) a tenth-order scheme is generated. This is the scheme with the highest formal accuracy amongst the schemes defined by (2.1). The coefficients of this scheme are

$$\alpha=\tfrac{1}{2},\quad \beta=\tfrac{1}{20},\quad a=\tfrac{17}{12},\quad b=\tfrac{101}{150},\quad c=\tfrac{1}{100}. \qquad (2.1.14)$$

Among the class of derivative approximations represented by (2.1) those which achieve the highest possible formal accuracy within each subclass of schemes (denoted by a specified computational stencil on both the l.h.s. and r.h.s. of (2.1)) are precisely the schemes obtained by a rational (or Padé) approximation of the first derivative

operator.[1] This Padé table is given by Kopal [23]. The formal rational approximation generates only these specific members of the multiparameter scheme (2.1.9).

An alternate and more effective way of classifying the schemes presented here is provided by their Fourier analysis. The Fourier analysis also quantifies the resolution characteristics of the schemes. It also provides a way to "optimize" the scheme from a multi-parameter family. These issues are fully discussed in Section 3.1.

### 2.2. *Approximation of Second Derivative*

The derivation of compact approximations for the second derivative proceeds exactly analogous to the first derivative. Once again the starting point is a relation of the form

$$\begin{aligned}\beta f''_{i-2}+\alpha f''_{i-1}+f''_i+\alpha f''_{i+1}+\beta f''_{i+2}&\\ =c\frac{f_{i+3}-2f_i+f_{i-3}}{9h^2}+b\frac{f_{i+2}-2f_i+f_{i-2}}{4h^2}&\\ +a\frac{f_{i+1}-2f_i+f_{i-1}}{h^2},&\end{aligned}\tag{2.2}$$

where $f''_i$ represents the finite difference approximation to the second derivative at node $i$. The relations between the coefficients $a, b, c$ and $\alpha, \beta$ are derived by matching the Taylor series coefficients of various orders. The first unmatched coefficient determines the formal truncation error of the approximation (2.2). These constraints are:

$$a+b+c=1+2\alpha+2\beta \quad \text{(second order)} \tag{2.2.1}$$

$$a+2^2b+3^2c=\frac{4!}{2!}(\alpha+2^2\beta) \quad \text{(fourth order)} \tag{2.2.2}$$

$$a+2^4b+3^4c=\frac{6!}{4!}(\alpha+2^4\beta) \quad \text{(sixth order)} \tag{2.2.3}$$

$$a+2^6b+3^6c=\frac{8!}{6!}(\alpha+2^6\beta) \quad \text{(eighth order)} \tag{2.2.4}$$

$$a+2^8b+3^8c=\frac{10!}{8!}(\alpha+2^8\beta) \quad \text{(tenth order).} \tag{2.2.5}$$

The form of these constraints is very close to those derived for the first derivative approximations but the multiplying factors on the r.h.s. are different. In the following discussion at least the first two of these constraints are imposed resulting in schemes with at least a fourth-order formal accuracy. For dependent variables which are periodic in $x$ the tridiagonal or pentadiagonal system defined by (2.2) at each node may be solved to yield the second derivatives. For the non-periodic case additional relations are required at the boundary (presented in Section 4.3).

By choosing $\beta=0$ and $c=0$ a one-parameter family of fourth-order schemes is generated. This family has

$$\beta=0, \quad c=0, \quad a=\tfrac{4}{3}(1-\alpha), \quad b=\tfrac{1}{3}(-1+10\alpha). \tag{2.2.6}$$

The truncation error on the r.h.s. of (2.2) for this and other schemes discussed in this section are listed in Table II. It may be noted that as $\alpha\to 0$ this family coincides with the well-known fourth-order central difference scheme. For $\alpha=\frac{1}{10}$ the classical Padé scheme is recovered. For $\alpha=\frac{2}{11}$ a sixth-order tridiagonal scheme is obtained. This scheme has

$$\alpha=\tfrac{2}{11}, \quad \beta=0, \quad a=\tfrac{12}{11}, \quad b=\tfrac{3}{11}, \quad c=0. \tag{2.2.7}$$

The particular members obtained with $\alpha=\frac{1}{10}$ and $\alpha=\frac{2}{11}$ were given by Collatz [22, p. 538].

A three-parameter family of fourth-order schemes is generated from (2.2) by considering $\beta\neq 0$ and $c\neq 0$. These satisfy

$$\begin{aligned}a&=\tfrac{1}{3}(4-4\alpha-40\beta+5c),\\ b&=\tfrac{1}{3}(-1+10\alpha+46\beta-8c).\end{aligned}\tag{2.2.8}$$

This class of schemes can be further specialized into a two-parameter family of sixth-order schemes, a one-parameter family of eighth-order schemes or a single tenth-

**TABLE II**

Truncation Error for Second Derivative Schemes

| Scheme | Max. l.h.s. stencil size | Max. r.h.s. stencil size | Truncation error in (2.2) |
|---|---|---|---|
| (2.2.6) | 3 | 5 | $\frac{-4}{6!}(11\alpha-2)\,h^4f^{(6)}$ |
| (2.2.7) | 3 | 5 | $\frac{-8\cdot 23}{11\cdot 8!}h^6f^{(8)}$ |
| (2.2.8) | 5 | 7 | $\frac{-4}{6!}(-2+11\alpha-124\beta+20c)\,h^4f^{(6)}$ |
| (2.2.9) | 5 | 7 | $\frac{-8}{8!}(9-38\alpha+214\beta)\,h^6f^{(8)}$ |
| (2.2.10) | 5 | 7 | $\frac{899\alpha-334}{2696400}h^8f^{(10)}$ |
| (2.2.11) | 5 | 7 | $\frac{619}{299043360}h^{10}f^{(12)}$ |

[1] The author is grateful to Dr. K. Shariff for explicitly verifying this equivalence (via the use of MACSYMA) and pointing out some errors in the Padé table of Kopal [23]. The coefficients of $\delta^4$ and $\delta^6$ in the expressions for $D_2^{(0)}$ and $D_3^{(0)}$ (on p. 553) are in error, they should read as $\frac{1}{180}$ and $\frac{1}{1512}$, respectively.

order scheme. The two-parameter sixth-order family is defined by

$$a=\frac{6-9\alpha-12\beta}{4},\quad b=\frac{-3+24\alpha-6\beta}{5},\quad c=\frac{2-11\alpha+124\beta}{20}. \tag{2.2.9}$$

When the eighth-order constraint is imposed (2.2.9) reduces to a one-parameter family of eighth-order schemes. These are defined by

$$\beta=\frac{38\alpha-9}{214},\quad a=\frac{696-1191\alpha}{428},\quad b=\frac{2454\alpha-294}{535},\quad c=\frac{1179\alpha-344}{2140}. \tag{2.2.10}$$

The particular eighth-order scheme with $\alpha=\frac{344}{1179}$ corresponding to $c=0$ in (2.2.10) was given by Collatz [22, p. 539].

Finally, when the tenth-order constraint is imposed a single tenth-order scheme is obtained. This scheme defined by

$$\beta=\tfrac{43}{1798},\quad \alpha=\tfrac{334}{899},\quad a=\tfrac{1065}{1798},\quad b=\tfrac{1038}{899},\quad c=\tfrac{79}{1798} \tag{2.2.11}$$

has the highest formal accuracy within the class of schemes defined by (2.2).

As before amongst the schemes defined by (2.2), those which maximize the formal accuracy (with a prescribed computational stencil) correspond precisely to the rational or Padé approximation of the second derivative operator. These have been given by Kopal [23, pp. 551–552] in operator notation. A comparison of schemes (2.2) by means of Fourier analysis is presented in Section 3.1, where comparisons are also made with other well-known schemes. This analysis brings out the spectral-like resolution of the schemes described here and also leads to more "optimal" schemes.

Compact schemes for evaluating higher derivatives are described in Appendix A. Schemes involving a cell-centered mesh in evaluating derivatives are described in Appendix B. Compact schemes for interpolation and filtering are presented in Appendix C.

## 3. FOURIER ANALYSIS OF ERRORS

This section presents a Fourier analysis of the errors associated with the approximations introduced in the last two sections. Comparison are made with the standard finite-difference schemes to judge the improvement in the error characteristics. Formal truncation error of the differencing schemes were given in the preceding sections. The use of Fourier analysis to characterize the errors of difference approximations is described extensively in [24]. It is a classical technique for comparing differencing schemes. It was used by Roberts and Weiss [25], Fromm [26], Oliger and Kreiss [14], Orszag [27–28] and by Swartz and Wendroff [10–13]. Fourier analysis of the standard Padé scheme was presented in [18] and comparisons were made with the second- and fourth-order central differences.

The Fourier analysis provides an effective way to quantify the resolution characteristics of the differencing approximations. This quantification may be used to further guide an optimization of the differencing schemes. In the following section the differencing errors are analyzed in terms of dispersion or phase error and anisotropy (in multi-dimensions). Comparisons are made throughout with the standard difference formulae. Examples of optimization of the schemes based on the resolution characteristics are also presented. All differencing approximations (for the interior nodes) studied here are of central difference form, thus there are no dissipative errors (from the differencing of conservative terms). The treatment of boundary errors (for non-periodic problems) is presented in Section 4, where the approximations appropriate for the near boundary nodes are also introduced. Local errors introduced by the boundary scheme are discussed along with the schemes and their effect on the global accuracy is presented in Section 4.4. This requires direct numerical tests on the performance of the first derivative schemes. Analysis of the stability properties of the overall scheme is deferred to Section 4.5 which is followed by a summary of the time step restrictions for stable explicit time advancement in Section 4.6.

### 3.1. *Fourier Analysis of Differencing Errors*

For the purposes of Fourier analysis the dependent variables are assumed to be periodic over the domain $[0, L]$ of the independent variable, i.e., $f_1 = f_{N+1}$ and $h = L/N$. The dependent variables may be decomposed into their Fourier coefficients

$$f(x)=\sum_{k=-N/2}^{k=N/2}\hat{f}_k\exp\left(\frac{2\pi ikx}{L}\right), \tag{3.1.1}$$

where $i=\sqrt{-1}$. Since the dependent variables are real-valued, the Fourier coefficients satisfy $\hat{f}_k=\hat{f}^*_{-k}$ for $1\leqslant k\leqslant N/2$ and $\hat{f}_0=\hat{f}^*_0$, where $*$ denotes the complex conjugate.

It is convenient to introduce a scaled wavenumber $w=2\pi kh/L=2\pi k/N$ and a scaled coordinate $s=x/h$. The Fourier modes in terms of these are simply $\exp(iws)$. The domain of the scaled wavenumber $w$ is $[0, \pi]$. The exact first derivative of (3.1.1) (with respect to $s$) generates a func-

tion with Fourier coefficients $\hat{f}'_k = iw\hat{f}_k$. The differencing error of the first derivative scheme may be assessed by comparing the Fourier coefficients of the derivative obtained from the differencing scheme $(\hat{f}'_k)_{fd}$ with the exact Fourier coefficients $\hat{f}'_k$. For central difference schemes it may be shown that $(\hat{f}'_k)_{fd} = iw'\hat{f}_k$, where the modified wavenumber $w'$ is real-valued. Each finite difference scheme corresponds to a particular function $w'(w)$. Exact differentiation corresponds to the straight line $w' = w$. Spectral methods provide $w' = w$ for $w \neq \pi$ (and $w' = 0$ for $w = \pi$). The range of wavenumbers $[2\pi/N, w_f]$ over which the modified wavenumber $w'(w)$ approximates the exact differentiation $w'(w) = w$ within a specified error tolerance defines the set of *well*-resolved waves. While, the value $w_f$, i.e., the shortest well resolved wave, certainly depends on the specific error tolerance it is quite reasonable to keep this error tolerance fixed when different finite difference schemes are compared. It should also be noted that $w_f$ depends only on the scheme and not on the number of points $N$ used in the descretization. In the following the error tolerance is defined as:

$$\frac{|w'(w) - w|}{w} \leqslant \varepsilon, \tag{3.1.3}$$

The fraction $r_1 \equiv 1 - w_f/\pi$ represents the fraction of poorly resolved waves for the first derivative scheme. This fraction is also independent of the number of points $N$. The fraction $e_1 \equiv w_f/\pi = 1 - r_1$ may be regarded as a measure of the *resolving efficiency* of a scheme. We note that the computational efficiency of a scheme is proportional to the resolving efficiency but also depends on the operation count of the numerical algorithm and its implementation. The leading order operation count for the spectral-like scheme described later in this section is $7N$ multiplies, $N$ divides, and $7N$ addition or subtraction operations when sparse matrix techniques are used [46] and the Cholesky decomposition of the symmetric portion of the associated matrix is computed (in $\mathbf{LDL}^T$ form) and saved for future use. For the tridiagonal schemes the operation count is $5N$, $N$, and $5N$ for multiply, divide, and add/subtract operations, respectively. For reference, a radix 2 FFT [47] requires $2N \log_2 N$ multiplies and $3N \log_2 N$ adds.

The difference schemes (2.1) correspond to

$$w'(w) = \frac{a \sin(w) + (b/2)\sin(2w) + (c/3)\sin(3w)}{1 + 2\alpha\cos(w) + 2\beta\cos(2w)}. \tag{3.1.4}$$

Plots of the modified wavenumber $w'$ against wavenumber $w$ are presented in Fig. 1 for a variety of schemes. In this manner the resolution characteristics of different schemes can be compared. From this plot the fraction $r_1$, representing the fraction of poorly resolved waves and the resolving efficiency $e_1 = 1 - r_1$ is determined. This is done for three different values of the error tolerance $\varepsilon$, viz., $\varepsilon = 0.1$ 0.01, and 0.001. The results quantify the resolution characteristics of the schemes and are tabulated in Table. III.

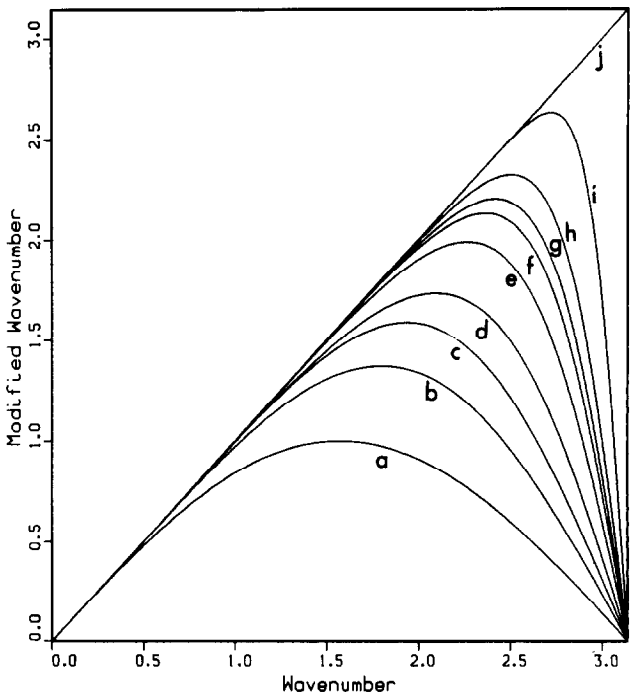


**FIG. 1.** Plot of modified wavenumber vs wavenumber for first derivative approximations: (a) second-order central differences; (b) fourth-order central differences; (c) sixth-order central differences; (d) standard Padé scheme ($\beta = 0 = c$, $\alpha = \frac{1}{2}$); (e) sixth-order tridiagonal scheme ($\beta = 0 = c$, $\alpha = \frac{1}{3}$); (f) eighth-order tridiagonal scheme ($\beta = 0$); (g) eighth-order pentadiagonal scheme ($c = 0$); (h) tenth-order pentadiagonal scheme; (i) spectral-like pentadiagonal scheme (3.1.6); (j) exact differentiation.

It is evident that compared to the standard second- and fourth-order central differences the compact schemes stay close to the exact differentiation over a wider range of wavenumbers. The tridiagonal sixth-order scheme (2.1.7) is better than the standard Padé scheme. Similarly, the eighth-

**TABLE III**

Resolving Efficiency $e_1(\varepsilon)$ of the First Derivative Schemes Shown in Fig. 1

| Scheme | $\varepsilon = 0.1$ | $\varepsilon = 0.01$ | $\varepsilon = 0.001$ |
|---|---|---|---|
| (a) | 0.25 | 0.08 | 0.02 |
| (b) | 0.44 | 0.23 | 0.13 |
| (c) | 0.54 | 0.35 | 0.23 |
| (d) | 0.59 | 0.35 | 0.20 |
| (e) | 0.70 | 0.50 | 0.35 |
| (f) | 0.75 | 0.58 | 0.44 |
| (g) | 0.77 | 0.61 | 0.48 |
| (h) | 0.81 | 0.68 | 0.56 |
| (i) | 0.90 | 0.83 | 0.79 |

order schemes (2.1.8) with $\alpha = \frac{3}{8}$, (2.1.11), and the tenth-order scheme (2.1.13) stay close to the exact differentiation over a progressively larger wavenumber range. It may also be noted that the improvements in the dispersive error for wavenumber range $(\pi/2, \pi)$ are not very sensitive to the reduction of the formal truncation error of the scheme. Also shown on Fig. 1 is a "spectral-like" scheme obtained from (2.1). This scheme has a formal fourth-order accuracy but considerably better resolution characteristics. For the purposes of constructing "spectral-like" schemes the following constraints were imposed:

$$w'(w_1) = w_1, \quad w'(w_2) = w_2, \quad w'(w_3) = w_3. \quad (3.1.5)$$

The scheme shown in Fig. 1 corresponds to $w_1 = 2.2$, $w_2 = 2.3$, and $w_3 = 2.4$. Its parameters are

$$\alpha = 0.5771439, \quad \beta = 0.0896406, \quad a = 1.3025166, \quad b = 0.9935500, \quad c = 0.03750245. \quad (3.1.6)$$

No attempt was made to optimize the choices for $w_1$, $w_2$, $w_3$. Schemes obtained for other choices of $w_1, w_2, w_3$ also share the characteristics of scheme (3.1.6).

