NASA Contractor Report 181803

ICASE REPORT NO. 89-13

ICASE

SPECTRAL METHODS FOR CFD

Thomas A. Zang

Craig L. Streett

M. Yousuff Hussaini

Contract No. NAS1-18605

February 1989

INSTITUTE FOR COMPUTER APPLICATIONS IN SCIENCE AND ENGINEERING
NASA Langley Research Center, Hampton, Virginia 23665

Operated by the Universities Space Research Association

National Aeronautics and Space Administration

Langley Research Center
Hampton, Virginia 23665

SPECTRAL METHODS FOR CFD

Thomas A. Zang[1] [2]
NASA Langley Research Center
Hampton, Virginia

Craig L. Streett[3]
NASA Langley Research Center
Hampton, Virginia

M. Y. Hussaini[4]
Institute for Computer Applications in Science and Engineering
NASA Langley Research Center
Hampton, Virginia

Abstract

One of the objectives of these notes is to provide a basic introduction to spectral methods with a particular emphasis on applications to computational fluid dynamics. Another objective is to summarize some of the most important developments in spectral methods in the last two years. The fundamentals of spectral methods for simple problems will be covered in depth, and the essential elements of several fluid dynamical applications will be sketched.

[1]This research was supported in part by the National Aeronautics and Space Administration under NASA Contract No. NAS1-18605 while the third author was in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA 23665.

[2]Research Scientist, Computational Methods Branch, High-Speed Aerodynamics Division.

[3]Research Scientist, Theoretical Aerodynamics Branch, High-Speed Aerodynamics Division

[4]Chief Scientist

SPECTRAL METHODS FOR CFD

Thomas A. Zang
Computational Methods Branch

Craig L. Streett
Theoretical Aerodynamics Branch

and

M. Yousuff Hussaini
I.C.A.S.E.

NASA Langley Research Center
Hampton, Virginia 23665

Contents

1 Introduction

As a tool for large-scale computations in fluid dynamics, spectral methods were prophesied in 1944, born in 1954, virtually buried in the mid-1960's, resurrected in 1969, evangalized in the 1970's, and catholicized in the 1980's. The use of spectral methods for meteorological problems was proposed by Blinova in 1944 and the first numerical computations were conducted by Silberman (1954). By the early 1960's computers had achieved sufficient power to permit calculations with hundreds of degrees of freedom. For problems of this size the traditional way of computing the nonlinear terms in spectral methods was outrageously expensive compared with finite-difference methods. Consequently, spectral methods fell out of favor. The expense of computing nonlinear terms remained a severe drawback until Orszag (1969) and Eliasen, Machenauer, and Rasmussen (1970) developed the transform methods that still form the backbone of many large-scale spectral computations. The original proselytes of spectral methods were meteorologists involved in global weather modeling and fluid dynamicists investigating isotropic turbulence. The converts who were inspired by the successes of these pioneers remained, for the most part, confined to these and closely related fields throughout the 1970's. During that decade spectral methods appeared to be well-suited only for problems governed by ordinary differential equations or by partial differential equations with periodic boundary conditions. And, of course,

the solution itself needed to be smooth. Some of the obstacles to wider application of spectral methods were: 1) poor resolution of discontinuous solutions; 2) inefficient implementation of implicit methods; and 3) drastic geometric constraints. All of these barriers have undergone some erosion during the 1980's, particularly the latter two. As a result, the applicability and appeal of spectral methods for computational fluid dynamics has broadened considerably.

The motivation for the use of spectral methods in numerical calculations stems from the attractive approximation properties of orthogonal polynomial expansions. Suppose, for example, that a function, $u(x)$, is expanded in a Legendre series on [-1,1]:

$$u^N(x) = \sum_{n=0}^{\infty} \hat{u}_n L_n(x), \tag{1}$$

where $L_n(x)$ is the Legendre polynomial of degree n. The classical form of the expansion coefficients (or spectra) is

$$\hat{u}_n = (k + \frac{1}{2}) \int_{-1}^{1} u(x) L_n(x) dx, \qquad for\ n \geq 0. \tag{2}$$

An integration-by-parts argument reveals that

$$n^p \hat{u}_n \to 0 \qquad as\ n \to \infty, \qquad for\ all\ p > 0, \tag{3}$$

provided that u is infinitely differentiable. Consequently, the approximation error decreases faster than algebraically. This rapid convergence is referred to as infinite-order accuracy, exponential convergence, or spectral accuracy. Our primary concern in this review is on numerical methods for partial differential equations that exhibit spectral accuracy for infinitely differentiable solutions.

The two most common types of finite approximations are based on truncation and interpolation. The Legendre series truncated after N terms is

$$P_N u(x) = \sum_{n=0}^{N} \hat{u}_n L_n(x). \tag{4}$$

It forms the foundation of Legendre Galerkin and tau methods. The Legendre interpolant of $u(x)$, on the other hand, is used for collocation approximations, and is given by

$$I_N u(x) = \sum_{n=0}^{N} \tilde{u}_n L_n(x), \tag{5}$$

where the coefficients, $\tilde{u}_n$, of the interpolant are determined by the condition,

$$I_N u(x_j) = u(x_j), \qquad for\ j = 0, 1, \ldots, N, \tag{6}$$

special, so-called collocation, points in [-1,1], denoted by x_j. For most problems, the optimal choice of these collocation points is

$$x_j = \begin{cases} -1 & j = 0 \\ j'th\ zero\ of\ L_N' & j = 1, 2, \ldots, N-1 \\ +1 & j = N \end{cases} . \tag{7}$$

This choice of collocation points yields an extremely accurate approximation to the integral appearing in Eq. (2):

$$\tilde{u}_n = \frac{1}{\gamma_n} \sum_{j=0}^{N} u(x_j) L_n(x_j) w_j, \qquad for\ j = 0, 1, \ldots, N, \tag{8}$$

where the Gauss-Lobatto weights,

$$w_j = \frac{2}{N(N+1)} \frac{1}{L_N^2(x_j)}, \qquad for\ j = 0, 1, \ldots, N, \tag{9}$$

and the normalization constants,

$$\gamma_n = \begin{cases} \frac{1}{n+\frac{1}{2}} & 0 \le n \le N-1 \\ \frac{2}{N} & n = N \end{cases} . \tag{10}$$

Whether (2) or (6) is used for the expansion coefficients, the expansion (1) is differentiated analytically to form the approximations to whatever derivatives are required for the problem at hand.

A graphical distinction between traditional approximations and spectral ones is provided in Figure 1 for the simple task of estimating the derivative of the function $1 + sin(2\pi x + \pi/4)$ on [-1,1] from the values of the function at a finite number of grid-points. A finite-difference or finite-element approximation uses local information to estimate derivatives, whereas a spectral approximation uses global information. In this figure a second-order (central) finite-difference approximation is compared with a Legendre spectral collocation approximation. The finite-difference approximation estimates the derivative at, say, $x = 0$, from the parabola which interpolates the function at $x = 0$ and the two adjacent grid-points. A separate parabola is used at each grid-point. The spectral approximation, on the other hand, uses all the available information about the function. If there are $N + 1$ grid-points, then the interpolating polynomial from which the derivative is extracted has degree N, and the same polynomial is used for all the grid-points. For small N the accuracy of the spectral and the finite-difference results are comparable. However, as N increases, the accuracy of the spectral approximation increases dramatically (its error decays exponentially), whereas the finite-difference result improves only algebraically.

The high accuracy achieved by the spectral method for a modest number of degrees of freedom on this simple one-dimensional approximation problem has been demonstrated repeatedly for multi-dimensional, nonlinear problems in fluid dynamics. This is particulary true for those computations, such as the numerical simulation of transition to turbulence, for which great accuracy is essential rather than a luxury. Figure 2, taken from a spectral simulation by Streett and Hussaini (1989), exemplifies this. The cascade of bifurcations leading to an aperiodic, weakly turbulent (chaotic) state in Taylor-Couette flow is an established experimental fact, but has proven to be an elusive goal for numerical simulations. In order to capture properly the dynamics of such states, long-time integrations of the Navier-Stokes equations are required. Numerical errors which are seemingly innocuous over a single time-step or in a steady-state result, nevertheless can accumulate to dangerous levels over long times. Hence, highly accurate calculations are imperative. The aperiodic time history and the resultant broadband frequency spectrum shown in the figure reflect the high accuracy achieved by the spectral simulation.