It is possible to optimize the scheme within a particular family of schemes (say defined by a given choice of the computational stencil on the l.h.s. and r.h.s of (2.1)). This is illustrated in Fig. 2 for the tridiagonal schemes defined by (2.1.6). Estimates for the fraction of poorly resolved waves $r_1$ and the resolving efficiency $e_1$ are tabulated in Table IV. Evidently, the member with $\alpha = \frac{5}{14}$ (fourth-order scheme (d)) provides a better resolution than the sixth-order scheme obtained with $\alpha = \frac{1}{3}$ (scheme (c)). We anticipate this to be a general feature of difference approximations. Other specific examples which illustrate this behaviour may be found elsewhere in this paper.

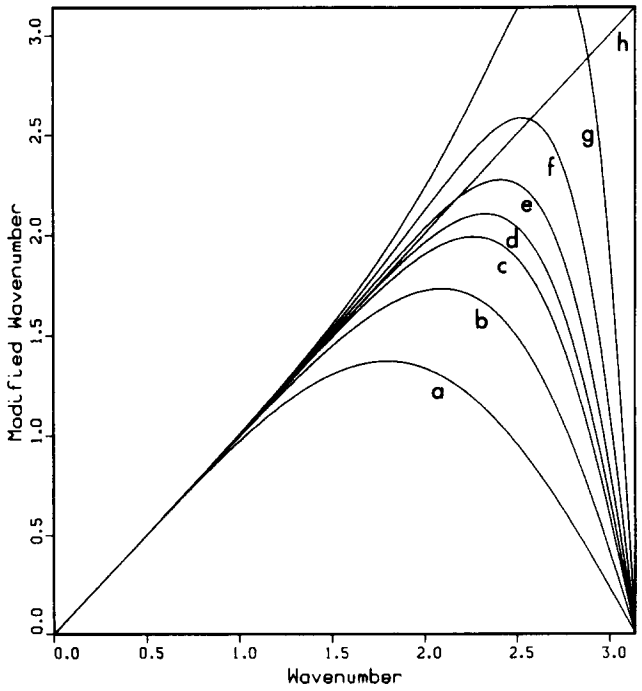


**FIG. 2.** Plot of modified wavenumbers vs wavenumber for first derivative approximations: (a) fourth-order central differences; (b) standard Padé scheme ($\beta = 0 = b = c$, $\alpha = \frac{1}{4}$); (c) sixth-order tridiagonal scheme ($\beta = 0 = c$, $\alpha = \frac{1}{3}$); (d) fourth-order tridiagonal scheme ($\beta = 0 = c$, $\alpha = \frac{5}{14}$); (e) fourth-order tridiagonal scheme ($\beta = 0 = c$, $\alpha = \frac{5}{13}$); (f) fourth-order tridiagonal scheme ($\beta = 0 = c$, $\alpha = \frac{5}{12}$); (g) fourth-order tridiagonal scheme ($\beta = 0 = c$, $\alpha = \frac{5}{11}$); (h) exact differentiation.

**TABLE IV**

Resolving Efficiency $e_1(\varepsilon)$ of the First Derivative Schemes Shown in Fig. 2

| Scheme | $\varepsilon = 0.1$ | $\varepsilon = 0.01$ | $\varepsilon = 0.001$ |
|---|---|---|---|
| (a) | 0.44 | 0.23 | 0.13 |
| (b) | 0.59 | 0.35 | 0.20 |
| (c) | 0.70 | 0.50 | 0.35 |
| (d) | 0.74 | 0.59 | 0.52 |
| (e) | 0.79 | 0.46 | 0.24 |
| (f) | 0.86 | 0.39 | 0.21 |
| (g) | 0.61 | 0.35 | 0.20 |

The dispersive error characteristics can be alternatively presented in terms of the error in the phase speed of waves of different wavenumber. It may be shown by considering the semi-discrete (exact time advancement) form of the advection equation

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} = 0 \quad (3.1.7)$$

that the phase speed for a wave of wavenumber $w$ is given by the finite difference scheme as $(c_p)_{fd} = w'(w)/w$. The partial differential equation (3.1.7) has the phase speed one for all wavenumbers, thus $(c_p)_{fd} - 1$ is the measure of phase error.[2] Figure 3 presents this information for a variety of finite difference schemes. Once again the improved phase error of the compact schemes is evident. Again, schemes with spectral-like resolution can be generated from (2.1) by not insisting on the highest possible formal accuracy.

In multi-dimensional problems the phase errors of the differencing schemes also appear in the form of anisotropy [24]. To illustrate this effect consider the 1D advection equation

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial l} = 0, \quad (3.1.8)$$

[2] For some of the schemes discussed in this paper the modified wavenumber $w'(w)$ exceeds $w$ over some intermediate range of wavenumbers. The phase speeds for this range of wavenumbers exceeds the exact phase speed and their phase error leads to the exact phase.

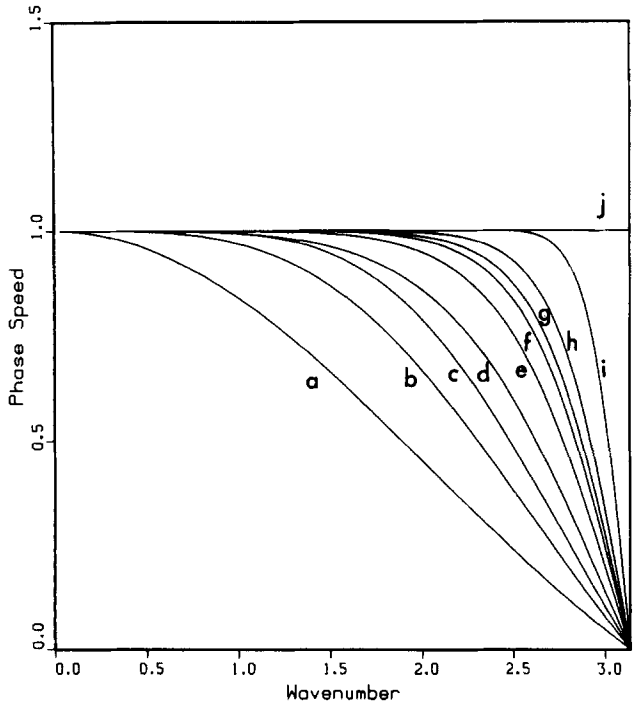


**FIG. 3.** Plot of phase speed vs wavenumber for first derivative approximations: (a) second-order central differences; (b) fourth-order central differences; (c) sixth-order central differences; (d) standard Padé scheme ($\beta = 0 = b = c$, $\alpha = \frac{1}{4}$); (e) sixth-order tridiagonal scheme ($\beta = 0 = c$, $\alpha = \frac{1}{3}$); (f) eighth-order tridiagonal scheme ($\beta = 0$); (g) eight-order pentadiagonal scheme ($c = 0$); (h) tenth-order pentadiagonal scheme; (i) spectral-like pentadiagonal scheme (3.1.6); (j) exact differentiation.

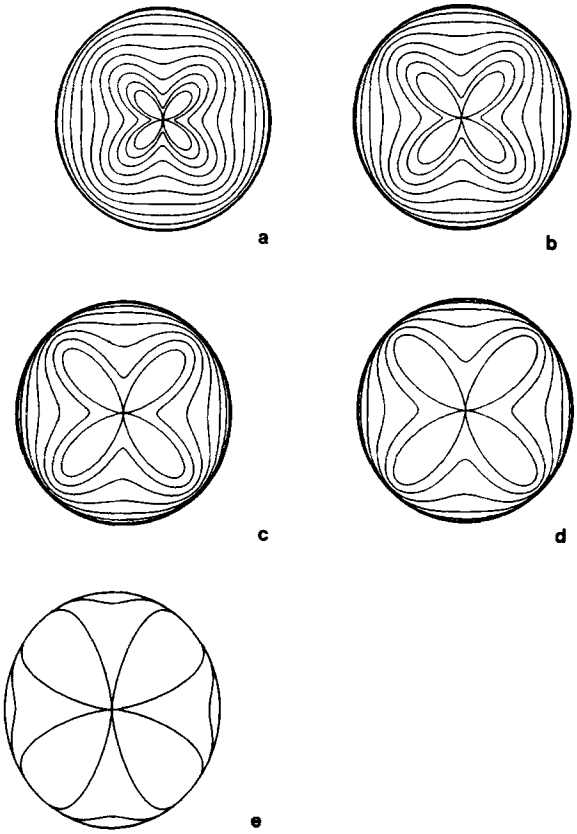


**FIG. 4.** Polar plot of phase speed anisotropy for first derivative approximations; the phase speed for wavenumber (magnitude) $w/\pi = \frac{1}{50}, \frac{5}{50}, \ldots, \frac{45}{50}, \frac{50}{50}$ are plotted: (a) second-order central differences; (b) fourth-order central differences; (c) standard Padé scheme ($\beta = 0 = b = c$, $\alpha = \frac{1}{4}$); (d) sixth-order tridiagonal scheme ($\beta = 0 = c$, $\alpha = \frac{1}{3}$); (e) spectral-like pentadiagonal scheme (3.1.6).

where $l$ is an arbitrary direction on a two-dimensional grid. While (3.1.8) has a phase speed of unity for all waves (any wavenumber and any orientation) the finite difference schemes generate different phase speeds for waves of differing wavenumber and orientation. It may be shown that

$$(c_p)_{fd}(w, \theta) = \frac{\cos\theta w'(w\cos\theta) + \sin\theta w'(w\sin\theta)}{w}, \quad (3.1.9)$$

where $\theta$ is the angle between the propagation direction and the $x$ axis. This anisotropic propagation is displayed in Fig. 4 for several finite difference schemes. The curves in this figure are polar plots of $(c_p)_{fd}$ at fixed wavenumber $w$. For each curve the radial distance at an angle $\theta$ represents $(c_p)_{fd}$ obtained for waves propagating in that direction. Curves are plotted for $w/\pi = \frac{1}{50}, \frac{5}{50}, \ldots, \frac{45}{50}, \frac{50}{50}$. The outermost curves (circles) correspond to small $w$, i.e., well resolved waves. For these waves the propagation is isotropic and phase speed is very close to unity. Shorter waves (larger $w$), usually, have smaller phase speeds and the propagation is anisotropic (with least error along $\pm 45°$ angles). The innermost curves correspond to the shortest waves resolved on the mesh. It may be seen that in compact schemes the anisotropic error is limited to a narrower range of short waves. For the spectral-like scheme the anisotropy is felt only by the shortest 20% of the waves (for any $\theta$).

We remind the reader that the improved resolution properties of the schemes described here also lead to the possibility of increased aliasing errors [1, 2, 27, 31, 32] when solving nonlinear problems. The relative importance of aliasing errors compared to pure differentiation errors depends on the nature of the physical problem (e.g., energy content of high wavenumbers compared to the most energetic scales), on the type of nonlinearity, (spatial) dimensionality of the problem, as well as on the specific numerical algorithms. If all scales are well resolved then the differentiation error may be expected to be more dominant. With marginal resolution of the short scales the aliasing errors may be more significant. Rather than attempting to analyze this complex issue in detail we remind the reader that since the differencing schemes presented here are spectral-like their aliasing behaviour may be expected to be akin to the spectral algorithms. When aliasing errors are dominant they may be removed either by following the

algorithms developed for the spectral methods (dealiasing by the use of FFTs, use of different forms of the conservation equations) [32–35] for periodic problems or by use of spectral-like filtering techniques (described in Section 6.2) for non-periodic cases. An alternative to these is provided in [6] via the use of a high order upwind biased scheme. More research is, however, needed to establish the best way of aliasing control.

The error analysis for the second and higher derivative approximations proceeds similarly to the analysis for the first derivative. The exact second derivative of (3.1.1) (with respect to $s$) generates a function with Fourier coefficients $\hat{f}''_k = -w^2\hat{f}_k$. The numerical approximations (2.2) correspond to $(\hat{f}''_k)_{fd} = -w''\hat{f}_k$, where

$$w''(w) = \frac{\begin{pmatrix} 2a(1-\cos(w)) + (b/2)(1-\cos(2w)) \\ + (2c/9)(1-\cos(3w)) \end{pmatrix}}{1 + 2\alpha\cos(w) + 2\beta\cos(2w)}. \tag{3.1.10}$$

The difference between $w''(w)$ and $w^2$ is a measure of error in the second derivative approximation. Plots of $w''(w)$ vs $w$ for different finite difference schemes are presented in Fig. 5.

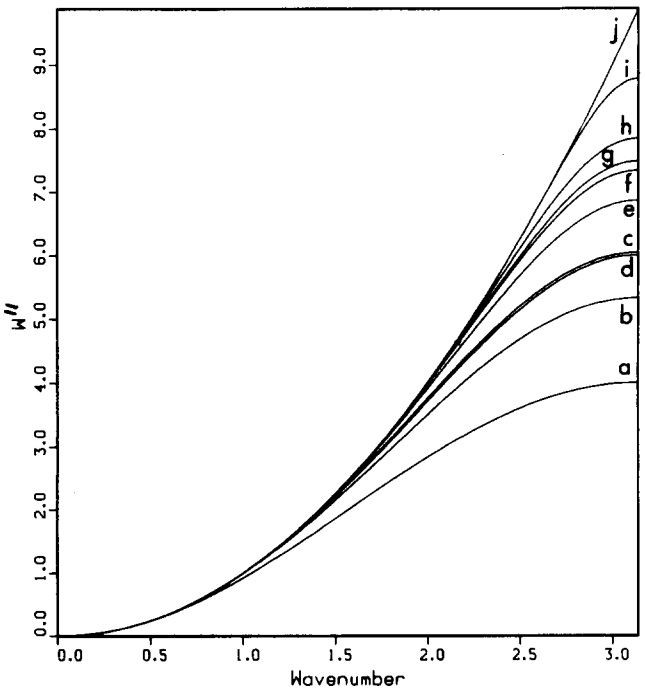


**FIG. 5.** Differencing error for second derivative vs wavenumber: (a) second-order central differences; (b) fourth-order central differences; (c) fourth-order central differences; (d) standard Padé scheme ($\beta = 0 = b = c$, $\alpha = \frac{1}{10}$); (e) sixth-order tridiagonal scheme ($\beta = 0 = c$, $\alpha = \frac{2}{11}$); (f) eighth-order tridiagonal scheme ($\beta = 0$); (g) eighth-order pentadiagonal scheme ($c = 0$); (h) tenth-order pentadiagonal scheme; (i) spectral-like pentadiagonal scheme (3.1.12); (j) exact differentiation.

The range of wavenumbers $[2\pi/N, w_s]$ which are well resolved may be defined by the error tolerance:

$$\frac{|w''(w) - w^2|}{w^2} \leqslant \varepsilon. \tag{3.1.11}$$

The fraction $r_2 \equiv 1 - w_s/\pi$ represents the fraction of poorly resolved waves for the second derivative scheme. This fraction is also independent of the number of points $N$. The fraction $e_2 \equiv w_s/\pi = 1 - r_2$ may be regarded as a measure of the resolving efficiency of a scheme. These estimates are tabulated in Table V for the schemes plotted on Fig. 5. The improvement of the compact schemes over the explicit central differences is evident. The largest differencing error which occurs near $w = \pi$ depends only weakly on the formal accuracy of the scheme. Once again spectral-like schemes can be constructed via a different optimization. One such scheme is also displayed in this figure. This scheme has a fourth-order formal accuracy. Its coefficients are given by

$$\alpha = 0.50209266, \quad \beta = 0.05569169, \quad a = 0.21564935,$$
$$b = 1.7233220, \quad c = 0.17659730. \tag{3.1.12}$$

For the purposes of constructing "spectral-like" schemes the following constraints were imposed on (2.1.4):

$$w''(w = w_1) = w_1^2, \quad w''(w = w_2) = w_2^2, \quad w''(w = w_3) = w_3^2. \tag{3.1.13}$$

The scheme shown in Fig. 5 was obtained with $w_1 = 2$, $w_2 = 2.4$, and $w_3 = 2.6$. These values of $(w_1, w_2, w_3)$ represent just one particular example. Schemes obtained for other choices of $(w_1, w_2, w_3)$ shared the same characteristics as the scheme (3.1.12).