One of the objectives of these notes is to provide a basic introduction to spectral methods with a particular emphasis on applications to computational fluid dynamics. Another objective is to summarize some of the most important developments in spectral methods in the last two years. The fundamentals of spectral methods for simple problems will be covered in depth, and the essential elements of several fluid dynamical applications will be sketched. The recent book by Canuto, Hussaini, Quarteroni, and Zang (1988) (referred to hereafter as CHQZ) contains a comprehensive discussion of both the implementation and the rigorous numerical analysis of spectral methods. The first part of these notes will rely heavily upon that text, with frequent references to it for details, and will also include some material from our earlier reviews (Hussaini,

Streett, and Zang 1983, Hussaini and Zang 1987). Moreover, the notation in these notes conforms closely to that used in the text. Since CHQZ contains an exhaustive list of references through 1986, we shall not bother here to cite the original sources for most of the methods that we discuss. Instead, the bibliography for these notes focuses on those references which have appeared since 1986 and which pertain to new developments in spectral methods rather than just additional applications of established algorithms.

Additional general references on spectral methods are the monographs by Gottlieb and Orszag (1977) and Mercier (1981), the proceedings of the ICASE Spectral Methods Workshop edited by Voigt, Gottlieb, and Hussaini (1984), and doubtless also the forthcoming proceedings of the International Conference on Spectral and High-Order Methods for Partial Differential Equations to be held in Como, Italy in June, 1989.

2 Spectral Approximations

This chapter provides a basic introduction to Fourier and Chebyshev spectral approximations. Fourier approximations to infinitely differentiable, periodic functions exhibit exponential convergence. They do not behave nearly so well for non-periodic functions, regardless of their differentiability. For such functions expansions in Jacobi polynomials are appropriate. Chebyshev polynomials are the most useful polynomials in this class. (Legendre polynomials also belong to this class, but they suffer from the lack of a fast transform.)

The focus of this chapter is on practical details. It closely parallels the discussion in CHQZ, Ch. 2. The extensive literature on the rigorous numerical analysis of spectral approximations is surveyed in CHQZ, Ch. 9.

2.1 Fourier Expansions

The continuous Fourier expansion of a function, $u(x)$, on $(0, 2\pi)$ is

$$u(x) = \sum_{k=-\infty}^{\infty} \hat{u}_k e^{ikx}, \tag{11}$$

where the Fourier coefficients,

$$\hat{u}_k = \frac{1}{2\pi} \int_0^{2\pi} u(x) e^{-ikx} dx. \tag{12}$$

The orthogonality condition is

$$\frac{1}{2\pi} \int_0^{2\pi} e^{ipx} dx = \delta_{p0}, \tag{13}$$

where δ_{kl} denotes the Kronecker delta-function.

Let us begin by analyzing the convergence of this series for a function which is infinitely differentiable and which, together with all its derivatives, has period 2π. Using integration-by-parts we obtain

$$\hat{u}_k = -\frac{1}{2\pi} \frac{1}{ik} u(x) e^{-ikx} \Big|_0^{2\pi} + \frac{1}{2\pi} \frac{1}{ik} \int_0^{2\pi} u'(x) e^{-ikx} dx. \tag{14}$$

Invoking periodicity,

$$\hat{u}_k = \frac{1}{2\pi} \frac{1}{ik} \int_0^{2\pi} u'(x) e^{-ikx} dx. \tag{15}$$

The Riemann-Lebesgue lemma implies that

$$\hat{u}_k = O(k^{-2}) \qquad as\ k \to \infty. \tag{16}$$

Repeated integrations-by-parts leads to

$$\hat{u}_k = \frac{1}{2\pi}\frac{1}{(ik)^p}\int_0^{2\pi} u^{(p)}(x)e^{-ikx}dx; \tag{17}$$

hence to the conclusion that, for all positive integers p,

$$\hat{u}_k = O(k^{-p-1}) \qquad as\ k \to \infty. \tag{18}$$

Thus,

$$k^p\hat{u}_k \to 0 \qquad as\ k \to \infty \tag{19}$$

for all positive integers p.

The truncated Fourier series is defined to be

$$P_N u(x) = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \hat{u}_k e^{ikx}. \tag{20}$$

The property (19) implies that the truncated Fourier series converges faster than any finite power of $1/N$. In theoretical discussions of Fourier series, the truncation is customarily defined by

$$P_N u(x) = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}} \hat{u}_k e^{ikx}, \tag{21}$$

rather than by (20). However, because of the properties of the Fast Fourier Transform (see below), the truncation (20) is the one that is invariably used in practice.

For a less well-behaved function the repeated integrations-by-parts must stop at some point, either because $u(x)$ has only finitely many derivatives or because for the q^{th} derivative $u^{(q)}(2\pi^-) \neq u^{(q)}(0^+)$. In such a case the Fourier coefficients, and hence also the truncated Fourier series, decay only algebraically.

The discrete Fourier expansion is closely related. It is inherently a finite series, and is typically based on values of $u(x)$, denoted by u_j, which are given at the collocation points,

$$x_j = 2\pi j/N, \qquad j = 0, 1, \ldots, N-1. \tag{22}$$

The discrete Fourier coefficients are

$$\tilde{u}_k = \frac{1}{N}\sum_{j=0}^{N-1} u_j e^{-ikx_j}, \qquad k = -\frac{N}{2}, -\frac{N}{2}+1, \ldots \frac{N}{2}-1. \tag{23}$$

The inversion formula is

$$u_j = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \tilde{u}_k e^{ikx_j}, \tag{24}$$

and the orthogonality condition is

$$\frac{1}{N}\sum_{j=0}^{N-1} e^{ipx_j} = \begin{cases} 1 & p = Nm,\ m = 0, \pm 1, \pm 2, \ldots \\ 0 & \text{otherwise} \end{cases}. \tag{25}$$

Associated with the discrete Fourier series is the interpolating trigonometric polynomial,

$$I_N u(x) = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \tilde{u}_k e^{ikx} \tag{26}$$

which is defined for all $x \in [0, 2\pi]$. Note that

$$u_j \equiv u(x_j) = I_N u(x_j). \tag{27}$$

It is customary to refer to $u(x)$ and u_j as the physical space representations of u, and to $\hat{u}_k$ and $\tilde{u}_k$ as its transform space representations.

For a discrete Fourier series, both the function itself and its Fourier coefficients are periodic:

$$u_{j+N} = u_j \tag{28}$$

$$\tilde{u}_{k+N} = \tilde{u}_k. \tag{29}$$

For a continuous Fourier series, of course, only the function itself is periodic.

The sums in (23) and (24) can be accomplished by means of the Fast Fourier Transform (FFT). The simplest FFT requires N to be a power of 2. If the data are fully complex it requires $5N\ log_2 N\ - 6N$ real operations, where addition and multiplication are counted as separate operations. In most applications, u is real, and therefore $\tilde{u}_k = \tilde{u}^*_{-k}$. In this case the operation count is halved. Fast Fourier Transforms which allow factors of 2, 3, 4, 5, and 6 are widely available (Temperton 1983), and they offer a 10-20% reduction in the operation count over the basic power of 2 FFT. For simplicity, we shall often use just $5N\ log_2 N$ as the operation count for a complex FFT. Thus, to compute the discrete Fourier coefficients of u_j requires $\frac{5}{2} N\ log_2 N$ operations; the cost for computing the grid-point values, u_j, given the discrete Fourier coefficients is the same. A more complete discussion of FFT's (including several FORTRAN programs) is available in CHQZ, Appendix B.

The FFT's that are typically used in spectral methods require N to be even. This accounts for the use of (20) rather than (21) for the truncated Fourier series. The $\frac{N}{2}$ mode is something of a nuisance (see also the remark after (37) below), and it is probably advisable in practice simply to leave it out of the finite Fourier representation altogether.

The discrete Fourier coefficients are related to the continuous ones by

$$\tilde{u}_k = \hat{u}_k + \sum_{\substack{m=-\infty \\ m \neq 0}}^{\infty} \hat{u}_{k+Nm} \qquad k = -\frac{N}{2}, -\frac{N}{2}+1, \ldots \frac{N}{2}-1. \tag{30}$$

This expression shows that the k^{th} mode (or Fourier coefficient) of the trigonometric interpolant of u depends not only on the k^{th} mode of u, but also on all the modes of u which "alias" the k^{th} one on the discrete grid. The $(k+Nm)^{th}$ frequency aliases the k^{th} frequency on a grid consisting of N points; they are indistinguisable at the nodes since $e^{i(k+Nm)x_j} = e^{ikx_j}$. This phenomenon is illustrated for $N = 8$ in Figure 3. Shown there are three sine waves with frequencies k = 6, -2, and -10. Superimposed upon each wave are the eight grid-point values of the function. In each case these grid-point values coincide with the k = -2 wave.

An equivalent formulation of (30) is

$$I_N u = P_N u + R_N u, \tag{31}$$

with

$$R_N u = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \Big(\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \hat{u}_{k+Nm} \Big) e^{ikx}. \tag{32}$$

The error, $R_N u$, between the interpolating polynomial and the truncated Fourier series is called the "aliasing error". It is orthogonal to the truncation error, $u - P_N u$, so that

$$|| u - I_N u ||^2 = || u - P_N u ||^2 + || R_N u ||^2 . \tag{33}$$

Hence, the error due to interpolation is at least as great as the error due to truncation.