**TABLE V**

Resolving Efficiency $e_2(\varepsilon)$ of the Second Derivative Schemes shown in Fig. 5

| Scheme | $\varepsilon = 0.1$ | $\varepsilon = 0.01$ | $\varepsilon = 0.001$ |
|---|---|---|---|
| (a) | 0.35 | 0.11 | 0.03 |
| (b) | 0.59 | 0.31 | 0.17 |
| (c) | 0.70 | 0.44 | 0.29 |
| (d) | 0.68 | 0.39 | 0.22 |
| (e) | 0.80 | 0.55 | 0.38 |
| (f) | 0.86 | 0.64 | 0.48 |
| (g) | 0.89 | 0.66 | 0.51 |
| (h) | 0.91 | 0.73 | 0.59 |
| (i) | 1.00 | 0.89 | 0.84 |

## 4. NON-PERIODIC BOUNDARIES

Many applications involve computations in domains with non-periodic boundaries. This section introduces approximations for the first and second derivatives for the near boundary nodes. These approximations are, of necessity, non-central or one-sided. In developing these boundary formulations emphasis has been placed on maintaining a discrete form of global conservation. Analysis of the local error inherent in the approximation is included. This is supplemented by numerical estimates of the global error. Typically the global error is dominated by the boundary error.

### 4.1. *Boundary Formulation for the First Derivative*

The first derivative at the boundary $i = 1$ may be obtained from a relation of the form[3]

$$f'_1 + \alpha f'_2 = \frac{1}{h}(af_1 + bf_2 + cf_3 + df_4), \qquad (4.1.1)$$

coupled to the relations (2.1) written for the interior nodes. Requiring (4.1.1) to be at least second-order accurate constrains the coefficients to

$$a = -\frac{3+\alpha+2d}{2}, \quad b = 2+3d, \quad c = -\frac{1-\alpha+6d}{2}. \qquad (4.1.2)$$

If higher order formal accuracy is desired schemes of third and fourth order may be derived. These are given by

$$a = -\frac{11+2\alpha}{6}, \quad b = \frac{6-\alpha}{2}, \quad c = \frac{2\alpha-3}{2}, \quad d = \frac{2-\alpha}{6}, \quad \text{(third order)} \qquad (4.1.3)$$

$$\alpha = 3, \quad a = -\tfrac{17}{6}, \quad b = \tfrac{3}{2}, \quad c = \tfrac{3}{2}, \quad d = -\tfrac{1}{6}. \quad \text{(fourth order)} \qquad (4.1.4)$$

The leading order truncation error (on the r.h.s. of (4.1.1)) for these boundary approximations are given by $((2-\alpha-6d)/3!)\,h^2 f_1^{(3)}$ for second-order schemes, by $(2(\alpha-3)/4!)\,h^3 f_1^{(4)}$ for third-order schemes and by $(6/5!)\,h^4 f_1^{(5)}$ for the fourth-order scheme. It may be noted that for the even order schemes the leading order truncation error is of dispersive type, while for the third-order schemes it is dissipative.

As discussed in Section 3.1 the Fourier analysis of the differencing schemes reveals the resolution characteristics of the schemes. In the following the same approach is used to compare different boundary approximations. It should be noted that this Fourier analysis of the boundary approximations can be justified only at a heuristic level [29], while the application of the analysis to the interior differencing is rigorous (for problems with periodic boundary conditions).

The modified wavenumber $w'$ (introduced in (3.1)) corresponding to (4.1.1) is in general complex. The real part of $w'$, indicated by $w'_r$, is associated with the dispersive error (when different from $w$) and the imaginary part, $w'_i$, is associated with the dissipative error. In Figs. 6–7 the real and imaginary parts of $w'$ are plotted for various boundary approximations. It may be noted that increasing the formal accuracy of the explicit approximations (shown by curves a, b, and c) reduces the dissipative error in the low-intermediate wavenumber range, but at the same time degrades the dispersive error for the intermediate wavenumbers. The second-order compact scheme (4.1.2) (shown by curve d) with $\alpha = 1$ and $d = 0$ discussed by Adam [17] is purely non-dissipative (i.e., $w'_i = 0$), its formal truncation error (in evaluating $f'_1$) is $\frac{1}{4}$ the truncation error of the explicit second-order form, but has a singular $w'_r$ at $w = \pi$. The third-order compact scheme (4.1.3) with $\alpha = 2$ (giving $d = 0$)

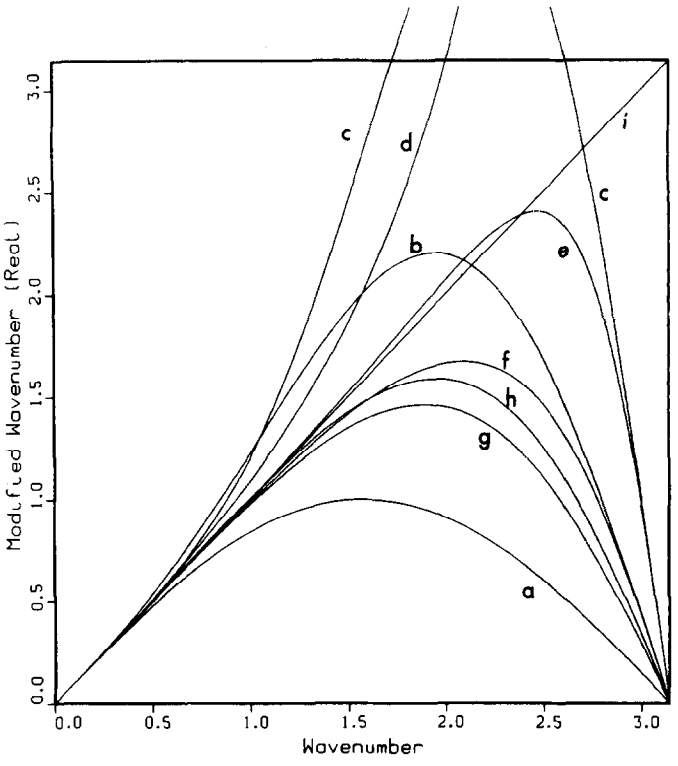


**FIG. 6.** Real part of modified wavenumber for first derivative boundary schemes: (a) first-order explicit scheme; (b) second-order explicit scheme; (c) third-order explicit scheme; (d) second-order compact scheme ($\alpha = 1$, $d = 0$); (e) third-order compact scheme ($\alpha = 2$, $d = 0$); (f) fourth-order compact scheme ($\alpha = 3$); (g) third-order compact scheme ($\alpha = 5$); (h) second-order compact scheme ($\alpha = 4$, $d = -\frac{1}{2}$); (i) exact differentiation.

[3] With this choice the boundary schemes can be used with a tridiagonal interior scheme without increasing the bandwidth.

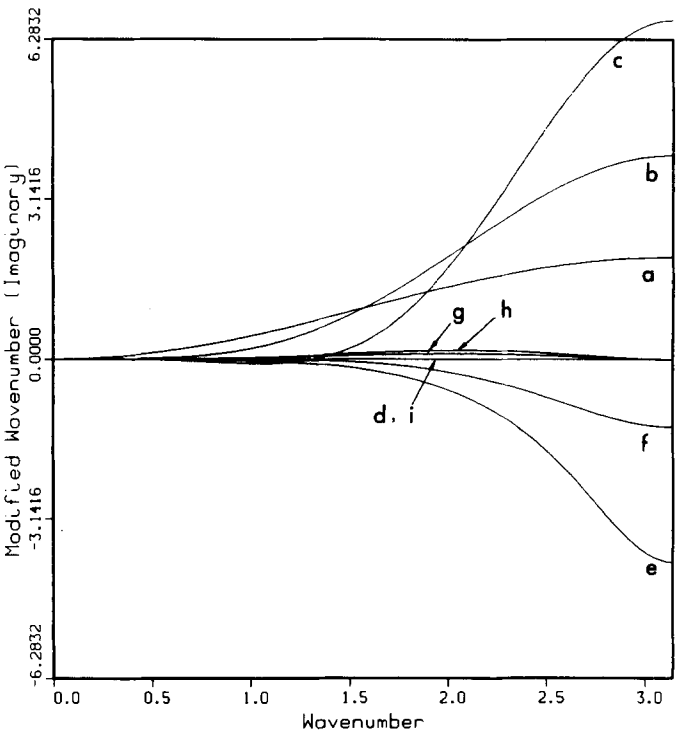


**FIG. 7.** Imaginary part of modified wavenumber for first derivative boundary schemes: (a) first-order explicit scheme; (b) second-order explicit scheme; (c) third-order explicit scheme; (d) second-order compact scheme ($\alpha = 1$, $d = 0$); (e) third-order compact scheme ($\alpha = 2$, $d = 0$); (f) fourth-order compact scheme ($\alpha = 3$); (g) third-order compact scheme ($\alpha = 5$); (h) second-order compact scheme ($\alpha = 4$, $d = -\frac{1}{2}$); (i) exact differentiation.

shown by curve e (also discussed in [17]), has quite small dispersive errors. Its formal truncation error (in evaluating $f'_1$) is $\frac{1}{9}$ of the third-order explicit form. Its dissipative error is also confined to high wavenumbers. Other members of (4.1.3) shown by curves f and g reduce the dissipative error but also degrade the dispersive error. Also shown on the plots is a second-order compact scheme with virtually no dissipative error. This scheme is obtained by adding the constraint $w'_i(w = \pi) = 0$ to (4.1.2). This constraint requires that $d = -\frac{1}{2}$, thus reducing (4.1.2) to a one-parameter ($\alpha$) family. The member with $\alpha = 4$ is plotted (shown as curve h). It may be noted again that the desired characteristics of a finite difference scheme are better achieved by directly optimizing the scheme (in Fourier space) rather than by seeking the scheme with the lowest possible truncation error.

It may also be noted that for many of the compact boundary schemes described here $w'_i$ has a sign opposite to that obtained with the explicit one-sided boundary formulas. For the second-order schemes defined by (4.1.2) it may be shown that $w'_i(w = \pi) = 4(2d + 1)/(1 - \alpha)$. Thus for the (first), second- and third-order explicit ($\alpha = 0$) boundary schemes the values of $w'_i(\pi)$ are (2), 4, and $\frac{20}{3}$, while for the compact schemes labelled (e) and (f) the values are $-4$ and $-\frac{4}{3}$, respectively.

It is erroneous to conclude from this behavior of $w'_i(\pi)$ that the compact boundary schemes (described above) may lead to unstable numerical schemes. The stability of the complete numerical scheme (including the boundary approximation) must be determined by the appropriate eigenvalue analysis. In the present instance this eigenvalue analysis requires a numerical solution. It, however, establishes that the boundary schemes presented in this section do in fact lead to stable numerical schemes. Details of such an eigenvalue analysis are deferred to Section 4.5. Numerical tests of the global performance of the complete difference scheme is presented in Section 4.4.

### 4.2. *Conservative Formulation for the First Derivative*

In this section an approach is presented for constructing difference approximations (for the near boundary nodes) which satisfy a discrete form of global conservation constraint. Such a treatment is particularly useful in the discretization of conservation laws. To motivate the discussion consider a conservation law of the form

$$\frac{\partial f}{\partial t} + \frac{\partial F}{\partial x} = 0, \tag{4.2.1}$$

over the domain $[a, b]$ and $F = F(f)$, with some initial and boundary conditions. Integrating (4.2.1) over the domain yields

$$\frac{d}{dt}\int_{x=a}^{x=b} f(x, t)\, dx = F|_{(x=a, t=t)} - F|_{(x=b, t=t)}, \tag{4.2.2}$$

showing that the total $f$ in the domain (i.e., integral of $f$) changes (in time) only due to the flux of $f$ at the boundary. This is a global conservation statement. We seek a formulation for the near boundary nodes such that the global conservation law (4.2.2) has a discrete analog for the difference approximations. As a result of this constraint the difference approximations imply the appropriate quadrature weights for discretizing the integral on the l.h.s. of (4.2.2).

We consider the system of linear equations (2.1), one equation for each interior node but not necessarily with the same coefficients, along with the boundary equation (4.1.1) (and the analogous equation at the other boundary). Formally this system may be written as

$$\mathbf{A}\hat{\mathbf{f}}' = \frac{1}{h}\mathbf{B}\hat{\mathbf{f}}, \tag{4.2.3}$$

where $\mathbf{A}$, $\mathbf{B}$ are $N \times N$ sparse matrices and $\hat{\mathbf{f}}$, $\hat{\mathbf{f}}'$ are $N$ vectors

representing the values of the function and its derivative at the nodes, respectively. In order to satisfy the global conservation constraint it is sufficient to require that the columns 2 through $N-1$ of the matrix **B** sum exactly to zero. This ensures that only the boundary nodes contribute to the boundary fluxes. Given a particular interior scheme (say with a particular value of $\alpha$ along with any of (2.1.6)–(2.1.13)) we show that near boundary approximations of the form (2.1.6)–(2.1.13) and (4.1.1) may be chosen to satisfy global conservation. This derivation is simpler when the interior scheme has $c=0$ and is presented first. The procedure is easily adapted to interior schemes with $c\neq 0$. This is summarized subsequently.

For the interior schemes with $c=0$ it is sufficient to consider the entries in the top left corner of **B**, viz.,

$$\mathbf{B}=\begin{pmatrix} -w_1 p & w_1 q & w_1 r & w_1 s & 0 & 0 & 0 & 0\,0\,0 & \cdots & 0 \\ -w_2 q' & 0 & w_2 q' & 0 & 0 & 0 & 0 & 0\,0\,0 & \cdots & 0 \\ -w_3 r'' & -w_3 q'' & 0 & w_3 q'' & w_3 r'' & 0 & 0 & 0\,0\,0 & \cdots & 0 \\ 0 & -\hat{r} & -\hat{q} & 0 & \hat{q} & \hat{r} & 0 & 0\,0\,0 & \cdots & 0 \\ 0 & 0 & -\hat{r} & -\hat{q} & 0 & \hat{q} & \hat{r} & 0\,0\,0 & \cdots & 0 \\ 0 & 0 & 0 & -\hat{r} & -\hat{q} & 0 & \hat{q} & \hat{r}\,0\,0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & -\hat{r} & -\hat{q} & 0 & \hat{q}\,\hat{r}\,0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & -\hat{r} & -\hat{q} & 0\,\hat{q}\,\hat{r} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\,\vdots\,\vdots & \ddots & \vdots \end{pmatrix} \tag{4.2.4}$$

In writing the entries of **B** we have explicitly allowed for the weights $w_1$, $w_2$, and $w_3$. The choice of the schemes at the near boundary nodes has been restricted. In particular at node 2, neighboring the boundary node, the standard Padé form is used, thus $q'=\frac{3}{4}$. The coefficients $p, q, r, s$ are given by (4.1.1) and other coefficients satisfy the constraints defined by (2.1.6)–(2.1.10) as appropriate. It may be seen that at least 10 nodes (four interior and three near boundary nodes from each boundary) are needed for the fluxes to telescope to the boundary fluxes. By imposing the specific global conservation restriction it follows that

$$\frac{q''(\alpha'')}{r''(\alpha'')}=\frac{q}{q+s}\frac{\hat{q}(\hat{\alpha})}{\hat{r}(\hat{\alpha})}+\frac{q-s}{q+s}, \tag{4.2.5}$$

where $\hat{\alpha}$ and $\alpha''$ are the specific values of the coefficient $\alpha$ of the schemes used in the interior and at node 3, respectively. Once specific choices about the family of schemes to be used in the interior and at node 3 are made, all of the coefficients appearing in (4.2.4) are determined. In the following we assume that the schemes used at node 3 and the interior belong to the same family.

The global conservation requirements lead to

$$\alpha''=\frac{(40\hat{\alpha}-1)q+7(4\hat{\alpha}-1)s}{16(\hat{\alpha}+2)q+8(1-4\hat{\alpha})s}, \tag{4.2.6}$$

$$w_1=\frac{2\hat{\alpha}+1}{2(q+s)}, \tag{4.2.7}$$

$$w_2=\frac{(8\hat{\alpha}+7)q-6(2\hat{\alpha}+1)r+(8\hat{\alpha}+7)s}{9(q+s)}, \tag{4.2.8}$$

$$w_3=\frac{4(\hat{\alpha}+2)q+2(1-4\hat{\alpha})s}{9(q+s)}, \tag{4.2.9}$$

if the scheme (2.1.6) is used in the interior.

The interior scheme (2.1.11) yields

$$\alpha''=\frac{(4209\hat{\alpha}-289)q+(2679\hat{\alpha}-799)s}{(3249\hat{\alpha}+2271)q+(561-1881\hat{\alpha})s}, \tag{4.2.10}$$

$$w_1=\frac{5(3\hat{\alpha}+1)}{12(q+s)}, \tag{4.2.11}$$

$$w_2=\frac{(33\hat{\alpha}+47)q-30(3\hat{\alpha}+1)r+(33\hat{\alpha}+47)s}{54(q+s)}, \tag{4.2.12}$$

$$w_3=\frac{(1083\hat{\alpha}-757)q+(187-627\hat{\alpha})s}{1080(q+s)}. \tag{4.2.13}$$

Once the weights $w_1$, $w_2$, $w_3$, and $\alpha''$ are determined the (implied) nodal weights in computing the integral on the l.h.s. of (4.2.2) become fixed. If a non-uniform mesh defined by a mathematical mapping $x=x(s)$ (not changing with time) between a uniformly spaced mesh $s$ and the physical coordinate $x$ is used, rewriting (4.2.2) as

$$\frac{d}{dt}\int_{x=a}^{x=b}\frac{dx(s)}{ds}f(x(s),t)\,ds=F|_{(x=a,t=t)}-F|_{(x=b,t=t)}, \tag{4.2.14}$$

shows that the near boundary formulation discussed above preserves global conservation (with the factor $(dx/ds)|_{x=x_i}$ included with the nodal weights for calculating the l.h.s of (4.2.14)). Such a definition for discretizing the integral naturally assigns more weight to larger computational cells.

Now we summarize the near boundary formulation for the interior schemes with $c\neq 0$. It becomes necessary to consider a larger number of entries of **B**, viz.

$$\begin{bmatrix} w_1 p & w_1 q & w_1 r & w_1 s & 0 \\ -w_2 q' & 0 & w_2 q' & 0 & 0 \\ -w_3 r'' & -w_3 q'' & 0 & w_3 q'' & w_3 r'' \\ -w_4 s''' & -w_4 r''' & -w_4 q''' & 0 & w_4 q''' \\ 0 & -\hat{s} & -\hat{r} & -\hat{q} & 0 \\ 0 & 0 & -\hat{s} & -\hat{r} & -\hat{q} \\ 0 & 0 & 0 & -\hat{s} & -\hat{r} \\ 0 & 0 & 0 & 0 & -\hat{s} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ w_4 r''' & w_4 s''' & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hat{q} & \hat{r} & \hat{s} & 0 & 0 & 0 & 0 & \cdots \\ 0 & \hat{q} & \hat{r} & \hat{s} & 0 & 0 & 0 & \cdots \\ -\hat{q} & 0 & \hat{q} & \hat{r} & \hat{s} & 0 & 0 & \cdots \\ -\hat{r} & -\hat{q} & 0 & \hat{q} & \hat{r} & \hat{s} & 0 & \cdots \\ -\hat{s} & -\hat{r} & -\hat{q} & 0 & \hat{q} & \hat{r} & \hat{s} & \cdots \\ 0 & -\hat{s} & -\hat{r} & -\hat{q} & 0 & \hat{q} & \hat{r} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{4.2.15}$$

A minimum of 14 nodes (six interior and four at each boundary) are needed to allow for the telescoping of the fluxes. The global conservation constraint requires that

$$\frac{r'''(\alpha''')}{s'''(\alpha''')} = \frac{\hat{r}(\hat{\alpha})}{\hat{s}(\hat{\alpha})} + 1, \tag{4.2.16}$$

where $\hat{\alpha}$ and $\alpha'''$ are the specific values of the coefficient $\alpha$ of the schemes used in the interior and at node 4, respectively. The weights may be obtained from

$$w_1 = \frac{\hat{q} + 2\hat{r} + 3\hat{s}}{q+s}, \quad w_4 = \frac{\hat{s}}{s'''},$$
$$w_2 = \frac{4}{3}\hat{r} + \hat{s} - (\hat{q} + 2\hat{r} + 3\hat{s})\frac{r}{q+s} + \frac{\hat{s}q'''}{s'''}, \tag{4.2.17}$$

and

$$w_3 r'' = \hat{q} + \hat{r} + \hat{s}\left(1 - \frac{q'''}{s'''}\right),$$
$$w_3 q'' = \frac{(q-s)(\hat{q}+\hat{s}) + (q-2s)\hat{s}}{q+s}. \tag{4.2.18}$$

The weights $w_1$, $w_2$, $w_4$ depend only on the interior scheme and the scheme at node 1, while the weight $w_3$ depends on the choice made for the scheme at node 3.

### 4.3. *Boundary Formulation for the Second Derivative*

The relations appropriate for near boundary nodes between the nodal values of a function and its second derivative may be derived by Taylor series expansions. The compact scheme analogous to (4.1.1) is given by

$$f_1'' + 11 f_2'' = \frac{1}{h^2}(13f_1 - 27f_2 + 15f_3 - f_4). \tag{4.3.1}$$

This relation is formally third-order accurate (truncation error on the r.h.s. is $h^3 \frac{1}{12} f^{(5)}$). The explicit expressions (at the boundary node) with second- and third-order formal accuracy are given by

$$f_1'' = \frac{1}{h^2}(2f_1 - 5f_2 + 4f_3 - f_4),$$
(second order) (4.3.2)

$$f_1'' = \frac{1}{h^2}\left(\frac{35}{12}f_1 - \frac{26}{3}f_2 + \frac{19}{2}f_3 - \frac{14}{3}f_4 + \frac{11}{12}f_5\right)$$
(third order). (4.3.3)

Their truncation errors are $\frac{11}{12}h^2 f^{(4)}$ and $\frac{5}{6}h^3 f^{(5)}$. The truncation error of the explicit third-order form is 10 times larger than that of the third-order compact form.

Boundary schemes for other neighboring nodes may be chosen from (2.2). It is possible to extend the global conservation considerations (Section 4.2) to the evaluation of the second derivatives. For these purposes it becomes necessary to introduce more general compact boundary schemes

$$f_1'' + \alpha f_2'' = \frac{1}{h^2}(af_1 + bf_2 + cf_3 + df_4 + ef_5). \tag{4.3.4}$$

Requiring second-order accuracy restricts the coefficients to

$$a = \alpha + 2 + e, \quad b = -(2\alpha + 5 + 4e),$$
$$c = \alpha + 4 + 6e, \quad d = -(1 + 4e). \tag{4.3.5}$$

The formal truncation error is $\frac{1}{12}(\alpha + 12e - 11)h^2 f^{(4)}$. By requiring third-order formal accuracy the coefficients are reduced to

$$a = \frac{11\alpha + 35}{12}, \quad b = -\frac{5\alpha + 26}{3}, \quad c = \frac{\alpha + 19}{2},$$
$$d = \frac{\alpha - 14}{3}, \quad e = \frac{11 - \alpha}{12}. \tag{4.3.6}$$

The formal truncation error is reduced to $((\alpha - 10)/12)h^3 f^{(5)}$.

### 4.4. *Numerical Tests (Including Boundary Scheme)*

For problems with periodic boundary conditions the Fourier analysis presented in Section 3.1 may be used to assess the global errors. This, however, is not possible with other forms of boundary conditions. The global errors may still be assessed by direct numerical tests. Such tests can be performed at various levels. The simplest of these is to compare the numerically calculated derivative with the known derivative of various test functions. In order to test the schemes on functions which contain a range of scales, such functions were numerically synthesized by taking a sum of different Fourier modes (representable on the mesh). The phases of the Fourier modes were chosen randomly (uniformly distributed over $[0, 2\pi]$) and their amplitudes were chosen to synthesize a prescribed energy spectrum. For the examples presented here the interval $[0, 1]$ was descretized into 128 intervals (i.e., 129 points when counting both end points). On this mesh Fourier modes $[0, 63]$ (for the wavenumber $k$ in (3.1.1)) may be represented. The mode $k = 64$ (or the $2\delta$ wave) is not included. In the examples to be discussed Fourier modes with $k$ in the range $[0, k_m]$ were included. The amplitude of the Fourier modes were equal and phases were random (white noise). For the numerical tests the conservative formulation of the first derivative scheme with $\alpha = \frac{1}{3}$, $\beta = 0 = c$ (2.1.7) were used. At the end points the third-order compact boundary scheme with $d = 0$ was used. This overall scheme was chosen as it has been used in several practical applications [36–37]. In Figs. 8a–b numerical tests of this overall scheme are shown for four classes of test functions (with $k_m$ varying from 9, 21, 31, and 63). One randomly chosen realization of the test functions is displayed for each class in Fig. 8a. On Fig. 8b the numerically computed first derivatives (shown with a dashed line) and the exact derivative of the test functions (shown with a solid line) are plotted. It may be noted that dominant error in the derivatives always occurs at the boundary. Only for cases with $k_m$ of 31 and 63 this localized boundary error is visible on the plots, and even in these cases the interior is virtually error-free.

### 4.5. *Eigenvalue Analysis of the Complete Differencing Scheme*

This section presents an eigenvalue analysis of the complete first derivative scheme. Such an analysis is necessary to establish that the overall differencing scheme generates a numerically stable algorithm. We begin by considering

$$\frac{\partial f}{\partial t} + c\frac{\partial f}{\partial x} = 0, \tag{4.5.1}$$

over a domain $[0, 1]$ with a prescribed boundary condition $f(x = 0, t) = 0$. For the present purposes there is no loss of

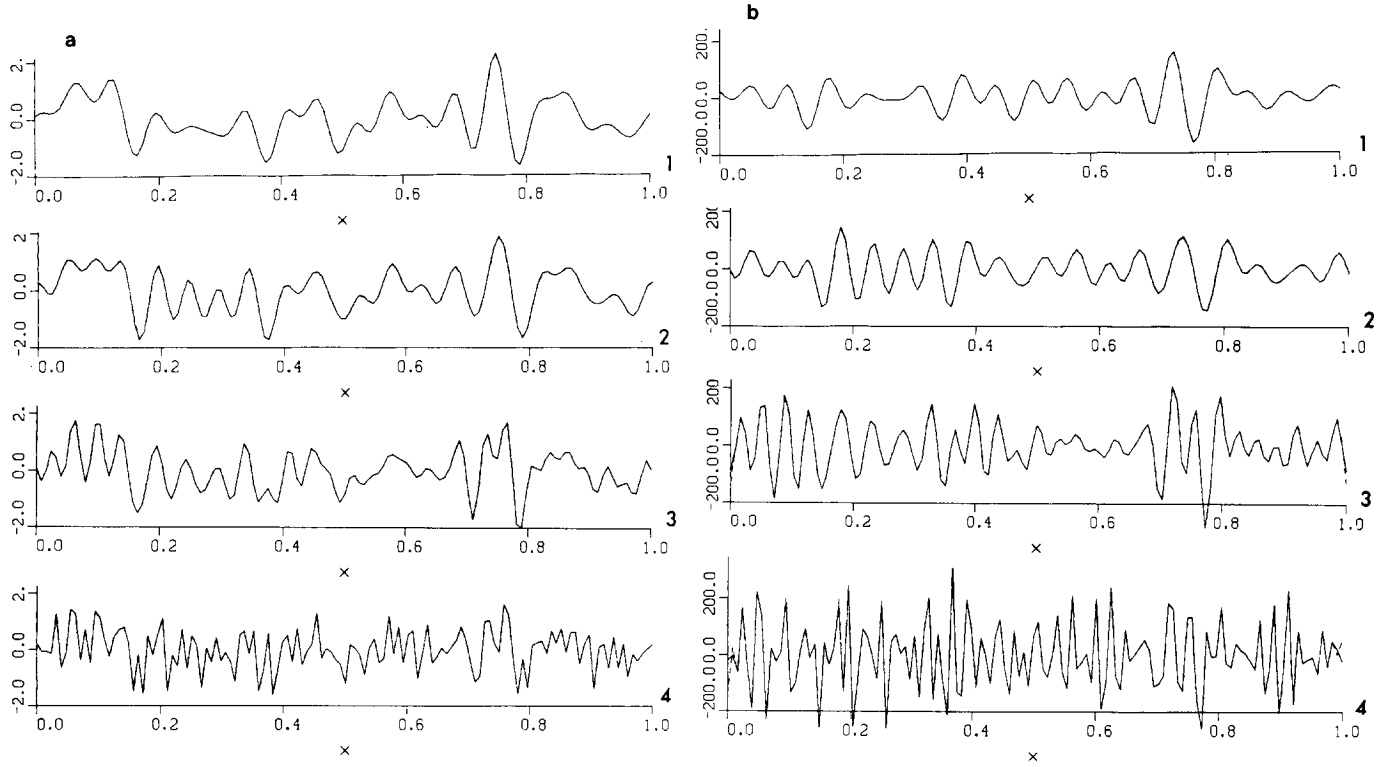


**FIG. 8.** Numerical tests on the derivative scheme; the domain is divided into 128 intervals. The test functions contain Fourier modes up to wavenumber $k_m$ with equal amplitude and random phases. One realization is plotted for each case with (1) $k_m = 9$; (2) $k_m = 21$; (3) $k_m = 31$; (4) $k_m = 63$; on (a). The corresponding exact derivative (solid line) and the calculated derivative (dashed line) are shown on (b).

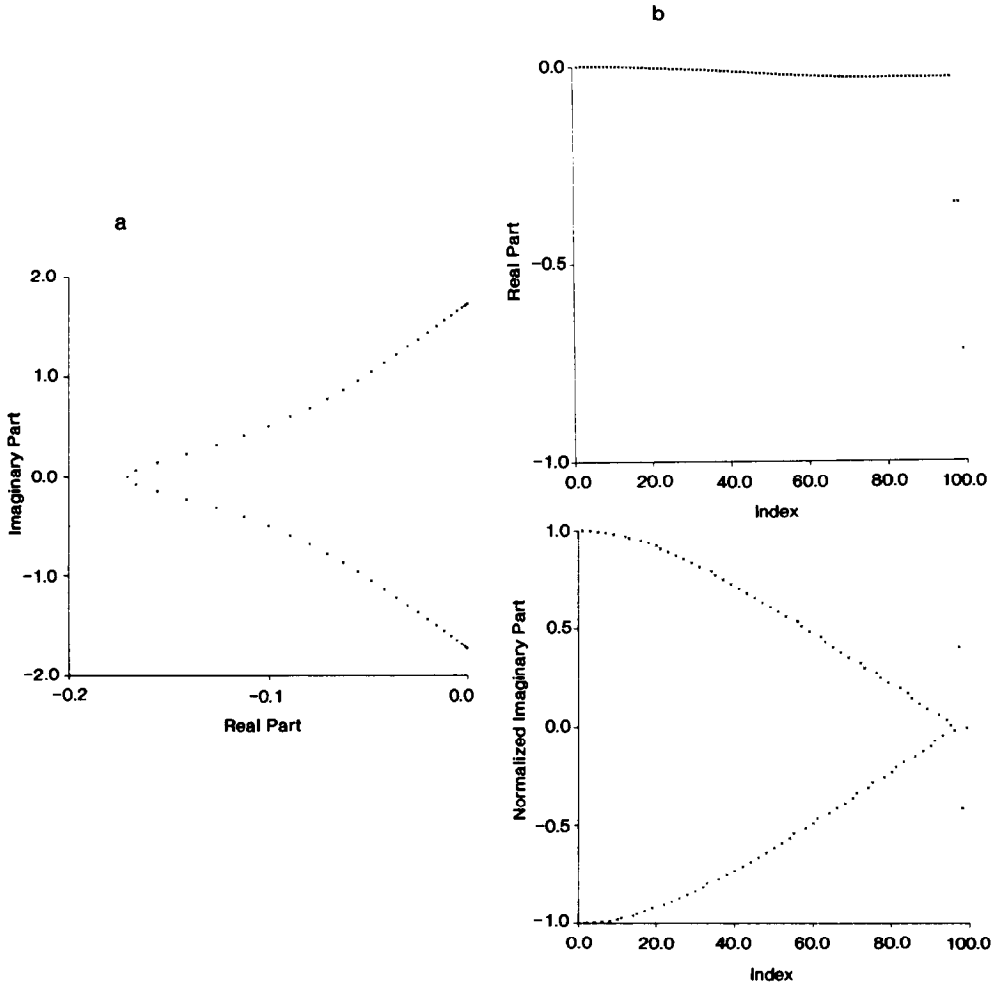


**FIG. 9.** Eigenvalue spectrum $\omega'$ for the complete derivative scheme; boundary formulation is non-conservative and (4.1.3) is used with $\alpha = 2$ ($d = 0$). $N = 49$. (a) Interior scheme (2.1.6) with $\alpha = \frac{1}{4}$, $\beta = b = c = 0$ (standard Padé); (b) interior scheme (2.1.6) with $\alpha = \frac{1}{3}$, $\beta = c = 0$.

generality in assuming this simple form for the boundary condition. The domain is descretized into $N$ equal intervals (i.e., the step size in $x$, $\Delta x = 1/N$). The derivative $\partial f/\partial x$ in (4.5.1) is evaluated by the differencing schemes described earlier in the paper. This linear operation may be formally written as

$$\mathbf{A}\hat{\mathbf{f}}' = \frac{1}{\Delta x}\mathbf{B}\hat{\mathbf{f}}, \tag{4.5.2}$$

where $\mathbf{A}$, $\mathbf{B}$ are $N \times N$ (sparse) matrices and $\hat{\mathbf{f}}$, $\hat{\mathbf{f}}'$ are $N$ vectors representing the values of the function and its derivative at the nodes $x_i = i/N$, respectively. Pre-multiplying the spatially descritized form of (4.5.1) by $\mathbf{A}$ and substituting (4.5.2) yields

$$\mathbf{A}\frac{\hat{\mathbf{df}}}{\mathbf{dt}} = -\frac{c}{\Delta x}\mathbf{B}\hat{\mathbf{f}}, \tag{4.5.3}$$

where $\hat{\mathbf{df}}/\mathbf{dt}$ is a $N$ vector of the time derivative of the nodal values.[4] Since (4.5.3) is a system of ODE's in time with constant coefficients we look for normal modes $\hat{\mathbf{f}} = \exp(\omega t)\tilde{\mathbf{f}}$ reducing (4.5.3) to an eigenvalue problem

$$-\frac{\omega\,\Delta x}{c}\mathbf{A}\tilde{\mathbf{f}} = \mathbf{B}\tilde{\mathbf{f}}, \tag{4.5.4}$$

involving the eigenvalue $\omega' = \omega\,\Delta x/c$. For the interior and boundary schemes of interest this matrix eigenvalue problem requires a numerical solution (performed with the IMSL routine EIGZF in double precision arithmetic). The eigenvalue $\omega'$ or $\omega$ is, in general, complex valued and depends on the size $N$ of the matrices $\mathbf{A}$, $\mathbf{B}$, the interior differencing scheme, and the boundary schemes. For the numerical stability of the overall differencing scheme it is required that all the eigenvalues of (4.5.4) lie in the left half

[4] In the numerical implementation the first derivative scheme also generates an approximation to $(\partial f/\partial x)|_{x=0}$ which is discarded in writing (4.5.3) since $f_{x=0}$ is prescribed by the boundary condition.

of the complex plane. A technical discussion of stability may be found in Strikwerda [38].