In a Fourier spectral method the derivative of a function, $u(x)$, is approximated by the analytic derivative of the finite series which approximates $u(x)$. In the case of the truncated Fourier series,

$$\frac{d(P_N u)}{dx} = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} ik\hat{u}_k e^{ikx}. \tag{34}$$

This is called the Fourier Galerkin derivative. Note that

$$\frac{d(P_N u)}{dx} = P_N(\frac{du}{dx}), \tag{35}$$

i.e., truncation and differentiation commute.

Likewise, for the interpolating Fourier series,

$$\frac{d(I_N u)}{dx} = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} ik\tilde{u}_k e^{ikx}. \tag{36}$$

This is called the Fourier collocation derivative. In this case, however,

$$\frac{d(I_N u)}{dx} \neq I_N(\frac{du}{dx}), \tag{37}$$

i.e., truncation and differentiation do not commute. Incidentally, the $\frac{N}{2}$ term in the sums (34) and (36) make a purely imaginary contribution to the derivative and are best left out entirely.

In a typical spectral Galerkin method, the fundamental representation is in transform space. Differentiation is accomplished by a simple multiplication. For the spectral collocation method, on the other hand, the fundamental representation is in physical space. In this case, Fourier differentiation using a transform method (employing two FFT's) is accomplished in three steps: (1) compute the discrete Fourier coefficients via (23); (2) compute the derivative coefficients $\tilde{u}_k^{(1)} = ik\tilde{u}_k$; (3) perform the sum (24) with $\tilde{u}_k^{(1)}$ in place of $\tilde{u}_k$. Steps (1) and (3) require $\frac{5}{2}N\ log_2 N - 3N$ operations each, and step (2) costs N multiplications. Thus, the total cost of the derivative is $5N\ log_2 N - 5N$ operations.

Fourier collocation differentiation can be represented by a matrix:

$$\frac{d(I_N u)}{dx}(x_k) = \sum_{l=0}^{N-1} D_{kl}^{(1)} u_l, \tag{38}$$

where

$$D_{kl}^{(1)} = \begin{cases} \frac{1}{2}(-1)^{k+l} cot[\frac{(k-l)\pi}{N}] & l \neq k \\ 0 & l = k \end{cases} . \tag{39}$$

For the second derivative we have

$$\frac{d^2(I_N u)}{dx^2}(x_k) = \sum_{l=0}^{N-1} D_{kl}^{(2)} u_l, \tag{40}$$

where

$$D_{kl}^{(2)} = \begin{cases} \frac{1}{2}(-1)^{k+l+1} csc^2[\frac{(k-l)\pi}{N}] & l \neq k \\ -(N^2-2)/12 & l = k \end{cases} . \tag{41}$$

Fourier collocation derivatives obviously may be calculated by simply performing the matrix-multiplies implied by (38) or (40). The operation count for this approach is $2N^2$. For large N, this is much larger than the operation count for the transform method. Moreover, the matrix-multiply approach is more susceptible to round-off errors. For small or moderate values of N, however, the matrix-multiply approach is competitive. Some timings are available in CHQZ, Sec. 2.1.3.

2.2 Chebyshev Expansions

Spectral methods for non-periodic problems are typically based upon Chebyshev polynomials. These are defined on [-1,1] by

$$T_k(x) = cos(k\, cos^{-1} x) \qquad k \geq 0. \tag{42}$$

The first several are $T_0(x) = 1, T_1(x) = x$, and $T_2(x) = 2x^2 - 1$. The continuous Chebyshev expansion of a function $u(x)$ on [-1,1] is

$$u(x) = \sum_{k=0}^{\infty} \hat{u}_k T_k(x), \tag{43}$$

where the Chebyshev coefficients,

$$\hat{u}_k = \frac{2}{\pi c_k} \int_{-1}^{1} u(x) T_k(x) w(x) dx, \tag{44}$$

with the Chebyshev weight,

$$w(x) = (1-x^2)^{-\frac{1}{2}}, \tag{45}$$

and

$$c_k = \begin{cases} 2 & k = 0 \\ 1 & otherwise \end{cases} . \tag{46}$$

The orthogonality relation is

$$\frac{2}{\pi c_k} \int_{-1}^{1} T_k(x) T_l(x)\, w(x) dx = \delta_{kl}. \tag{47}$$

The change of variables,

$$x = cos\theta, \tag{48}$$

the definition,

$$v(\theta) = u(cos\theta), \tag{49}$$

and (42) reduce (43) to

$$v(\theta) = \sum_{k=0}^{\infty} \hat{u}_k cosk\theta. \tag{50}$$

Thus, the Chebyshev coefficients of $u(x)$ are precisely the Fourier coefficients of $v(\theta)$. This new function is automatically periodic. If $u(x)$ is infinitely differentiable, then $v(\theta)$ will be infinitely differentiable. Hence, straightforward integration-by-parts arguments lead to the conclusion that the Chebyshev coefficients of an infinitely differentiable function will decay faster than algebraically. Note that this holds regardless of the boundary conditions.

Some important properties of Chebyshev polynomials are the boundary values,

$$T_k(\pm 1) = (\pm 1)^k, \tag{51}$$
$$T_k'(\pm 1) = (\pm 1)^k k^2, \tag{52}$$

and the recursion relations,

$$2xT_k(x) = T_{k+1}(x) + T_{k-1}(x), \qquad k \geq 1, \tag{53}$$
$$2T_k(x) = \frac{1}{k+1}T_{k+1}'(x) - \frac{1}{k-1}T_{k-1}'(x), \qquad k \geq 1. \tag{54}$$

The discrete Chebyshev expansion is based on the Gauss-Lobatto collocation points,

$$x_j = cos\frac{\pi j}{N}, \qquad j = 0, 1, \ldots, N. \tag{55}$$

The discrete Chebyshev coefficients are

$$\tilde{u}_k = \frac{2}{N\bar{c}_k}\sum_{j=0}^{N}\frac{1}{\bar{c}_j}u_j cos\frac{k\pi j}{N}, \qquad k = 0, 1, \ldots, N, \tag{56}$$

where

$$\bar{c}_k = \begin{cases} 2 & k = 0 \text{ or } N \\ 1 & 1 \leq k \leq N-1 \end{cases}. \tag{57}$$

The inversion formula is

$$u_j = \sum_{k=0}^{N}\tilde{u}_k cos\frac{\pi jk}{N}. \tag{58}$$

The Gauss collocation points are also sometimes used:

$$x_j = cos\frac{\pi(j-\frac{1}{2})}{N}, \qquad j = 1, 2, \ldots, N. \tag{59}$$

Special versions of the FFT may be used for evaluating the sums in (56) and (58). See CHQZ, Appendix B for a discussion of how to adapt the standard complex FFT for the Chebyshev transforms based on the Gauss-Lobatto and Gauss points. Schumann and Sweet (1988) discuss how a wide variety of sine and cosine transforms may be implemented with the standard FFT; this includes transforms which take advantage of special symmetries.

The truncated Chebyshev series is defined to be

$$P_N u(x) = \sum_{k=0}^{N}\hat{u}_k T_k(x), \tag{60}$$

and the Chebyshev interpolating polynomial is

$$I_N u(x) = \sum_{k=0}^{N}\tilde{u}_k T_k(x). \tag{61}$$

Differentiation in Chebyshev spectral methods is based upon the relation,

$$\frac{du}{dx} = \sum_{k=0}^{\infty} \hat{u}_k^{(1)} T_k(x), \tag{62}$$

where the Chebyshev coefficients, $\hat{u}_k^{(1)}$, of the derivative are related to those of the function itself by the recursion relation,

$$2k\hat{u}_k = \bar{c}_{k-1}\hat{u}_{k-1}^{(1)} - \hat{u}_{k+1}^{(1)}, \qquad k \geq 1. \tag{63}$$

(See CHQZ, Sec. 2.4.2 for the derivation of this formula.) This relation obviously generalizes to

$$2k\hat{u}_k^{(q-1)} = \bar{c}_{k-1}\hat{u}_{k-1}^{(q)} - \hat{u}_{k+1}^{(q)}, \qquad k \geq 1. \tag{64}$$

Explicit expressions for the first and second derivative coefficients are

$$\hat{u}_k^{(1)} = \frac{2}{c_k} \sum_{\substack{p=k+2 \\ p+k \text{ odd}}}^{\infty} p\hat{u}_p, \tag{65}$$

and

$$\hat{u}_k^{(2)} = \frac{1}{c_k} \sum_{\substack{p=k+2 \\ p+k \text{ even}}}^{\infty} p(p^2-k^2)\hat{u}_p, \tag{66}$$

For the truncated Chebyshev series (60), the derivative coefficients, $\hat{u}_k^{(1)}$, are zero for $k \geq N$, and the derivative coefficients for $0 \leq k \leq N-1$ may be computed in decreasing order by

$$\bar{c}_k\hat{u}_k^{(1)} = \hat{u}_{k+2}^{(1)} + 2(k+1)\hat{u}_{k+1}, \qquad k = N-1, N-2, \ldots, 0. \tag{67}$$

The Chebyshev Galerkin derivative is just $d(P_N u)/dx$. Unlike the Fourier case, truncation and differentiation do not commute: in general, $d(P_N u)/dx$ is a polynomial of degree $N-1$, whereas $P_N(du/dx)$ is a polynomial of degree N. A similar formula holds for the coefficients, $\tilde{u}_k^{(1)}$, for the derivatives of the Chebyshev interpolant. Higher-order derivative coefficients are computed by iterating on this recursion relation.

Chebyshev collocation differentiation can also be represented by matrices. For the first derivative,

$$\frac{d(I_N u)}{dx}(x_k) = \sum_{l=0}^{N} D_{kl}^{(1)} u_l, \tag{68}$$

where

$$D_{kl}^{(1)} = \begin{cases} (\bar{c}_k/\bar{c}_l)\,(-1)^{k+l}/(x_k - x_l) & 0 \leq k,l \leq N, l \neq k \\ -x_l/[2(1-x_l^2)] & 1 \leq k = l \leq N-1 \\ (2N^2+1)/6 & k = l = 1 \\ -(2N^2+1)/6 & k = l = N \end{cases}; \tag{69}$$

and for the second derivative (Ehrenstein and Peyret 1987),

$$\frac{d^2(I_N u)}{dx^2}(x_k) = \sum_{l=0}^{N} D_{kl}^{(2)} u_l, \tag{70}$$

where

$$D_{kl}^{(2)} = \begin{cases} [(-1)^{k+l}/\bar{c}_k]\,[x_k^2 + x_k x_l - 2]/[(1-x_k^2)(x_k - x_l)^2] & 1 \le k \le N-1,\ o \le l \le N;\ l \ne k \\ -[(N^2-1)(1-x_k^2)+3]/[3(1-x_k^2)^2] & 1 \le k = l \le N-1 \\ (2/3)[(-1)^l/\bar{c}_l][(2N^2+1)(1-x_l)-6]/(1-x_l)^2 & k = 0,\ 1 \le l \le N \\ (2/3)[(-1)^{N+l}/\bar{c}_l][(2N^2+1)(1+x_l)-6]/(1+x_l)^2 & k = N,\ 0 \le l \le N-1 \\ (N^4-1)/15 & k = l = 0, N \end{cases} \tag{71}$$

General expressions for the first- and second-derivative matrices for Jacobi polynomials are available in Deville and Mund (1989). For an explicit expression for the inverse of the Chebyshev first-derivative matrix (with homogeneous Dirichlet boundary conditions on one end) see Funaro (1988a).

2.3 Galerkin, Tau, and Collocation Approximations

On some domain Ω with boundary $\partial\Omega$, consider the differential equation,

$$\begin{aligned} L(u) &= f, && in\ \Omega, \\ B(u) &= 0, && on\ \partial\Omega_b, \end{aligned} \tag{72}$$

where L is a differential operator, B is a linear boundary operator, u is the solution, f is the data, and $\partial\Omega_b$ is that subset of $\partial\Omega$ on which boundary conditions are enforced. The boundary conditions are assumed here for simplicity to be homogeneous.

In order to use spectral methods to solve this differential problem, a set of discrete equations must be employed. Three strategies for deriving a discrete system are in common use: Galerkin, tau, and collocation. All of these may be viewed as particular cases of spectral weighted residual methods. The key elements are the set of trial functions, X_N, the set of test functions, Y_N, and discrete approximations, L_N and B_N, to the operators, L and B. Another important element is an inner product on Ω, denoted by (u, v). CHQZ, Secs. 10.3 and 10.4 have described a general formulation of Galerkin, tau, and collocation methods. Here we shall be content with presenting a straightforward prescription for these methods.

The starting point for choosing the trial and test functions is a set, $\mathcal{P}_N$, of algebraic polynomials on Ω of degree less than or equal to N. Let $\{p_k\}_{k=0}^N$ comprise an orthogonal basis for $\mathcal{P}_N$, and assume that these polynomials are ordered by increasing degree. For notational simplicity, we describe the procedure for approximations to non-periodic problems, in which case Jacobi polynomials are appropriate. Both the trial and test functions are contained in subsets of $\mathcal{P}_N$. The trial functions necessarily satisfy the boundary conditions:

$$X_N = \{v \in \mathcal{P}_N \mid Bv = 0\ on\ \partial\Omega_b\}. \tag{73}$$

The approximate solution is denoted by u^N, and the residual is just

$$R(u^N) = L(u^N) - f. \tag{74}$$

The weighted residual orthogonality statement is equivalent to

$$\begin{aligned} &u^N \in X_N, \\ &(L(u^N) - f, v) = 0, \qquad for\ all\ v \in Y_N. \end{aligned} \tag{75}$$

The Galerkin weighted residual approximation results from setting $Y_N = X_N$, and using the continuous inner product,

$$(u, v) = \int_\Omega uv\, w\, d\Omega, \tag{76}$$

where w is a weight function, in (75). Let $\{\phi_k\}_{k \in J}$ be a basis for X_N (and in the case of Galerkin methods, also for Y_N). The ϕ_k are linear combinations of the polynomials p_k such that the boundary conditions are satisfied. The approximate solution can be expressed as

$$u^N = \sum_{k \in J} \hat{u}_k \phi_k. \tag{77}$$

Eq. (75) equivalent to the requirement that

$$\begin{aligned} &u^N \in X_N, \\ &(L(u^N), \phi_k) = (f, \phi_k), \qquad\qquad k \in J. \end{aligned} \tag{78}$$

We have assumed that $L_N = L$, as occurs for most applications. In the Galerkin approximation the unknowns are the expansion coefficients, $\hat{u}_k$, of the basis functions for X_N.

The tau approximation is very similar, but the test functions are not required individually to satisfy the boundary conditions. The set, X_N, remains as given by (73), but $Y_N = \mathcal{P}_{N-M}$, where M is the number of boundary conditions. Although the solution still has the form (77) of an expansion in terms of the trial functions, it is generally more convenient to work with it in the form,

$$\begin{aligned} &u^N = \sum_{k=0}^{N} \hat{u}_k p_k, \\ &B(u^N) = 0, \qquad\qquad \text{on } \partial\Omega_b, \end{aligned} \tag{79}$$

of an expansion in terms of the orthogonal basis for $\mathcal{P}_N$. The condition (75) is equivalent to (79) plus

$$(L(u^N), p_k) = (f, p_k), \qquad k = 0, 1, \ldots, N - M. \tag{80}$$

Observe that the residual is not required to be orthogonal to the M highest-order expansion functions. These conditions are dropped in favor of enforcing the boundary conditions on u^N. Note that the unknowns in the expansion are the coefficients of the basis functions for $\mathcal{P}_N$ and not of the basis functions for X_N. In the tau method there is no need to construct explicitly the combinations of polynomials which satisfy the boundary conditions.

Collocation methods are at first sight quite different. A set of collocations points $x_j, j = 1, 2, \ldots, N$, is chosen, and the approximate solution is required to satisfy (72) exactly at the collocation points:

$$\begin{aligned} B_N(u^N)(x_j) &= 0, \qquad &&\textit{at all boundary points}, \\ L_N(u^N)(x_j) &= f(x_j), \qquad &&\textit{at all internal points}, \end{aligned} \tag{81}$$

where $B_N(u)$ and $L_N(u)$ are the discrete approximations to $B(u)$ and $L(u)$ which result from taking derivatives via collocation. In this approximation the unknowns are the nodal values of the function.

Nevertheless, the collocation method can be placed within the framework of weighted residual methods. Loosely speaking, one keeps the condition (79) on the trial functions and chooses the

Dirac delta-functions, $\delta(x - x_j)$, as the test functions. A more precise analogy is based on discrete inner product,

$$(u, v)_N = \sum_{j=0}^{N} u_j v_j \, w_j, \tag{82}$$

which results from applying the appropriate Gaussian quadrature formula, rather than the continuous inner product (76). (The quadrature rule, in fact, determines the collocation points, x_j, and the weights, w_j.) As shown in CHQZ, Sec. 10.4.3, the collocation method can then be placed in the same variational framework as the Galerkin method.

In these notes we shall customarily use $\hat{u}_k$ for the expansion coefficients of Galerkin and tau methods (in which they are the fundamental variables), but $\tilde{u}_k$ for the expansion coefficients of collocations methods (in which they are intermediate variables used for the approximation of derivatives).