It should be noted that the present analysis (i.e., method of lines) is based on the semi-discrete form (no time differentiation errors) of (4.5.1). The effect of particular time descretization may be analyzed in an analogous manner.

Eigenvalues are calculated for the tridiagonal differencing scheme family (2.1.6) when combined with either globally conservative or non-conservative boundary forms (in the sense of (4.2)). Boundary forms given by (4.1.1) as well as explicit boundary forms are considered. In all cases considered the eigenvalues of (4.5.4) satisfy the stability constraint. In Fig. 9 two examples of the eigenvalue spectrum are shown. In all cases (with a single exception to be noted below) the real part of the eigenvalue $\omega'$ with the least negative real part approaches the imaginary axis like $N^{-3}$ as $N$ is increased. The asymptotic behaviour was inferred from the numerically obtained eigenvalues for $N$ up to 200. Note that the analysis of Warming and Beam [39] which uses a second-order interior scheme also predicts the same asymptotic behaviour. The only exception to the $N^{-3}$ behaviour is found for standard Padé scheme ($\alpha = \frac{1}{4}$ in (2.1.6)) when combined with the second-order compact boundary scheme $\alpha = 4$, $d = -\frac{1}{2}$ (giving $w_i'(\pi) = 0$) which has a zero eigenvalue (generalized GKS eigenvalue [39]). The same boundary scheme when combined with other members of the family (2.1.6) does not possess this zero eigenvalue.

The behaviour of the eigenvalue $\omega'$ with most negative real part depends on the interior scheme. For the standard Padé scheme this eigenvalue migrates towards the origin but at a rate slower than $N^{-1}$ (with any boundary scheme). The rate of approach is consistent with the $\log N/N$ approach predicted by [39] for second-order interior schemes. For other members of the family (2.1.6) this eigenvalue ($\omega'$) remains fixed as $N$ is increased (with any boundary scheme). In this case there are additional complex eigenvalues (adjacent to the one with the most negative real part) which also remain fixed. In all cases the eigenvalue spectrum falls into the three classes of eigenvalues defined by Warming and Beam [39] with no unstable eigenvalues.

In Fig. 10 the influence of changing the boundary scheme

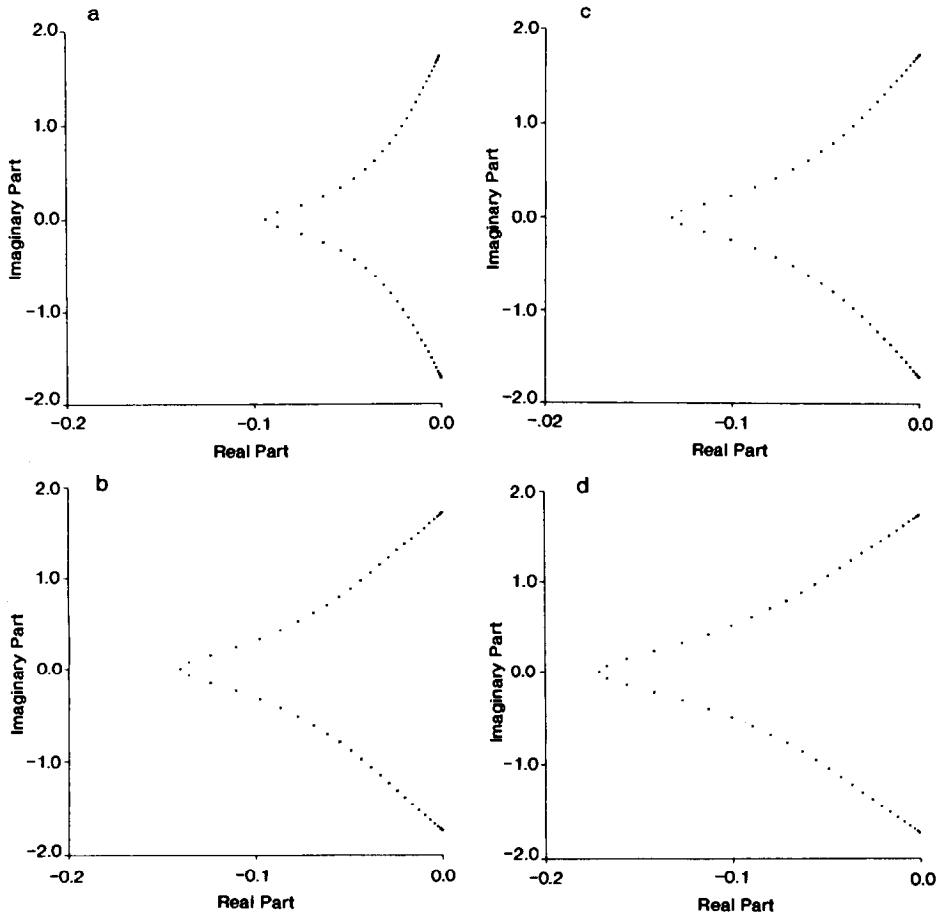


**FIG. 10.** Effect of boundary formulation on the eigenvalue spectrum $\omega'$ for the complete derivative scheme; boundary formulation is non-conservative and interior scheme (2.1.6) with $\alpha = \frac{1}{4}$, $\beta = b = c = 0$ is used. $N = 49$. (a) First-order explicit scheme at the boundary; (b) second-order explicit scheme at the boundary; (c) second-order compact scheme at the boundary ($\alpha = 1$, $d = 0$); (d) third-order compact scheme at the boundary ($\alpha = 2$, $d = 0$).

with a fixed interior scheme (standard Padé) on the eigenvalue spectrum is shown. It may be noted that, with a fixed number of nodes $N$, as the local accuracy of the boundary scheme is increased the eigenvalue spectrum shifts towards more negative eigenvalues. This unexpected trend serves as a caution against associating the eigenvalue spectrum with the overall performance (accuracy) of the differencing scheme.

### 4.6. *Stability Limits for Explicit Schemes*

This section summarizes the time step limits which need to be maintained for a numerically stable time advancement. One may anticipate (by analogy with the spectral methods) that the differentiation schemes presented here would require smaller time steps than are possible with the standard second-order central difference schemes. In order to keep the analysis simple the stability bounds are presented only for model problems with periodic boundary conditions. Furthermore, the analysis is restricted to the cases of pure advection and pure diffusion operators. For the more general non-periodic case the calculation of precise stability bounds requires a numerical calculation of the spectral radius of the associated matrices (as presented in Section 4.5).

First we consider the pure advection case (on a periodic domain):

$$\frac{\partial f}{\partial t} + c\frac{\partial f}{\partial x} = 0. \tag{4.6.1}$$

The stability limit in this case is given by (e.g., [30])

$$\frac{c\,\Delta t}{\Delta x} \leqslant \frac{\sigma_i}{w'_m}, \tag{4.6.2}$$

where $\Delta x$ and $\Delta t$ are the step sizes in $x$ and $t$, respectively, $[-i\sigma_i, i\sigma_i]$ is the segment of the imaginary axis in the stable region for the time advancement scheme, and $w'_m$ is the maximum value of the modified wavenumber for the first derivative operator (i.e., maximum ordinate achieved on Fig. 1 for the first derivative scheme). For the schemes displayed on Fig. 1 (labelled (a)–(j)) the values of $w'_m$ are 1.0, 1.372, 1.586, 1.732 (or $\sqrt{3}$), 1.989, 2.133, 2.205, 2.324, 3.142 (or $\pi$), respectively. We remind the reader that the values of $\sigma_i$ for the second-, third-, and fourth-order Runge–Kutta schemes are 0, $\sqrt{3}$, and 2.85, respectively.

Next we consider the pure diffusion case (on a periodic domain):

$$\frac{\partial f}{\partial t} = \nu\frac{\partial^2 f}{\partial x^2}. \tag{4.6.3}$$

The stability limit in this case is given by (e.g., [30])

$$\frac{\nu\,\Delta t}{(\Delta x)^2} \leqslant \frac{\sigma_r}{w''_m}, \tag{4.6.4}$$

where $\Delta x$ and $\Delta t$ are the step sizes in $x$ and $t$, respectively, $[-\sigma_r, 0]$ is the segment of the real axis in the stable region for the time advancement scheme, and $w''_m$ is the maximum value of $w''$ for the second derivative operator (i.e., maximum ordinate achieved on Fig. 5 for the second derivative scheme). For the schemes displayed on Fig. 5 (labelled (a)–(j)) the values of $w''_m$ are 4.0, 5.333, 6.044, 6.0, 6.857, 7.324, 7.471, 7.838, 9.108, 9.860 (or $\pi^2$), respectively. Typically $w''$ achieves its maximum value at $w = \pi$. In such a case

$$w''_m = \frac{4(a + c/9)}{1 - 2\alpha + 2\beta}. \tag{4.6.5}$$

We remind the reader that the values of $\sigma_r$ for the second-, third-, and fourth-order Runge–Kutta schemes are 2, 2.5, and 2.9, respectively. It may be noted that, as expected, the spectral schemes require the most severe time stepping restriction when explicit time advancement is used.

## 5. APPLICATIONS

This section illustrates application of the differencing schemes developed in this paper to some problems from fluid mechanics. Some general remarks are made first about these applications.

The differencing schemes described here provide an improved resolution of the short length scales. Furthermore, the schemes have a pure central difference form (except near the boundaries); i.e., they have no built-in artificial dissipation. It is, therefore, necessary that the applications to which they are (directly) applied be such that there is a (physically) well defined cutoff for the shortest scales. In other words the shortest scales should be determined *physically and not numerically*. This rules out (direct) applications to problems with discontinuities (in variables and their derivatives). This is not to say that the present method is inapplicable to inviscid problems. It is, however, restricted to problems with smooth solutions. They may have a wide range of scales but the solution is required to be smooth. In the following an example is presented for resolving the structure of a shock wave and we note that the smallest length scale is determined by viscosity and heat conduction. In treating the viscous or diffusion terms the second derivative approximation (presented in Section 2.1) should be used. This non-conservative form of the viscous terms is chosen over the conservative form with two consecutive applications of the first derivative operator. This

provides a finite damping of the grid to grid oscillations (or $2\delta$ waves) which remain undamped otherwise. The viscous and diffusion terms involving fluid property variations are expanded out using chain rule and then differenced in the manner just indicated. The time step restrictions for stable time advancement with Runge–Kutta methods were summarized in Section 4.6.

The spectral-like nature of the differencing schemes also makes it necessary to use accurate boundary conditions. Boundary conditions which may seem suitable with low order schemes (and with built-in dissipation) may not perform well with the schemes described here. Adaptation of the "non-reflecting" boundary conditions described by Thompson [40] performs well. These were used in [36–37] in applications to compressible mixing layers. A full discussion of the different boundary conditions, comparisons with other methods commonly used with low order schemes, and application to reacting and non-reacting flows are presented in [41]. Applications to 3D incompressible mixing layers and wakes (along with spectral methods in two directions) is described by Buell [42, 43].

An example of compressible mixing layer evolution is presented in Fig. 11. The case displayed is a mixing layer between two supersonic streams of Mach number 2.0 and 1.2. The static pressure and temperature of the two streams are equal, generating a convective Mach number [44] of 0.4. The computation is conducted with 601 by 161 grid points in the streamwise (uniform mesh) and cross-stream direction (with clustering of points in the shearing region). Non-reflecting boundary conditions are imposed on the top, bottom, and outflow boundaries. The inflow is prescribed (hyperbolic tangent velocity profile, Crocco–Buseman relation for temperature) and contains 4% (peak value relative to the velocity difference) fluctuations in the transverse velocity. The fluctuations are confined to the shearing region by means of a Gaussian shape function. The fluctuation has the fundamental and its first two subharmonic components. Additional details may be found in [36]. Figure 11a shows the evolution of the vorticity field. The rollup and pairing of vortices can be clearly seen. In Fig. 11b the evolution of a passive scalar field is shown which illustrates the formation of braid regions in between the vortices. The scalar gradients intensify in the braid region and their peak value is determined by an equilibrium between the straining and diffusion. The spectral-like property of the present schemes allows these sharpened gradients to be resolved with 7–8 grid points across the braid region.

Another example where the viscous (diffusion) process determines the small scale cutoff is presented in Figs. 12–13. The compressible Navier–Stokes equations are solved to resolve the structure of a shock wave. For the example presented the supersonic flow upstream of the shock has a Mach number of 1.5 and $\gamma = \frac{7}{5}$. The calculation uses a uniform mesh with seven to eight grid points across the shock wave and which is stationary in the calculation frame of reference. The profiles of pressure, temperature, density, velocity, and Mach number vary smoothly from the