2.4 Generalizations

In many cases a mapping of the desired interval, (a, b), onto the standard interval – $(0, 2\pi)$ for periodic boundary conditions and $(-1, 1)$ otherwise – is needed. A simple linear mapping is straightforward. However, often a nonlinear mapping can improve the quality of the approximation. Some types of mappings are discussed in CHQZ, Sec. 2.5. Of particular interest is the recent work of Boyd. See the references in the bibliography of CHQZ and in the present bibliography.

An important unresolved issue is how severe a stretching is permitted. The current folklore is that the truncated series must resolve not only the function itself, but also the mapping, in order to achieve a good approximation.

In more than one dimension the standard choice for the basis functions is just the tensor product of one-dimensional basis functions. (See, however, Bisseling and Kosloff (1988) for an application using a Fourier method on an isotropic grid.) This, of course, limits conventional spectral methods to extremely simple geometries. This limitation can be surmounted by spectral domain decomposition methods, which are discussed in Chapter 4.

Some subtle issues arise in the imposition of Neumann or Robin boundary conditions and the treatment of coordinate singularities. See CHQZ, Ch. 3. In some circumstances there are advantages to enforcing a combination of the differential equation and the boundary conditions at the boundary. This applies to both elliptic (Funaro 1988b) and to hyperbolic (Funaro and Gottlieb 1988) problems.

3 Elementary Applications to Model Problems

In this chapter we describe the spatial discretizations that result from spectral approximations to simple PDE's. For evolution problems the time-discretizations used in spectral methods are quite conventional; see CHQZ, Ch. 4 for a complete discussion.

3.1 Some Fourier Methods For the Wave Equation

An elementary example is provided by the model problem,

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \tag{83}$$

with periodic boundary conditions on $(0, 2\pi)$, and the initial condition,

$$u(x,0) = sin(\pi cos x). \tag{84}$$

The exact solution,

$$u(x,t) = sin[\pi cos(x-t)], \tag{85}$$

has the Fourier expansion,

$$u(x,t) = \sum_{k=-\infty}^{\infty} \hat{u}_k(t) e^{ikx}, \tag{86}$$

where the Fourier coefficients,

$$\hat{u}_k(t) = sin(\frac{k\pi}{2}) J_k(\pi) e^{-ikt}, \tag{87}$$

and $J_k(t)$ is the Bessel function of order k. The asymptotic properties of the Bessel functions imply that

$$k^p \hat{u}_k(t) \to 0 \qquad as\ k \to \infty \tag{88}$$

for all positive integers p, consistent with the result (19). As a result, the truncated Fourier series,

$$u^N(x,t) = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \hat{u}_k(t) e^{ikx}, \tag{89}$$

converges faster than any finite power of 1/N.

The prescription (78) readily extends to the semi-discrete (discrete in space, continuous in time) Galerkin approximation to (83) and produces the classical result,

$$\frac{d\hat{u}_k}{dt} + ik\hat{u}_k = 0, \qquad k = -\frac{N}{2}, \ldots, \frac{N}{2} - 1. \tag{90}$$

Let $u_j(t)$ denote the approximation to $u(x_j, t)$, and recall (27). Then the semi-discrete collocation approximation (81) to (83) is

$$\frac{\partial u_j}{\partial t}(t) + \frac{\partial (I_N u)}{\partial x}(x_j, t) = 0, \qquad j = 0, 1, \ldots, N-1. \tag{91}$$

N	Fourier Galerkin	Fourier Collocation	Second-order	Fourth-order
8	9.87 (-2)	1.62 (-1)	1.11 (0)	9.62 (-1)
16	2.55 (-4)	4.97 (-4)	6.13 (-1)	2.36 (-1)
32	1.05 (-11)	1.03 (-11)	1.99 (-1)	2.67 (-2)
64	6.22 (-13)	9.55 (-12)	5.42 (-2)	1.85 (-3)

Table 1: Maximum Error for the Periodic Wave Equation

An illustration of the superior accuracy available from the spectral method for this problem is provided in Table 1 and Figure 4. Shown in the table are the maximum errors at $t = 1$ for the Fourier Galerkin and the Fourier collocation methods as well as for second-order and fourth-order finite-difference methods. The time discretization was the classical fourth-order Runge-Kutta

method. In all cases the time-step was chosen so small that the temporal discretization error was negligible. The figure compares the second-order finite-difference and Fourier Galerkin solutions for $N = 32$ with the exact answer. Note that the major error in the finite-difference solution is one of phase rather than amplitude. In many problems the very low phase errors of spectral methods is a significant advantage.

Because the solution is infinitely differentiable, the convergence of the spectral method on this problem is more rapid than any finite power of $1/N$. (The error for the $N = 64$ spectral result is so small that it is swamped by the round-off error of these 60-bit calculations.) In most practical applications the benefit of the spectral method is not the extraordinary accuracy available for large N, but rather the small size of N necessary for a moderately accurate solution.

3.2 Some Chebyshev Methods for the Wave Equation

The PDE is again (83), but the initial condition is now

$$u(x,0) = sin(\alpha\pi x), \tag{92}$$

and the boundary condition is

$$u(-1,t) = u_L(t) = sin[\alpha\pi(1-t)]. \tag{93}$$

The function,

$$u(x,t) = sin[\alpha\pi(x-t)], \tag{94}$$

is one solution to the problem (83), (92), and (93). It has the Chebyshev expansion,

$$u(x,t) = \sum_{k=0}^{\infty} \hat{u}_k(t) T_k(x), \tag{95}$$

where

$$\hat{u}_k(t) = \frac{2}{c_k} sin(\frac{k\pi}{2} - \alpha\pi t) J_k(\alpha\pi). \tag{96}$$

Clearly, the truncated series, $P_N u$, converges at an exponential rate. Note that this result holds whether or not α is an integer. In contrast, the Fourier coefficients of $u(x,t)$ are

$$\hat{u}_k(t) = \frac{i}{2\pi} e^{i\alpha\pi t} \frac{sin\pi(\alpha+k)}{\alpha+k} - \frac{i}{2\pi} e^{-i\alpha\pi t} \frac{sin\pi(\alpha-k)}{\alpha-k}. \tag{97}$$

For non-integer α these decay extremely slowly – only as $O(k^{-1})$ – and the Fourier series exhibits a Gibbs phenomenon (see CHQZ, Sec. 2.1.4).

For the present problem the semi-discrete Chebyshev tau formulation (79) - (80) is

$$\frac{d\hat{u}_k}{dt} + \hat{u}_k^{(1)} = 0, \qquad k = 0, 1, \ldots, N-1, \tag{98}$$

$$\sum_{k=0}^{N} (-1)^k \hat{u}_k = u_L(t), \tag{99}$$

along with the obvious initial condition on the Chebyshev coefficients. Consistent with (80), the highest-order condition ($k = N$) for the differential equation has been dropped in (98) in favor of the boundary condition (99). The property (51) has been employed in (99). Note that no special formula is required for the derivative at $j = 0$ (or $x = 1$). In a typical finite-difference

method a special one-sided formula would be utilized at the right boundary. The Chebyshev collocation approximation is

$$\frac{\partial u_j}{\partial t}(t) + \frac{\partial (I_N u)}{\partial x}(x_j, t) = 0, \qquad j = 0, 1, \ldots, N-1, \tag{100}$$

with

$$u_N(t) = u_L(t), \tag{101}$$

and the initial condition.

N	Truncated Series	Chebyshev Collocation	Fourier Collocation	Second-order
4	1.24 (0)	1.49 (0)	1.85 (0)	1.64 (0)
8	1.25 (-1)	6.92 (-1)	1.92 (0)	1.73 (0)
16	7.03 (-6)	1.50 (-4)	2.27 (0)	1.23 (0)
32	1.62 (-13)	3.45 (-11)	2.28 (0)	3.34 (-1)
64	1.79 (-13)	9.55 (-11)	2.27 (0)	8.44 (-2)

Table 2: Maximum Error for the Non-periodic Wave Equation

Results pertaining to $\alpha = 2.5$ at $t = 1$ for a truncated Chebyshev series, a Chebyshev collocation method, a Fourier collocation method, and a second-order finite-difference method are given in Table 2. For this non-periodic problem Fourier, spectral methods are quite inappropriate (and the failure to converge in the maximum norm reflects the Gibbs phenomenon), but the Chebyshev spectral method is far superior to the finite-difference method.

The Chebyshev collocation points are the extreme points of $T_N(x)$. Note that they are not evenly distributed in x, but rather are clustered near the endpoints. The smallest mesh size scales as $1/N^2$, and the magnitude of the largest eigenvalue of the Chebyshev first-derivative operator grows as N^2. While the Chebyshev Gauss-Lobatto distribution contributes to the quality of the Chebyshev approximation and permits the use of the FFT in evaluating the series, it also places a severe time-step limitation on explicit methods for evolution equations.

3.3 Some Explicit Chebyshev Methods for the Heat Equation

The heat equation,

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \tag{102}$$

is the natural parabolic linear model problem. For our simple example the spatial domain is (-1,1), the initial condition is

$$u(x, 0) = sin\pi x, \tag{103}$$

and the boundary conditions are

$$u(-1, t) = u(+1, t) = 0. \tag{104}$$

The exact solution is then

$$u(x, 0) = e^{-\pi^2 t} sin\pi x. \tag{105}$$

The semi-discrete tau equations are

$$\frac{d\hat{u}_k}{dt} = \hat{u}_k^{(2)}, \qquad k = 0, 1, \ldots, N-2, \tag{106}$$

with

$$\sum_{k=0(mod\ 2)}^{N} \hat{u}_k \quad = \sum_{k=1(mod\ 2)}^{N} \hat{u}_k \quad = 0. \tag{107}$$

(The notation $k = 0\ (mod\ 2)$ means that only the contributions from the even values of k are included in the sum.) The Chebyshev coefficients, $\hat{u}_k^{(2)}$, of the approximation to the second spatial derivative can be obtained from $\hat{u}_k$ by two applications of the recursion relation (63). In this tau approximation the dynamical equations for the two highest-order coefficients are dropped in favor of the equations for the boundary conditions. Eq. (107) follows from (104) and the property (51). The Chebyshev collocation approximation is

$$\frac{\partial u_j}{\partial t}(t) = \frac{\partial^2 (I_N u)}{\partial x^2}(x_j, t), \qquad j = 1, 2, \ldots, N, \tag{108}$$

with

$$u_0(t) = u_N(t) = 0. \tag{109}$$

For both the tau and collocation approximations above, implementation of explicit time-discretization schemes is straightforward. The results at $t = 1$ are given in Table 3. The maximum errors shown there have been boosted up by the factor e^{π^2} so that they represent relative errors. On the whole the collocation results are the best. It goes almost without saying that finite-difference results are far inferior to any of these spectral approximations.

N	Truncated Series	Tau	Collocation
8	2.44 (-4)	1.61 (-3)	4.58 (-4)
10	5.76 (-6)	2.12 (-5)	8.25 (-6)
12	9.42 (-8)	3.19 (-7)	1.01 (-7)
14	1.14 (-9)	3.35 (-9)	1.10 (-9)
16	1.05 (-11)	8.39 (-11)	2.09 (-11)

Table 3: Maximum Error for Chebyshev Approximations to the Heat Equation

3.4 An Implicit Chebyshev Method for the Heat Equation

The time-step restriction for explicit Chebyshev methods for the heat equation is very severe, scaling as $1/N^4$. Implicit time-discretizations overcome this stability limit at the price of requiring the solution of implicit equations.

We write the Chebyshev tau approximation resulting from a Crank-Nicolson discretization of (106) - (107) as

$$-\frac{1}{2}\Delta t\, \hat{u}_k^{(2)}(t+\Delta t) + \hat{u}_k(t+\Delta t) = \frac{1}{2}\Delta t\, \hat{u}_k^{(2)}(t) + \hat{u}_k(t), \qquad k = 0, 1, \ldots, N-2, \tag{110}$$

with

$$\sum_{k=0(mod\ 2)}^{N} \hat{u}_k(t+\Delta t) = 0, \tag{111}$$

$$\sum_{k=1(mod\ 2)}^{N} \hat{u}_k(t+\Delta t) = 0. \tag{112}$$

We re-write (110) as

$$\hat{u}_k^{(2)} - \lambda \hat{u}_k = \hat{f}_k, \tag{113}$$

where $\lambda = 2/\Delta t$, the coefficients on the left-hand side are at $t + \Delta t$, and

$$\hat{f}_k = -\hat{u}_k^{(2)}(t) - \lambda \hat{u}_k(t). \tag{114}$$

Eqs. (111) - (113) form an implicit system for the coefficients $\hat{u}_k$ for $k = 0, 1, \ldots, N$, with (66) used to express $\hat{u}_k^{(2)}$ in terms of $\hat{u}_k$. The corresponding linear system is upper triangular. The solution process for such a system requires N^2 operations.

A far more efficient solution procedure is obtained by re-arranging the equations. By invoking the recursion relation (63) twice one obtains

$$\begin{aligned} &\frac{c_{k-2}\lambda}{4k(k-1)}\hat{u}_{k-2} + \Big(1 - \frac{e_k\lambda}{2(k^2-1)}\Big)\hat{u}_k + \frac{e_{k+2}\lambda}{4k(k+1)}\hat{u}_{k+2} = \\ &\frac{c_{k-2}}{4k(k-1)}\hat{f}_{k-2} - \frac{e_k}{2(k^2-1)}\hat{f}_k + \frac{e_{k+2}}{4k(k+1)}\hat{f}_{k+2}, \qquad k = 2, 3, \ldots, N, \end{aligned} \tag{115}$$

where

$$e_k = \begin{cases} 1 & 0 \le k \le N-2 \\ 0 & k \ge N-1 \end{cases}, \tag{116}$$

in place of (113). (See CHQZ, Sec. 5.1.2 for the complete derivation.) Note that the even and odd coefficients are uncoupled. By writing the equations in the order (111), and then (115) for $k = 2, 4, \ldots, N$ one obtains a linear system which is quasi-tridiagonal: the first row is full, and the remaining rows have a standard tridiagonal structure. A similar ordering is chosen for the odd coefficients. This ordering has been chosen to minimize the round-off errors arising from a specially tailored Gauss elimination procedure which performs no pivoting (and works from the "bottom up" rather than the more customary "top down"). Assuming that the coefficients in (115) have already been calculated, the solution of both systems takes only $16N$ operations, a substantial improvement over the $O(N^2)$ cost of the upper triangular system.

The coefficient of $\hat{u}_k$ in (115) is the largest coefficient and it is desirable for it to be on the main diagonal. The system is not diagonally dominant and, in practice, round-off errors are a mild problem: typically 4 digits are lost for $N = 128$. The accuracy may be increased through the use of iterative improvement or double-precision.

The $O(N)$ operation count for a single step of an implicit Chebyshev tau approximation to the one-dimensional heat equation is comparable to the cost of a finite-difference method. The transformation to physical space takes an extra $O(N \log_2 N)$ operations. Hence, the cost of this spectral method is larger than that of a finite-difference method by a factor $N/\log_2 N$. However, the spectral method requires far fewer grid-points to achieve a high accuracy.

The class of problems for which such efficient implicit Chebyshev tau methods apply is quite limited. Only the simplest variable coefficients are permitted. Tuckerman (1989b) has provided a characterization of such problems and given a general prescription for manipulating the implicit equations into a banded form.

An implicit collocation method does not reduce to a quasi-tridiagonal system. In this approach the cost of solving the implicit equations is $O(N^3)$.

3.5 Chebyshev Methods for the Poisson Equation

As usual the discussion will begin with a linear model problem, but this time in two spatial dimensions. That problem is the Poisson equation,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f, \tag{117}$$

on the square, $(-1,1)^2$, with homogeneous Dirichlet boundary conditions. The choice,

$$f(x,y) = -2\pi^2 sin\pi x\ sin\pi y, \tag{118}$$

corresponds to the analytical solution ,

$$u(x,y) = sin\pi x\ sin\pi y. \tag{119}$$

A Chebyshev expansion in both x and y is appropriate for this problem. Assume, for simplicity, that polynomials of degree less than or equal to N are used in both directions. The Chebyshev tau approximation is

$$\hat{u}^{(2,0)}_{mn} + \hat{u}^{(0,2)}_{mn} = \hat{f}_{mn}, \qquad m,n = 0,1,\ldots,N-2, \tag{120}$$

with

$$\sum_{m=0(mod\ 2)}^{N} \hat{u}_{mn} = \sum_{m=1(mod\ 2)}^{N} \hat{u}_{mn} = 0, \qquad n = 0,1,\ldots,N, \tag{121}$$

$$\sum_{n=0(mod\ 2)}^{N} \hat{u}_{mn} = \sum_{n=1(mod\ 2)}^{N} \hat{u}_{mn} = 0, \qquad m = 0,1,\ldots,N. \tag{122}$$

In this notation, $\hat{u}^{(2,0)}_{mn}$ and $\hat{u}^{(0,2)}_{mn}$ denote the Chebyshev coefficients of $\frac{\partial^2 u}{\partial x^2}$ and $\frac{\partial^2 u}{\partial y^2}$, respectively. The boundary conditions (121) - (122) contain 4 redundant conditions corresponding to the 4 corner points of the square. The collocation approximation reads

$$\frac{\partial^2 (I_N u)}{\partial x^2}(x_j,y_k) + \frac{\partial^2 (I_N u)}{\partial y^2}(x_j,y_k) = f(x_j,y_k), \qquad j,k = 1,2,\ldots,N-1, \tag{123}$$

$$u_{jk} = 0, \qquad j,k = 0\ or\ N, \tag{124}$$

where x_j and y_k denote the Chebyshev-Gauss-Lobatto points in x and y, respectively.