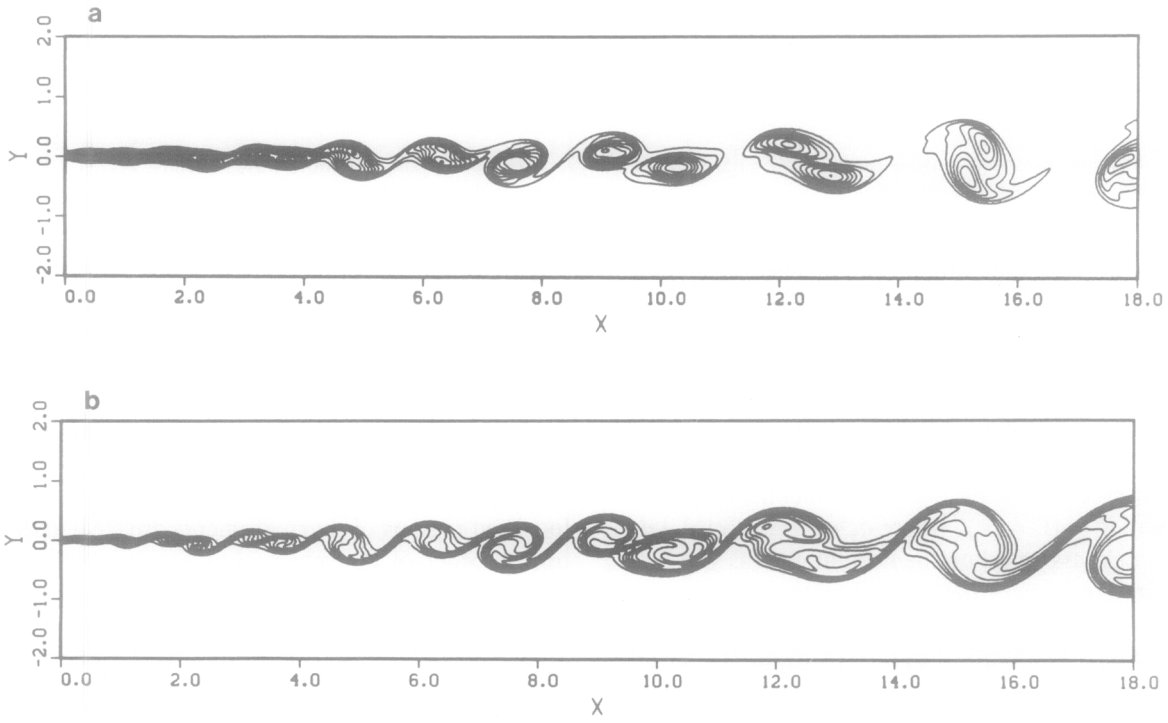


**FIG. 11.** Spatial evolution of a compressible mixing layer: (a) vorticity field; (b) scalar field.

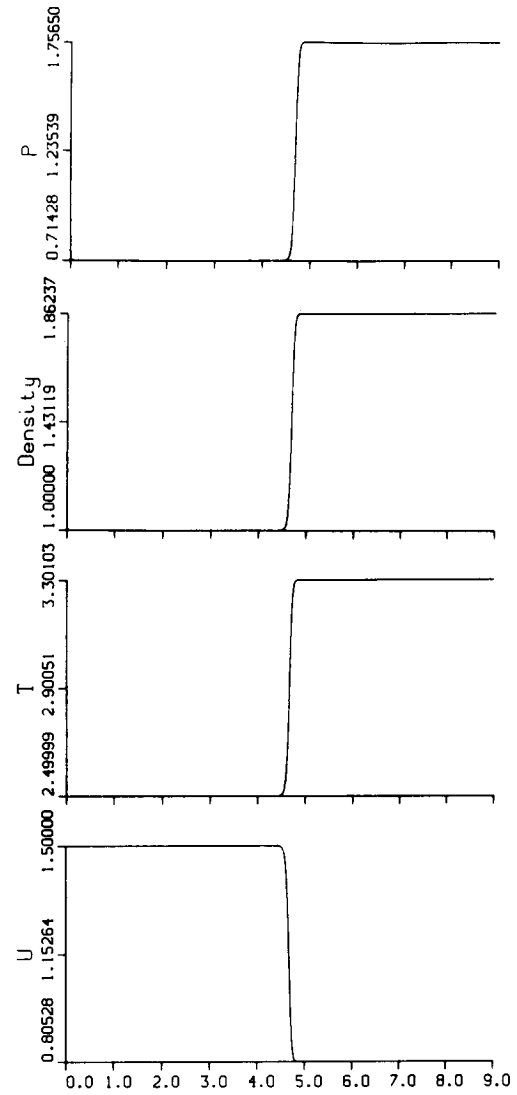


**FIG. 12.** Example of shock wave structure; $M_1 = 1.5$. Profiles of pressure, density, temperature, and velocity.

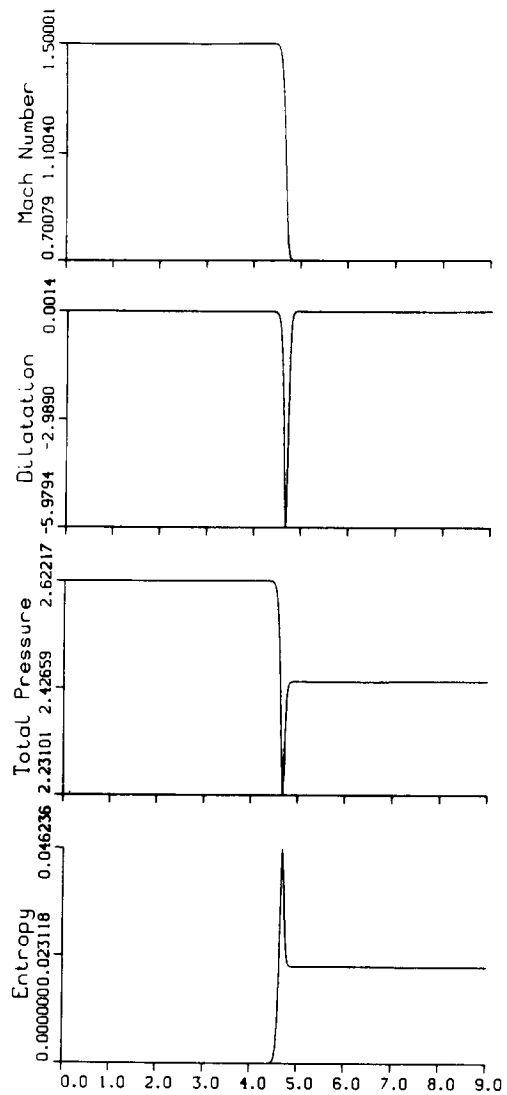


**FIG. 13.** Example of shock wave structure; $M_1 = 1.5$. Profiles of Mach number, dilatation, total pressure, and entropy.

upstream to downstream values. The profiles of dilatation, entropy, and total pressure through the shock wave are a more sensitive test of the resolution. The non-monotone behaviour of these is expected on physical grounds. For weak shocks (small pressure jump across the shock) the structure of the shock wave is analytically known and agrees perfectly with the computations. It may be noted that the Reynolds number based on the shock thickness $R$ is about six to seven ($R = \Delta U\, \delta_\tau / \nu_*$, where $\delta_\tau = \int_{-\infty}^{+\infty} (\tau(x)/\tau_m)\, dx$ is the shock thickness based on the normal viscous stress $\tau$, $\tau_m$ is the maximum value of the normal viscous stress, and $\nu_*$ is the kinematic viscosity evaluated at the sonic reference temperature $T_*$). For weak shocks asymptotic theory [45] gives a falue $32\gamma/3(\gamma+1)$ for this Reynolds number which for a diatomic gas is 6.222. Other definitions of shock thickness (e.g., based on the maximum dilatation) also give a Reynolds number in the range 5.8–6.4 over a Mach number range 1.05–2.0. As may be expected, the thickness measures based on maximum gradients in velocity, density, temperature, pressure or the measures based on viscous stresses or heat flux underestimate (by a factor of 3 or so) the "visual" extent of the shock wave structure due to the curvature of the profiles.

It should be stressed that the resolution with which localized regions of sharp gradients are treated depends not on the order of the scheme but on the resolution characteristics ($w'$ vs $w$) for the intermediate wavenumbers. We further note that the schemes described here are not monotone (or TVD). If they are used to differentiate a step function the computed derivative has grid to grid ($2\delta$) oscillation (even with an arbitrarily fine mesh); however, these oscillations are exponentially damped away from the

location of the step change. This may be contrasted with spectral differentiation which has the well-known Gibbs phenomena and the associated oscillations (in derivative) are damped only algebraically like $x^{-1}$.

Physical discussion of the results obtained by using the schemes discussed here for non-reacting and reacting compressible turbulent flows will be described in future publications.

## 6. SUMMARY

Compact finite difference schemes for the evaluation of derivatives, interpolation, and filtering have been presented and analyzed. Comparisons were made throughout with other well-known schemes. The emphasis has been on improving the representation of a range of wavenumbers rather than accurate resolution of a single wave. Fourier analysis provides the tool for this kind of optimization and leads to a definition of resolving efficiency. The family of schemes presented reduce to the Padé schemes if the constraint of maximal formal accuracy with a specified computational stencil is imposed. In this sense the schemes presented may be considered modified or generalized Padé schemes. Their improved resolution (spectral-like behaviour) has been demonstrated in a variety of applications. The present approach may fruitfully be extended to treat more general operators.

## APPENDIX A: HIGHER DERIVATIVES

Compact approximations for third and higher derivatives can be constructed analogous to the discussion in Sections 2.1 and 2.2. In the following we restrict our discussion to tridiagonal approximations with at least fourth-order formal accuracy. Once again the starting point is a relation of the form:

$$\alpha f'''_{i-1} + f'''_i + \alpha f'''_{i+1} = b\frac{f_{i+3} - 3f_{i+1} + 3f_{i-1} - f_{i-3}}{8h^3} + a\frac{f_{i+2} - 2f_{i+1} + 2f_{i-1} - f_{i-2}}{2h^3}, \quad \text{(A.1)}$$

where $f'''_i$ represents the finite difference approximation to the third derivative at node $i$. Each term on the r.h.s. is formally second-order accurate. The relations between the coefficients $a$, $b$ and $\alpha$ are derived by matching the Taylor series coefficients of various orders.

The constraints of fourth-order formal accuracy lead to a one-parameter family of schemes defined by

$$a = 2, \quad b = 2\alpha - 1. \quad \text{(A.2)}$$

The truncation error of this approximation (on the r.h.s. of (A.1)) is $(42/7!)(16\alpha - 7)\, h^4 f^{(7)}$.

It may be noted that as $\alpha \to 0$ the explicit fourth-order accurate form for the third derivative is recovered. For $\alpha = \frac{1}{2}$ the coefficient $b$ vanishes, generating the most compact tridiagonal scheme

$$\alpha = \tfrac{1}{2}, \quad a = 2, \quad b = 0 \quad \text{(A.3)}$$

for the third derivative. This scheme is also given by Collatz [22, p. 539]. For $\alpha = \frac{7}{16}$ a sixth-order tridiagonal scheme is obtained. This scheme with coefficients

$$\alpha = \tfrac{7}{16}, \quad a = 2, \quad b = -\tfrac{1}{8} \quad \text{(A.4)}$$

is the scheme with the highest formal accuracy amongst the schemes (A.1). Its truncation error is $(36/9!)\, h^6 f^{(9)}$.

We now summarize the resolution characteristics of the third derivative schemes. By taking the exact third derivative of (3.1.1) (with respect to $s$) a function with Fourier coefficients $\hat{f}'''_k = iw^3 \hat{f}_k$. The numerical approximations (A.1) correspond to $(\hat{f}'''_k)_{fd} = -iw'''\hat{f}_k$, where

$$w'''(w) = \frac{\alpha(2\sin(w) - \sin(2w)) + (b/4)(3\sin(w) - \sin(3w))}{1 + 2\alpha\cos(w)}. \quad \text{(A.5)}$$

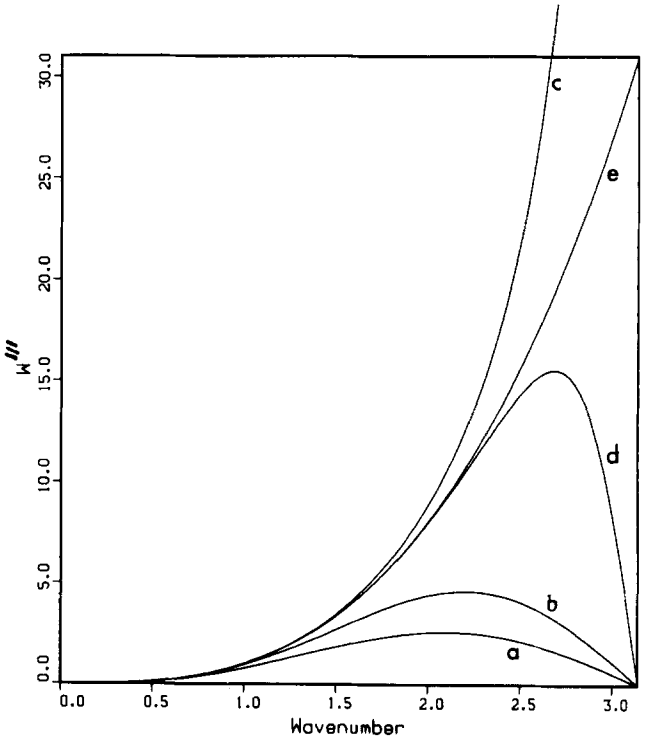


**FIG. 14.** Differencing error for third derivative vs wavenumber: (a) second-order central differences; (b) fourth-order central differences; (c) fourth-order tridiagonal scheme; (d) sixth-order tridiagonal scheme; (e) exact differentiation.

Plots of $w'''(w)$ vs. $w$ for different finite difference schemes are presented in Fig. 14. The most compact tridiagonal scheme improves over the explicit schemes but becomes singular as $w \to \pi$. This singular nature is absent in the sixth-order tridiagonal scheme and the differencing errors are further reduced.

For the fourth derivatives the starting point is a relation of the form

$$\alpha f''''_{i-1} + f''''_i + \alpha f''''_{i+1} = b\frac{f_{i+3} - 9f_{i+1} + 16f_i - 9f_{i-1} + f_{i-3}}{6h^4} + a\frac{f_{i+2} - 4f_{i+1} + 6f_i - 4f_{i-1} + f_{i-2}}{h^4}, \quad \text{(A.6)}$$

where $f''''_i$ represents the finite difference approximation to the fourth derivative at node $i$. Once again each term on the r.h.s. is formally second-order accurate.

The constraints of fourth-order formal accuracy lead to a one-parameter family of schemes defined by

$$a = 2(1-\alpha), \quad b = 4\alpha - 1. \quad \text{(A.7)}$$

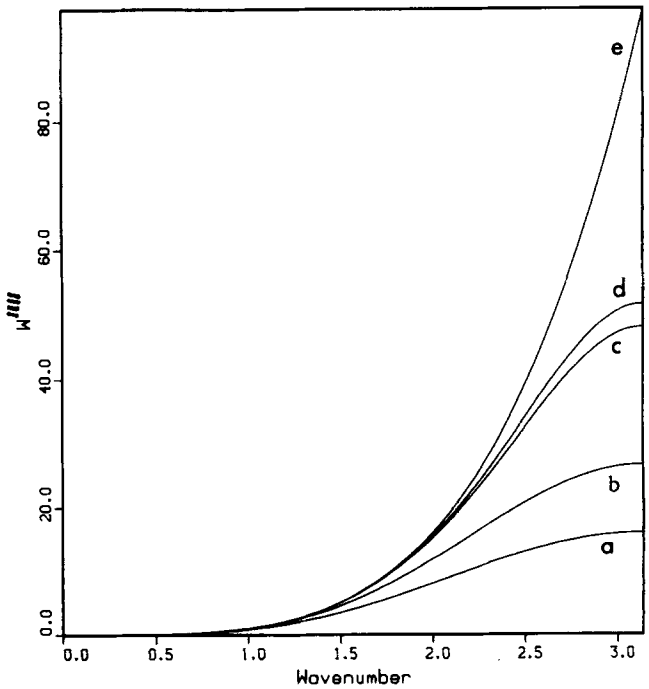


**FIG. 15.** Differencing error for fourth derivative vs wavenumber: (a) second-order central differences; (b) fourth-order central differences; (c) fourth-order tridiagonal scheme; (d) sixth-order tridiagonal scheme; (e) exact differentiation.

The truncation error of this approximation (on the r.h.s. of (A.6)) is $((7-26\alpha)/240)\,h^4 f^{(8)}$.

As $\alpha \to 0$ the explicit fourth-order accurate form for the fourth derivative is recovered. For $\alpha = \frac{1}{4}$ the coefficient $b$ vanishes and the most compact tridiagonal scheme for the fourth derivative

$$\alpha = \tfrac{1}{4}, \quad a = \tfrac{3}{2}, \quad b = 0, \quad \text{(A.8)}$$

is obtained. This scheme was given by Collatz [22, p. 539]. For $\alpha = \frac{7}{26}$ a sixth-order tridiagonal scheme is obtained. This scheme with coefficients

$$\alpha = \tfrac{7}{26}, \quad a = \tfrac{19}{13}, \quad b = \tfrac{1}{13} \quad \text{(A.9)}$$

is the scheme with the highest formal accuracy amongst the schemes (A.6). Its truncation error is $\frac{193}{393120}\,h^6 f^{(10)}$.

Finally, the exact fourth derivative of (3.1.1) (with respect to $s$) is a function with Fourier coefficients $\hat{f}''''_k = w^4 \hat{f}_k$. The numerical approximations (A.6) correspond to $(\hat{f}''''_k)_{fd} = w''''\hat{f}_k$, where

$$w''''(w) = \frac{\left(2a(\cos(2w) - 4\cos(w) + 3) + (b/3)(\cos(3w) - 9\cos(w) + 8)\right)}{1 + 2\alpha\cos(w)}. \quad \text{(A.10)}$$

Plots of $w''''(w)$ vs. $w$ for different finite difference schemes are presented in Fig. 15, showing the improvement of the compact schemes over explicit central differences.

## APPENDIX B: COMPACT SCHEMES ON A CELL-CENTERED MESH

This appendix presents the compact finite difference schemes for the first and second derivatives on a cell-centered mesh. In the formulation presented below the nodes on which the derivatives are evaluated are staggered by a half-cell ($h/2$) from the nodes on which the function values are prescribed. Such grid configurations arise naturally from a finite-volume discretization of conservation equations. As shown below the cell-centered schemes have better resolution characteristics for wavenumbers ($w$) near $\pi$ than the schemes discussed in Section 2. It should be further noted that the use of cell-centered schemes for differentiation also makes it necessary to use accurate interpolation schemes. Compact schemes for mid-point interpolation are presented in Appendix C.1. The resolution characteristics of the interpolation scheme when combined with those of the cell-centered differentiation determine the resolution characteristics of the overall scheme.

B.1. *First Derivative on a Cell-Centered Mesh*

Starting from an approximation of the form

$$\beta f'_{i-2} + \alpha f'_{i-1} + f'_i + \alpha f'_{i+1} + \beta f'_{i+2} = c\frac{f_{i+5/2} - f_{i-5/2}}{5h} + b\frac{f_{i+3/2} - f_{i-3/2}}{3h} + a\frac{f_{i+1/2} - f_{i-1/2}}{h}, \quad \text{(B.1.1)}$$

the constraints on the coefficients are derived by matching the Taylor series coefficients (at least up to fourth order). Tridiagonal schemes analogous to the standard Padé scheme are obtained with $\beta = 0$, $c = 0$. These fourth-order schemes are defined by

$$\beta = 0, \quad a = \tfrac{3}{8}(3 - 2\alpha), \quad b = \tfrac{1}{8}(22\alpha - 1). \quad \text{(B.1.2)}$$

The truncation error (on the r.h.s. of B.1.1)) is $((9 - 62\alpha)/1920)\, h^4 f^{(5)}$.

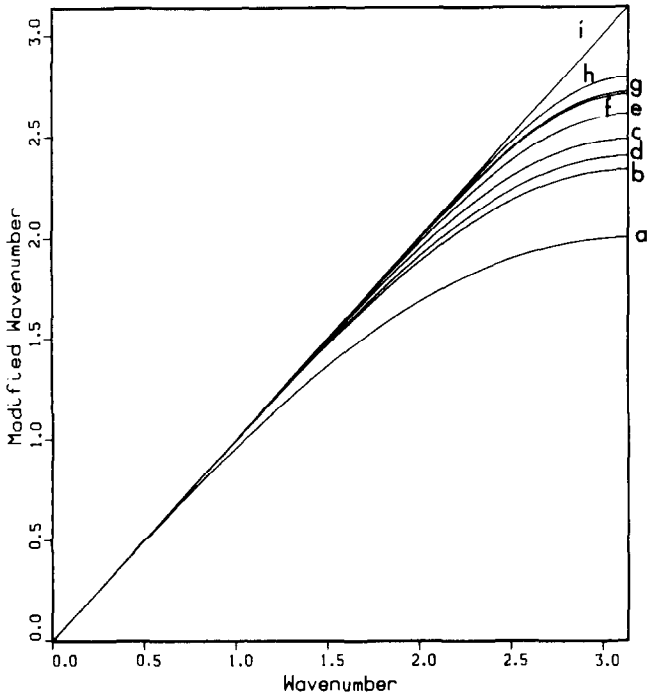


**FIG. 16.** Plot of modified wavenumber vs wavenumber for the cell-centered first derivative approximations: (a) second-order central differences; (b) fourth-order central differences; (c) sixth-order central differences; (d) fourth-order tridiagonal scheme ($\beta = b = c = 0$, $\alpha = \frac{1}{22}$); (e) sixth-order tridiagonal scheme ($\beta = c = 0$); (f) eighth-order tridiagonal scheme ($\beta = 0$, $\alpha = \frac{75}{554}$); (g) eighth-order pentadiagonal scheme ($c = 0$); (h) tenth-order pentadiagonal scheme; (i) exact differentiation.

For $\alpha = \frac{1}{22}$ the coefficient $b$ vanishes, generating the most compact scheme. For $\alpha = \frac{9}{62}$ a sixth-order tridiagonal scheme is generated. This sixth-order scheme has small differencing errors over a larger range of wavenumbers. Characterization of the differencing error via a plot of the modified wavenumber of the scheme is presented in Fig. 16. Also shown on this figure are other higher order schemes generated by (B.1.1.). As expected the schemes on the cell-centered mesh have considerably lower differencing errors compared to the unstaggered schemes (Fig. 1). The sixth order tridiagonal scheme ($\alpha = \frac{9}{62}$ with (B.1.2)) is quite close to exact differentiation. Further improvements can be obtained by allowing for $\beta \neq 0$ and $c \neq 0$. Once again it is observed that the resolution of the short scales is not very sensitive to the formal accuracy or truncation error of the scheme.

Coefficients of the higher order schemes obtained from (B.1.1) and their truncation error (on the r.h.s. of (B.1.1)) are summarized in Table VI.

**TABLE VI**

Cell-Centered First Derivative Schemes

| Scheme | Parameters | Truncation error in (B.1.1) |
|---|---|---|
| (B.1.3) | $a = \frac{1}{8}(9 - 6\alpha - 78\beta + 16c)$ | |
| | $b = \frac{1}{8}(-1 + 22\alpha + 94\beta - 24c)$ | $\frac{1}{1920}(9 - 62\alpha + 1618\beta - 384c)\, h^4 f^{(5)}$ |
| (B.1.4) | $a = \frac{225 - 206\alpha - 254\beta}{192}$ | |
| | $b = \frac{414\alpha - 114\beta - 25}{128}$ | |
| | $c = \frac{9 - 62\alpha + 1618\beta}{384}$ | $\frac{(75 - 354\alpha + 2614\beta)}{107520} h^6 f^{(7)}$ |
| (B.1.5) | $\beta = \frac{354\alpha - 75}{2614}$ | |
| | $a = \frac{37950 - 39275\alpha}{31368}$ | |
| | $b = \frac{65115\alpha - 3550}{20912}$ | |
| | $c = \frac{25669\alpha - 6114}{62736}$ | $\frac{(96850 - 288529\alpha)}{1686343680} h^8 f^{(9)}$ |
| (B.1.6) | $\beta \approx \frac{9675}{577058}, \alpha \approx \frac{96850}{288529}$ | |
| | $a = \frac{683425}{865587}, b \approx \frac{505175}{577058}$ | |
| | $c = \frac{69049}{1731174}$ | $\frac{939109}{818997645312} h^{10} f^{(11)}$ |

B.2. *Second Derivative on a Cell-Centered Mesh*

For obtaining compact approximations for the second derivatives we start from the equation:

$$\beta f''_{i-2} + \alpha f''_{i-1} + f''_i + \alpha f''_{i+1} + \beta f''_{i+2} = 4c\frac{f_{i+5/2} - 2f_i + f_{i-5/2}}{25h^2} + 4b\frac{f_{i+3/2} - 2f_i + f_{i-3/2}}{9h^2} + 4a\frac{f_{i+1/2} - 2f_i + f_{i-1/2}}{h^2}, \quad \text{(B.2.1)}$$

the constraints on the coefficients are derived by matching the Taylor series coefficients (at least up to fourth order). Tridiagonal schemes analogous to the standard Padé scheme are obtained with $\beta = 0$, $c = 0$. These fourth-order schemes are defined by

$$\beta = 0, \quad a = \tfrac{3}{8}(3 - 10\alpha), \quad b = \tfrac{1}{8}(46\alpha - 1). \quad \text{(B.2.2)}$$

The truncation error (on the r.h.s. of (B.2.1)) is $\frac{1}{640}(1 + 2\alpha)\, h^4 f^{(6)}$.

For $\alpha = \frac{1}{46}$ the coefficient $b$ vanishes generating the "standard" Padé scheme. For $\alpha = -\frac{1}{2}$ a sixth-order tridiagonal scheme is generated. This sixth-order scheme is singular at $w = \pi$. The singular nature may be removed by either using $\alpha \neq -\frac{1}{2}$ for tridiagonal schemes or by using schemes with $c \neq 0$ or $\beta \neq 0$. Characterization of the differencing error of schemes generated by (B.2.1) is presented in Fig. 17. Also shown on this figure are other higher order schemes generated by (B.2.1). It may be noted that even the fourth- and sixth-order explicit forms have quite small differencing errors. Thus on a cell-centered mesh compact schemes can provide almost exact second derivatives. The eighth-order tridiagonal scheme slightly overpredicts the second derivative for $w = \pi$ and the pentadiagonal schemes are virtually indistiguishable from the exact differentiation. The coefficients of the higher order schemes are summarized in Table VII.

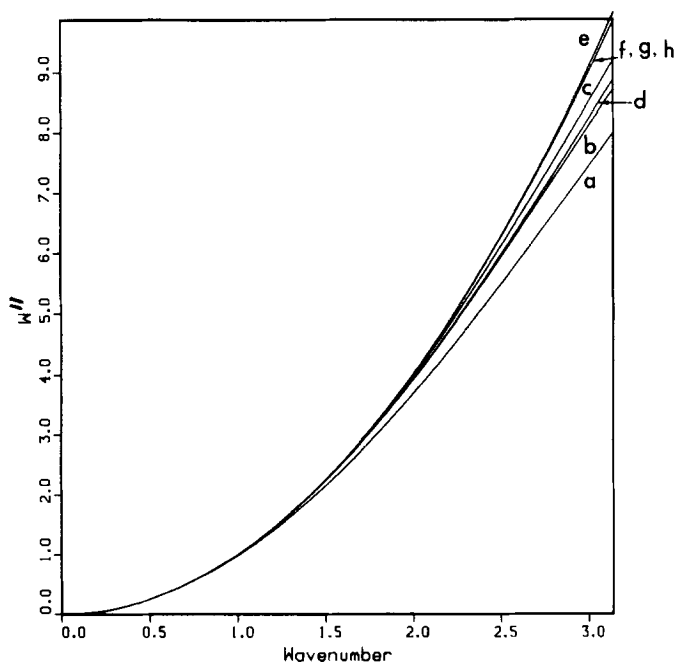


**FIG. 17.** Differencing error vs wavenumber for the cell-centered second derivative approximations: (a) second-order central differences; (b) fourth-order central differences; (c) sixth-order central differences; (d) fourth-order tridiagonal scheme ($\beta = c = 0$, $\alpha = \frac{1}{46}$); (e) eighth-order tridiagonal scheme ($\beta = 0$); (f) eighth-order pentadiagonal scheme ($c = 0$); (g) tenth-order pentadiagonal scheme; (h) exact differentiation.

**TABLE VII**

Cell-Centered Second Derivative Schemes

| Scheme | Parameters | Truncation error in (B.2.1) |
|---|---|---|
| (B.2.3) | $a = \frac{1}{8}(9 - 30\alpha - 174\beta + 16c)$ | |
| | $b = \frac{1}{8}(-1 + 46\alpha + 190\beta - 24c)$ | $\frac{1}{1920}(3 + 6\alpha + 1926\beta - 128c)\, h^4 f^{(6)}$ |
| (B.2.4) | $a = \frac{75 - 234\alpha + 534\beta}{64}$ | |
| | $b = \frac{718\alpha - 2738\beta - 25}{128}$ | |
| | $c = \frac{3 + 6\alpha + 1926\beta}{128}$ | $\frac{(225 + 1234\alpha - 9854\beta)}{1290240} h^6 f^{(8)}$ |
| (B.2.5) | $\beta = \frac{1234\alpha + 225}{9854}$ | |
| | $a = \frac{26850 - 51465\alpha}{19708}$ | |
| | $b = \frac{115515\alpha - 26950}{39416}$ | |
| | $c = \frac{76119\alpha - 14466}{39416}$ | $\frac{-(28750 + 81583\alpha)}{3178506240} h^8 f^{(10)}$ |
| (B.2.6) | $\alpha = -\frac{28750}{81583}, \beta = -\frac{3475}{163166}$ | |
| | $a = \frac{186225}{81583}, b = -\frac{16475}{9598}$ | |
| | $c = -\frac{51159}{163166}$ | $\frac{950813}{1389453778944} h^{10} f^{(12)}$ |

## APPENDIX C: COMPACT SCHEMES FOR INTERPOLATION AND FILTERING

Compact schemes can be easily constructed for interpolation and filtering applications. It is useful to analyze such operations in the wavenumber (Fourier) space of the

modes represented on the mesh [24, Chap. 5]. It may be anticipated that the compact schemes provide a greater control over the shape of the transfer function (in the wavenumber space). Once again the desirable characteristics of the transfer function may be optimized more effectively by not insisting on the highest formal accuracy. Compact schemes for mid-point interpolation are presented first. Similar schemes may be derived for interpolation at other (uncentered) points. Filtering applications are described next.

C.1. *Midpoint Interpolation*

Beginning with an approximation of the form

$$\beta\hat{f}_{i-2}+\alpha\hat{f}_{i-1}+\hat{f}_i+\alpha\hat{f}_{i+1}+\beta\hat{f}_{i+2} = \frac{c}{2}(f_{i+5/2}+f_{i-5/2})+\frac{b}{2}(f_{i+3/2}+f_{i-3/2}) + \frac{a}{2}(f_{i+1/2}+f_{i-1/2}), \tag{C.1.1}$$

where $\hat{f}_i$ represents the interpolated values at node $x_i$, schemes of different formal accuracy may be derived by matching the Taylor series coefficients of various order. Typically we require at least a fourth-order formal accuracy. The transfer function associated with (C.1.1),

$$T(w)=\frac{a\cos(w/2)+b\cos(3w/2)+c\cos(5w/2)}{1+2\alpha\cos(w)+2\beta\cos(2w)}, \tag{C.1.2}$$

is used to further optimize the schemes.

The fourth-order family (with three parameters) is defined by

$$a=\tfrac{1}{8}(9+10\alpha-14\beta+16c), \quad b=\tfrac{1}{8}(-1+6\alpha+30\beta-24c). \tag{C.1.3}$$

**TABLE VIII**

Truncation Error for the Interpolation Scheme

| Scheme | Max. l.h.s. stencil size | Max. r.h.s. stencil size | Truncation error in (C.1.1) |
|---|---|---|---|
| (C.1.3) | 5 | 6 | $\frac{1}{128}(3-10\alpha+70\beta-128c)\,h^4f^{(4)}$ |
| (C.1.4) | 5 | 6 | $\frac{1}{1024}(5-14\alpha+42\beta)\,h^6f^{(6)}$ |
| (C.1.5) | 5 | 6 | $\frac{10-21\alpha}{28672}\,h^8f^{(8)}$ |
| (C.1.6) | 5 | 6 | $\frac{1}{258048}\,h^{10}f^{(10)}$ |

Its formal truncation error (on r.h.s. of (C.1.1)) is tabulated in Table VIII along with those of other schemes. The sixth-order family (with two parameters) is defined by

$$a=\tfrac{1}{64}(75+70\alpha-42\beta), \quad b=\tfrac{1}{128}(270\beta+126\alpha-25), \quad c=\tfrac{1}{128}(70\beta-10\alpha+3). \tag{C.1.4}$$

This may be specialized into an eighth-order (one-parameter) family

$$\beta=\frac{14\alpha-5}{42}, \quad a=\frac{7\alpha+10}{8}, \quad b=\frac{189\alpha-50}{112}, \quad c=\frac{5\alpha-2}{48} \tag{C.1.5}$$

and, further, into a single tenth-order scheme,

$$\alpha=\tfrac{10}{21}, \quad \beta=\tfrac{5}{126}, \quad a=\tfrac{5}{3}, \quad b=\tfrac{5}{14}, \quad c=\tfrac{1}{126}. \tag{C.1.6}$$

The transfer function of several interpolation schemes is plotted in Fig. 18 against the wavenumber $w$. The improvement of the compact schemes over the explicit interpolation schemes is evident. Also shown on the figure is a "spectral-

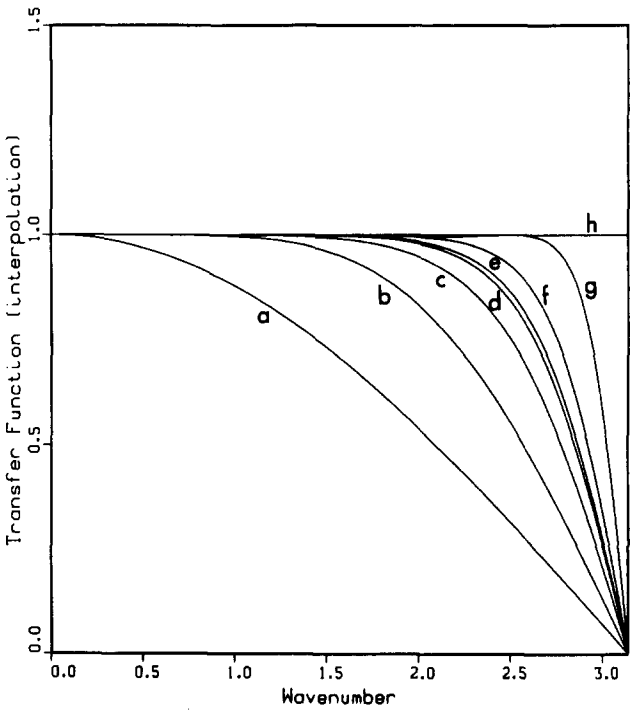


**FIG. 18.** Interpolation transfer function vs wavenumber for different explicit and compact schemes: (a) second-order explicit; (b) sixth-order explicit; (c) sixth-order tridiagonal scheme ($\beta=c=0$); (d) eighth-order tridiagonal scheme ($\beta=0$); (e) eighth-order pentadiagonal scheme ($c=0$); (f) tenth-order pentadiagonal scheme; (g) spectral-like pentadiagonal scheme (C.1.8); (h) exact.

like" interpolation scheme. This (formally fourth-order) scheme is obtained by requiring

$$T(w_*)=1, \quad \frac{dT(w=w_*)}{dw}=0. \tag{C.1.7}$$

The scheme shown in the figure is obtained for $w_*=2.5$. Its parameters are

$$\alpha=0.5579841, \quad \beta=0.07425165, \quad a=1.733442, \quad b=0.5105776, \quad c=0.02045137. \tag{C.1.8}$$