N	Truncated Series	Tau	Collocation
8	2.88 (-4)	2.79 (-3)	1.17 (-4)
10	6.79 (-6)	5.26 (-5)	2.33 (-6)
12	1.09 (-7)	8.86 (-7)	3.12 (-8)
14	1.34 (-9)	1.09 (-8)	3.27 (-10)
16	1.19 (-11)	9.15 (-11)	2.73 (-12)

Table 4: Maximum Error for Chebyshev Approximations to Poisson's Equation

N	Truncated Series	Tau	Collocation
8	2.88 (-4)	2.79 (-3)	1.17 (-4)
10	6.79 (-6)	5.26 (-5)	2.33 (-6)
12	1.09 (-7)	8.86 (-7)	3.12 (-8)
14	1.34 (-9)	1.09 (-8)	3.27 (-10)
16	1.19 (-11)	9.15 (-11)	2.73 (-12)

Table 4: Maximum Error for Chebyshev Approximations to Poisson's Equation

3.6 Fourier Methods for the Burgers Equation

Our last, and not so elementary, example is for the Burgers equation,

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \nu\frac{\partial^2 u}{\partial x^2}, \tag{125}$$

on $(0, 2\pi)$ with periodic boundary conditions. The test problem is chosen so that the exact solution is

$$u(x,t) = -2\nu\frac{\frac{\partial \phi}{\partial x}(x - ct, t+1)}{\phi(x-ct, t+1)}, \tag{126}$$

where $c = 1$, and

$$\phi(x,t) = \sum_{n=-\infty}^{\infty} e^{-[x-(2n+1)\pi]^2/4\nu t}. \tag{127}$$

Unlike the previous examples, this problem is nonlinear. The Fourier Galerkin approximation to the PDE is

$$\frac{\partial \hat{u}}{\partial t} + \sum_{\substack{p+q=k \\ p,q=-\frac{N}{2},\ldots,\frac{N}{2}-1}}^{\infty} \hat{u}_p iq\hat{u}_q = -k^2\nu\hat{u}_k, \qquad k = -\frac{N}{2},\ldots,\frac{N}{2}-1. \tag{128}$$

The convolution sum in this equation requires $O(N^2)$ operations if evaluated in a straightforward manner. Fortunately, transform methods permit it to be evaluated in only $O(N\ log_2 N)$ operations.

Let us consider a general one-dimensional convolution sum of the form,

$$\sum_{\substack{p+q=k \\ p,q=-\frac{N}{2},\ldots,\frac{N}{2}-1}}^{\infty} \hat{u}_p\hat{v}_q. \tag{129}$$

The basic approach is to transform $\hat{u}_k$ and $\hat{v}_k$ to physical space, to perform there a pointwise multiplication, and then to transform the result back to Fourier space. We introduce the discrete transforms,

$$U_j = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \hat{u}_k e^{ikx_j}, \qquad j = 0, 1, \ldots, N-1, \tag{130}$$

$$V_j = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \hat{v}_k e^{ikx_j}, \qquad j = 0, 1, \ldots, N-1, \tag{131}$$

and define the physical space product,

$$W_j = U_j V_j, \qquad j = 0, 1, \ldots, N-1, \tag{132}$$

and its discrete Fourier transform,

$$\hat{W}_k = \frac{1}{N} \sum_{j=0}^{N-1} W_j e^{-ikx_j}, \qquad k = -\frac{N}{2}, \ldots, \frac{N}{2} - 1. \tag{133}$$

Use of the discrete orthogonality relation (25) leads to

$$\hat{W}_k = \sum_{\substack{p+q=k \\ p,q=-\frac{N}{2},\ldots,\frac{N}{2}-1}} \hat{u}_p \hat{v}_q + \sum_{\substack{p+q=k\pm N \\ p,q=-\frac{N}{2},\ldots,\frac{N}{2}-1}} \hat{u}_p \hat{v}_q. \tag{134}$$

The first term on the right-hand side is the desired result. The second term is called the aliasing error. Hence, this naive transform method does not produce a true Galerkin method. For this reason it is referred to as a pseudospectral method.

The transform method, however, can be modified to produce a true de-aliased, Galerkin convolution sum. Two techniques are available. One is called the 3/2-rule and it involves performing the transforms (130) and (131) on a grid with $M \geq \frac{3}{2}N$ points. In this case one simply pads the original Fourier coefficients $\hat{u}_k$ and $\hat{v}_k$ with zeroes for $\frac{N}{2} < k \leq \frac{M}{2}$ before performing the transforms. The multiplication (132) at physical points is performed on M points, as is the transform (133). CHQZ, Sec. 3.2 demonstrate that the result of this method is a set of Fourier coefficients, $\hat{W}_k$, which agree with the desired coefficients, $\hat{w}_k$, for $k = -\frac{N}{2}, \ldots, \frac{N}{2} - 1$. This transform method costs roughly 50% more than the naive transform method discussed above, but it does produce a de-aliased result.

The second de-aliasing technique involves applying the naive transform method twice, once as described above and a second time using a grid in physical space which is shifted with respect to the standard one by π/N. The average of these two results is alias-free. (See CHQZ, Sec. 3.2 for the details.)

In multiple dimensions the 3/2-rule generalizes in an obvious manner. However, the phase shift de-aliasing technique requires evaluations on multiple grids, e.g., on 8 separate grids in three dimensions. Moreover, the phase shift method is the more expensive of the two. CHQZ, Sec. 7.2.2 furnish a detailed comparison of de-aliasing techniques for three-dimensional convolution sums.

A Fourier collocation method for the Burgers equation is much more straightforward. It reads

$$\frac{\partial u_j}{\partial t}(t) + u_j \frac{\partial (I_N u)}{\partial x}(x_j, t) = \nu \frac{\partial^2 (I_N u)}{\partial x^2}(x_j, t), \qquad j = 0, 1, \ldots, N-1. \tag{135}$$

One can show (CHQZ, Sec. 3.2.5) that the collocation method (135) is algebraically equivalent to the Fourier approximation (128) employing the pseudospectral transform method. Collocation methods do include aliasing errors. However, for problems with even more complicated nonlinearities than those of the Burgers equation, collocation methods are much simpler to implement than Galerkin methods.

The aliasing error can be removed from the collocation method for this problem. One simply forces the upper-third of the coefficients in transform space to be zero at every time step. (This can be achieved at nominal expense here if the viscous term is treated implicitly and a transform method is used for the solution.) This approach is called the 2/3-rule.

Figure 5 shows typical results for second-order finite-difference and for Fourier collocation methods, and Table 5 gives a comparison between spectral and finite-difference methods.

N	Fourier Collocation	Second-order	Fourth-order
16	2.1 (-1)	2.9 (-2)	2.0 (-1)
32	3.5 (-2)	8.2 (-2)	2.5 (-2)
64	3.6 (-4)	2.6 (-2)	5.2 (-4)
128	6.1 (-8)	3.3 (-3)	2.5 (-5)

Table 5: Maximum Error for Fourier Approximations to Burgers Equation

There has been a long controversy over the seriousness of the aliasing errors. Recent numerical and theoretical results establish that aliased and de-aliased spectral methods yield the same asymptotic rate of convergence. CHQZ, Secs. 4.6, 7.2.4, 7.3.5, and 11.3 summarize the results available as of a few years ago. Recently, Maday and Quarteroni (1989) have extended these results to the Korteweg-de Vries equation.

4 Domain Decomposition Methods

The previous chapters have concentrated on spectral methods for problems in simple domains. There have been a number of recent developments on the use of spectral techniques in more general geometries and even on their use in simple geometries in ways that improve resolution for problems with disparate scales. The basic idea is to partition the complete domain of the problem into several subdomains. The approximation is spectral if increased accuracy is obtained by increasing the order of approximation in a fixed number of subdomains, rather than by resorting to a further partitioning. One situation in which this domain decomposition approach is useful is illustrated in Figure 6. In this particular example it is clear that at least two subdomains are required in order to use spectral methods at all. In the case illustrated in Figure 6 the partitions intersect only at the boundaries. It is also possible to use subdomains which overlap. This is illustrated in Figure 7. Additional advantages may arise by using separate subdomains instead of a single subdomain which is their union. The partitioning illustrated in Figure 6 leads to a distribution of Chebyshev collocation points which improves the resolution of the approximation. Moreover, one would expect the corresponding algebraic problem to be better-conditioned because there is a less extreme ratio of the largest to smallest grid spacings. Finally, the use of subdomains facilitates the implementation of spectral methods on parallel computers, especially those with local memory.