C.2. *Filtering*

Beginning with an approximation of the form

$$\beta\hat{f}_{i-2}+\alpha\hat{f}_{i-1}+\hat{f}_i+\alpha\hat{f}_{i+1}+\beta\hat{f}_{i+2} = af_i+\frac{d}{2}(f_{i+3}+f_{i-3})+\frac{c}{2}(f_{i+2}+f_{i-2}) +\frac{b}{2}(f_{i+1}+f_{i-1}), \tag{C.2.1}$$

where $\hat{f}_i$ represents the filtered values at node $x_i$, the filtering application discussed here is the removal of short length scales. The problem is most naturally formulated in terms of the transfer function associated with (C.2.1)

$$T(w)=\frac{a+b\cos(w)+c\cos(2w)+d\cos(3w)}{1+2\alpha\cos(w)+2\beta\cos(2w)}, \tag{C.2.2}$$

For the filters we require $T(\pi)=0$. For all schemes discussed here this constraint automatically provides $(dT/dw)(\pi)=0$. For some of the filtering schemes discussed here we pose additional constraints

$$\frac{d^2T}{dw^2}(\pi)=0, \quad \frac{d^4T}{dw^4}(\pi)=0. \tag{C.2.3}$$

Schemes of different formal accuracy may be derived by matching the Taylor series coefficients of various orders, along with the filtering constraints. Typically we require at least a fourth-order formal accuracy. If the constraints (C.2.3) are not imposed a three-parameter family of fourth-order schemes is generated. Its coefficients are

$$a=\tfrac{1}{8}(5+6\alpha-6\beta+16d), \quad b=\tfrac{1}{2}(1+2\alpha+2\beta-2d), \quad c=\tfrac{-1}{8}(1-2\alpha-14\beta+16d). \tag{C.2.4}$$

The truncation error for this scheme is listed in Table IX

**TABLE IX**

Truncation Error for the Filtering Schemes

| Scheme | Max. l.h.s. stencil size | Max. r.h.s. stencil size | Truncation error in (C.2.1) |
|---|---|---|---|
| (C.2.4) | 5 | 7 | $\frac{1}{16}(1-2\alpha+2\beta-32d)\,h^4f^{(4)}$ |
| (C.2.5) | 5 | 7 | $\frac{1}{64}(1-2\alpha+2\beta)\,h^6f^{(6)}$ |
| (C.2.6) | 5 | 7 | $\frac{1}{16}(3-2\alpha-10\beta)\,h^4f^{(4)}$ |
| (C.2.7) | 5 | 7 | $\frac{1}{80}(2-3\alpha)\,h^6f^{(6)}$ |
| (C.2.8) | 5 | 7 | $\frac{1}{40}h^6f^{(6)}$ |

along with those for other schemes. If the sixth-order coefficients in (C.2.1) are also matched then (C.2.4) reduces to

$$a=\tfrac{1}{16}(11+10\alpha-10\beta), \quad b=\tfrac{1}{32}(15+34\alpha+30\beta), \quad c=\tfrac{1}{16}(-3+6\alpha+26\beta), \quad d=\tfrac{1}{32}(1-2\alpha+2\beta). \tag{C.2.5}$$

Attempts to match any further Taylor series coefficients reduce (C.2.1) to an identity.

A different family of schemes is generated when the first of (C.2.3) is imposed in addition to formal fourth-order accuracy. This two-parameter family of schemes is defined by

$$a=\tfrac{1}{4}(2+3\alpha), \quad b=\tfrac{1}{16}(9+16\alpha+10\beta), \quad c=\tfrac{1}{4}(\alpha+4\beta), \quad d=\tfrac{1}{16}(6\beta-1). \tag{C.2.6}$$

Posing the sixth-order constraint on (C.2.6) generates

$$\beta=\frac{3-2\alpha}{10}, \quad a=\frac{2+3\alpha}{4}, \quad b=\frac{6+7\alpha}{8}, \quad c=\frac{6+\alpha}{20}, \quad d=\frac{2-3\alpha}{40}. \tag{C.2.7}$$

If the second of the filtering constraints (C.2.3) is also imposed on (C.2.7), a single scheme with coefficients

$$\alpha=0, \quad \beta=\tfrac{3}{10}, \quad a=\tfrac{1}{2}, \quad b=\tfrac{3}{4}, \quad c=\tfrac{3}{10}, \quad d=\tfrac{1}{20} \tag{C.2.8}$$

is obtained.

A different way to optimize the schemes is to pose specific conditions on the shape of the transfer function $T(w)$. This

allows the location of the cutoff to be adjusted. This may be done by requiring

$$T(w_1)=v_1, \quad T(w_2)=v_2, \tag{C.2.9}$$

where $w_1$, $w_2$, $v_1$, and $v_2$ are specified. These constraints along with (C.2.6) determine all the coefficients.

Tridiagonal schemes may be obtained as special cases of (C.2.4), (C.2.5), or (C.2.6). In Fig. 19 the transfer function for the tridiagonal schemes with $\beta=0$ and $d=0$ is compared to explicit filtering schemes of second-, fourth-, and sixth-order accuracy. Evidently the compact schemes are much better low pass filters. By reducing the coefficient $\alpha$ towards 0.5 the filtering effect is confined more towards the shorter waves. In Fig. 20 the transfer function for the pentadiagonal schemes is shown. It is seen that the choice of the thresholds in (C.2.9) determines the cutoff wavenumber of the filtering. Schemes (b) and (c) on this figure correspond to the coefficients

$$\alpha=0.4627507, \quad \beta=0.2265509, \quad a=0.8470630, \quad b=1.166845, \quad c=0.3422386, \quad d=0.02245659, \tag{C.2.10.a}$$

$$\alpha=0.6522474, \quad \beta=0.1702929, \quad a=0.9891856, \quad b=1.321180, \quad c=0.3333548, \quad d=0.001359850. \tag{C.2.10.b}$$

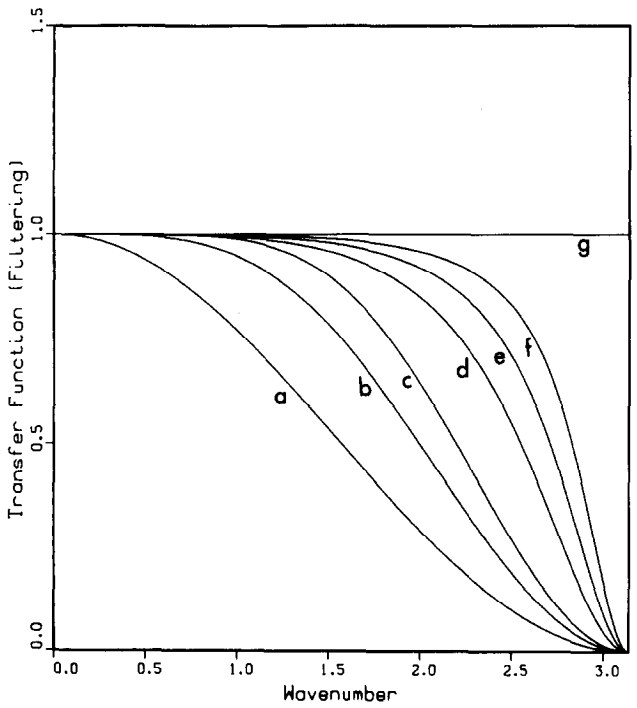


**FIG. 19.** Filtering transfer function vs wavenumber for different explicit and tridiagonal compact schemes: (a) second-order explicit; (b) fourth-order explicit; (c) sixth-order explicit; (d) fourth-order tridiagonal scheme ($\alpha=0.4, \beta=d=0$); (e) fourth-order tridiagonal scheme ($\alpha=0.45$, $\beta=d=0$); (f) fourth-order tridiagonal scheme ($\alpha=0.475$. $\beta=d=0$); (g) no filtering.

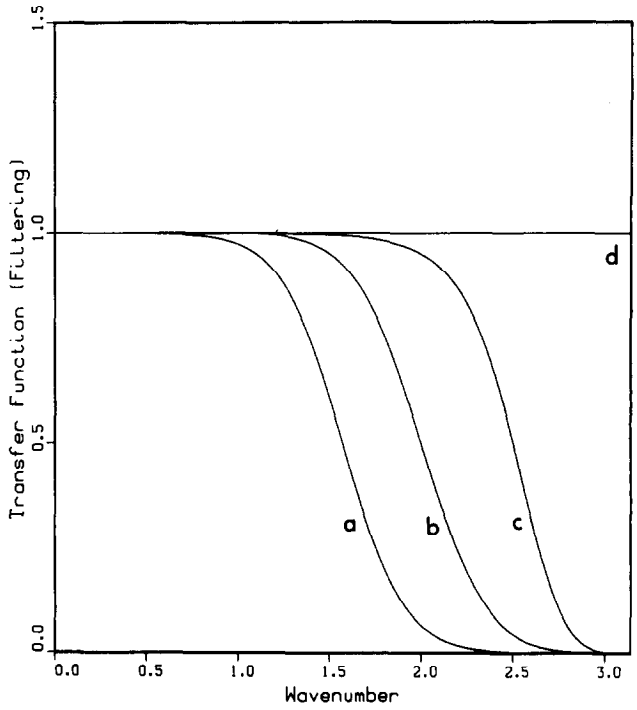


**FIG. 20.** Filtering transfer function vs wavenumber for pentadiagonal compact schemes: (a) sixth-order scheme (C.2.7); (b) fourth-order scheme (C.2.6) with ($w_1=1.5$, $v_1=0.95$, $w_2=2.0$, $v_2=0.5$); (c) fourth-order scheme (C.2.6) with ($w_1=2.0$, $v_1=0.95$, $w_2=2.5$, $v_2=0.5$); (d) no filtering.

Implementation of the filtering schemes on domains with non-periodic boundaries requires the near boundary nodes to be treated separately. The boundary schemes can also be devised to exactly filter the $w=\pi$ waves. Explicit near boundary formulas (of fourth-order formal accuracy and exact filtering of $w=\pi$) are

$$\hat{f}_1=\tfrac{15}{16}f_1+\tfrac{1}{16}(4f_2-6f_3+4f_4-f_5), \tag{C.2.11.a}$$

$$\hat{f}_2=\tfrac{3}{4}f_2+\tfrac{1}{16}(f_1+6f_3-4f_4+f_5), \tag{C.2.11.b}$$

$$\hat{f}_3=\tfrac{5}{8}f_3+\tfrac{1}{16}(-f_1+4f_2+4f_4-f_5). \tag{C.2.11.c}$$

Their formal truncation errors are given by $-(3/2\cdot 4!)h^4f_1^{(4)}$, $(11/8\cdot 4!)h^4f_2^{(4)}$, and $-(3/2\cdot 4!)h^4f_3^{(4)}$, respectively.

## ACKNOWLEDGMENTS

This work started with the support from the Director's discretionary fund of NASA–Ames Research Center. Its continued support from CTR through research sponsored by ONR and AFOSR is gratefully acknowledged. The author is grateful to Dr. A. Wray and Professor P. Moin for introducing him to the Fourier analysis of differencing schemes and for many friendly discussions. He is thankful to them and to Dr. T. Pulliam, Dr. H. Lomax, and Dr. R. F. Warming for remarks on a draft of this paper.

## REFERENCES


1. D. Gottlieb and S. A. Orszag, *Numerical Analysis of Spectral Methods* (SIAM, Philadelphia, 1977).
2. C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods in Fluid Dynamics* (Springer-Verlag, New York, 1987).
3. R. Rogallo and P. Moin, *Annu. Rev. Fluid Mech.* **16**, 99 (1984).
4. J. Kim, P. Moin, and R. D. Moser, *J. Fluid Mech.* **177**, 133 (1987).
5. P. Spalart, *J. Fluid Mech.* **187**, 61 (1988).
6. M. M. Rai and P. Moin, AIAA paper, AIAA-89-0369, Reno, 1989.
7. A. T. Patera, *J. Comput. Phys.* **54**, 468 (1984).
8. N. K. Ghaddar, G. E. Kariadakis, and A. T. Patera, *Numer. Heat Transfer* **9**, 277 (1986).
9. G. E. Karniadakis, *Appl. Numer. Math.* 5 (1989).
10. B. Swartz and B. Wendroff, in Lecture Notes in Mathematics, Vol. 363, *Proceedings, Conference on the Numerical Solution of Differential Equations*, Dundee, 1973, edited by G. A. Watson (Springer-Verlag, New York/Berlin, 1974), p. 153.
11. B. Swartz, in *Mathematical Aspects of Finite Elements in Partial Differential Equations*, edited by C. de Boor (Academic Press, New York, 1974), p. 279.
12. B. Swartz and B. Wendroff, *SIAM J. Numer. Anal.* **11**, 979 (1974).
13. B. Swartz, in *Advances in Computer Methods for Partial Differential Equations*, edited by R. Vichnevetsky (AICA, Ghent, Belgium, 1975), p. 17.
14. H. O. Kreiss and J. Oliger, *Tellus* **3**, 99 (1972).
15. H. O. Kreiss, S. A. Orszag, and M. Israeli, *Annu. Rev. Fluid Mech.* **6**, 281 (1974).
16. R. S. Hirsh, *J. Comput. Phys.* **19**, 90 (1975).
17. Y. Adam, *J. Comput. Phys.* **24**, 10 (1977).
18. N. N. Mansour, P. Moin, W. C. Reynolds, and J. H. Ferziger, *Turbulent Shear Flows I* (Springer-Verlag, New York/Berlin, 1977), p. 386.
19. S. G. Rubin and P. K. Khosla, *J. Comput. Phys.* **24**, 217 (1977).
20. Y. Adam, *Comput. Math. Appl.* **1**, 393 (1975).
21. W. J. Goedheer and J. H. H. M. Potters, *J. Comput. Phys.* **61**, 269 (1985).
22. L. Collatz, *The Numerical Treatment of Differential Equations* (Springer-Verlag, New York, 1966), p. 538.
23. Z. Kopal, *Numerical Analysis* (Wiley, New York, 1961), p. 552.
24. R. Vichnevetsky and J. B. Bowles, *Fourier Analysis of Numerical Approximations of Hyperbolic Equations* (SIAM, Philadelphia, 1982).
25. K. V. Roberts and N. O. Weiss, *Math. Comput.* **20**, 272 (1966).
26. J. E. Fromm, *Phys. Fluids (Suppl. II)*, II-3, II-12 (1969).
27. S. A. Orszag, *J. Fluid Mech.* **49**, 75 (1971).
28. S. A. Orszag, *Stud. Appl. Math.* **49**, 395 (1971).
29. J. A. Trapp and J. D. Ramshaw, *J. Comput. Phys.* **20**, 238 (1976).
30. R. Vichnevetsky, *Math. Comput. Simul.* **21**, 170 (1979).
31. J. R. Herring, S. A. Orszag, R. H. Kraichnan, and D. G. Fox, *J. Fluid Mech.* **66**, 417 (1974).
32. R. Rogallo, NASA TM-81315, 1981.
33. T. A. Zang, *Appl. Num. Math.*, to appear.
34. K. Horiuti, *J. Comput. Phys.* **71**, 343 (1987).
35. R. M. Kerr, *J. Fluid Mech.* **153**, 31 (1985).
36. S. K. Lele, AIAA paper, AIAA-89-0374, Reno, 1989.
37. N. D. Sandham and W. C. Reynolds, Thermosciences Division, Mech. Engrg., Stanford University, Report No. TF-45, 1989; AIAA paper, AIAA-89-0371, Reno, 1989.
38. J. Strikwerda, *J. Comput. Phys.* **34**, 94 (1980).
39. R. F. Warming and R. M. Beam, in *Proceedings, International Symposium on Computational Fluid Dynamics, Sydney, Australia*, edited by G. de Vahl Davis and C. Fletcher, (North-Holland, Amsterdam, 1988), p. 749; *Lecture Notes in Physics*, Vol. 264 (Springer-Verlag, New York/Berlin, 1986), p. 647.
40. K. Thompson, *J. Comput. Phys.* **68**, 1 (1987).
41. T. Poinsot and S. K. Lele, *J. Comput. Phys.*, to appear (1992); Center for Turbulence Research, Stanford University, CTR Manuscript 102, 1989.
42. J. C. Buell, *J. Comput. Phys.* **95**, 313–338 (1991); in *Proceedings, Turbulent Shear Flows - 7 (1989)*, Vol. 1, p. 9.2.1; in *Proceedings, Tenth Australasian Fluid Mechanic Conference, Melbourne, 1989.*
43. J. C. Buell and P. Huerre, in *Proceedings, 1988 Summer Program*; Report CTR-S88, Center for Turbulence Research, Stanford University, 1988, p. 19.
44. D. Papamoschou and A. Roshko, *J. Fluid Mech.* **197**, 453 (1988).
45. P. A. Lagerstrom, *Theory of Laminar Flows, High Speed Aerodynamics and Jet Propulsion, 4*, edited by F. K. Moore (Princeton Univ. Press, Princeton, NJ, 1964), pp. 175, 453.
46. A. George and J. W-H. Liu, *Computer Solution of Large Sparse Positive Definite Systems* (Prentice–Hall, Englewood Cliffs, NJ, 1981), p. 51.
47. E. O. Brigham, *The Fast Fourier Transform* (Prentice–Hall, Englewood Cliffs, NJ, 1974), p. 194.