A crucial aspect of any domain decomposition method is the manner in which solutions on contiguous domains are matched. Patching methods take a classical (pointwise) view of the differential equation. If the equation has order d, then at the interface of contiguous domains the solution and all its derivatives of order up to $d-1$ must be continuous. For second-order problems this is typically enforced by requiring that the solution, u, be continuous and that its normal derivatives, $\partial u/\partial n$, match at the interface, in the sense that they are equal and opposite.

The condition on the normal derivative may be replaced with a continuity condition on any other directional (but not tangential) derivative. These continuity conditions are discretized by enforcing them at selected points, and thus, are satisfied exactly by any approximation.

Alternatively, the differential equation may be posed variationally. The standard weak formulation is based on an integration-by-parts procedure over the entire physical domain. This implicitly assumes continuity of u and its first $d-1$ derivatives at all points, and, in particular, at the interfaces between subdomains. In most cases continuity of u is built into the choice of trial functions. Continuity of the derivatives of u occurs as a natural interface condition. By reversing the integration-by-parts procedure on the decomposed domain, one can deduce directly from the variational principle the continuity of the derivatives of u at the interfaces. In a variational domain decomposition method, the natural interface conditions are satisfied with increasing precision as the order of approximation is increased.

In an overlapping domain method only the continuity of the function is imposed. But this matching occurs on the boundary of the intersection between two adjacent domains. In the discrete case this is enforced at selected points on the boundary.

4.1 Patching Methods

A one-dimensional domain decomposition for the interval, $\Omega = (a,b)$, illustrated in Figure 8. It is broken into two subdomains, $\Omega_1 = (a_1, a_2)$ and $\Omega_2 = (a_2, a_3)$, where $a_1 = a$ and $a_3 = b$. The solution to a given differential equation on Ω is denoted by u. Its restriction to Ω_s is denoted by u_s, for $s = 1,2$. The integer, N_s, is used to indicate the degree of approximation on Ω_s. The approximate solution on Ω is denoted by u^N and its restriction to Ω_s by u_s^N.

Consider a patching method for the linear problem,

$$Lu \equiv -\nu \frac{d^2u}{dx^2} + \alpha \frac{du}{dx} + \lambda u = f, \qquad in\ \Omega. \tag{136}$$

The ODE problem may be formulated as

$$\begin{aligned} &Lu_1 = f, && in\ \Omega_1, \\ &u_1(a_1) = 0, \end{aligned} \tag{137}$$

$$\begin{aligned} &Lu_2 = f, && in\ \Omega_2, \\ &u_2(a_3) = 0, \end{aligned} \tag{138}$$

$$\begin{aligned} u_1(a_2) &= u_2(a_2), \\ \frac{du_1}{dx}(a_2) &= \frac{du_2}{dx}(a_2). \end{aligned} \tag{139}$$

One can show that this formulation is equivalent to (136). The Chebyshev collocation equations for (137) - (139) are apparent. They constitute $N_1 + N_2 + 2$ independent conditions for the $N_1 + N_2 + 2$ unknown nodal values.

Let us consider now a nonlinear generalization, namely,

$$L(u) \equiv \frac{dG(u)}{dx} = f, \tag{140}$$

where the flux, $G(u)$, is given by

$$G(u) = g(u) - \nu \frac{du}{dx}, \tag{141}$$

and $g(u)$ is a nonlinear function of u. The appropriate interface conditions are continuity of u and G. Thus, the condition,

$$G(u_1(a_2)) = G(u_2(a_2)), \tag{142}$$

replaces the last condition in (139).

An alternative patching method consists of replacing the pointwise flux balance condition (142) by an integral version over the two subdomains adjacent to the interface point:

$$G(u_1(a_1)) - \int_{a_1}^{a_2} f(u_1)dx = G(u_2(a_3)) + \int_{a_2}^{a_3} f(u_2)dx. \tag{143}$$

The correct discrete enforcement of (143) requires that appropriate quadrature formulas be employed. Gottlieb and Streett (in preparation) have studied this matter carefully.

Additional considerations arise in two or more dimensions. The interface is now a curve rather than a point. The collocation points on the two domains may not coincide, as illustrated in Figure 9. In this case appropriate interpolation formulas are needed. Moreover, there may be internal cross points, as illustrated in Figure 10. Special care is needed at these points. The integral flux balance condition (143) extends easily to such points. Further details are provided in CHQZ, Sec. 13.2.

Hyperbolic problems require different interface conditions than elliptic ones. Consider the one-dimensional wave equation (83) on Ω. A patching method for this problem reads

$$\begin{aligned} &\frac{\partial u_1}{\partial t} + \frac{\partial u_1}{\partial x} = 0, \qquad \text{on } \Omega_1, \\ &u_1(a_1, t) = u_L(t), \end{aligned} \tag{144}$$

$$\frac{\partial u_2}{\partial t} + \frac{\partial u_2}{\partial x} = 0, \qquad \text{on } \Omega_2, \tag{145}$$

with

$$u_1(a_2, t) = u_2(a_2, t). \tag{146}$$

The collocation equations are obvious, except at the domain interface. One choice is to average the contributions from Ω_1 and Ω_2. However, the temporal stability is dependent upon the particular values of N_1 and N_2. The preferred choice is to collocate (144) rather than (145) since information is travelling to the right; in other words, the interface condition is upwinded.

The upwind interface condition extends in a natural way to quasilinear hyperbolic systems, such as

$$\frac{\partial \mathbf{u}}{\partial t} + A(\mathbf{u})\frac{\partial \mathbf{u}}{\partial x} = 0, \qquad \text{on } \Omega, \tag{147}$$

where $\mathbf{u} \in \mathrm{R}^n$, and $A(\mathbf{u})$ is an $n \times n$ matrix, by a diagonalization procedure. Convergence of the multi-domain discretization of general hyperbolic systems of conservation laws, including the case in which the interface coincides with a shock, is discussed in Quarteroni (1989). An iteration-by-subdomain procedure, taking into account the direction of transfer of characteristic information at the interface, is also discussed there.

Patching methods for the spatial part of hyperbolic problems have generally been coupled with explicit time-discretizations. Elliptic problem, of course, require implicit methods as do implicit time-discretizations of parabolic problems. The matrices which represent the global algebraic system have a block structure due to the domain decomposition, and only adjacent subdomains (or blocks) are coupled. Block Gauss elimination, then, should be considered. So, too, should static condensation (Przmieniecki 1963). These methods are most attractive in the context of time-dependent problems in which the factorization of the global matrices need only be performed once.

An influence-matrix technique is a useful alternative. This enables the solution to the global system to be obtained at the price of two local solutions to Dirichlet problems for each subdomain. The influence-matrix is constructed by solving on each subdomain Dirichlet problems that have boundary data which is 1 at a particular interface point and 0 elsewhere, and then computing the resultant errors in the patching condition at all the interface points. In the first step, arbitrary values are assigned to the interior interface nodes. Then the residuals for the patching conditions are evaluated. These will be nonzero, but the influence-matrix is used to infer the correct interface values. The local solutions are then re-computed with the correct interface values and the desired solution is obtained.

CHQZ, Sec. 13.2.3 cover some of the early solution algorithms. More recent work on the influence-matrix method has been reported by Pulicani (1986), Gottlieb and Hirsh (1988), and Phillips and Karageorghis (1989). Block preconditioners of the influence-matrix have been discussed by Quarteroni and Sacchi-Landriani (1988, 1989).

As mentioned before, domain decomposition methods are often used to improve resolution in problems possessing both large- and small-scale structures. An example of this (taken from Macaraeg and Streett 1986) is the solution of the Burgers equation, written in the form,

$$\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial (u^2)}{\partial x} = \nu \frac{\partial^2 u}{\partial x^2}, \qquad x \in (-1,1), \tag{148}$$

$$u(-1,t) = u(1,t) = 0, \qquad t > 0, \tag{149}$$

$$u(x,0) = -sin\pi x, \qquad x \in (-1,1), \tag{150}$$

and with $\nu = 0.01/\pi$. Table 6 presents a comparison of a conventional single-domain Chebyshev collocation method with a computation which employed three domains. (A mapping was also employed to improve resolution of the internal "shock".) Observe that the multi-domain solution requires only half as many points to achieve accuracy comparable to the single-domain method. Moreover, the decay rate of the multi-domain solution exhibits spectral accuracy.

Method	Discretization	Relative Error in Maximum Slope
single-domain	64	1.31 (-4)
multi-domain	12/13/12	1.99 (-4)
multi-domain	20/21/20	3.23 (-5)
multi-domain	32/33/32	2.14 (-6)

Table 6: Comparison of Single-domain and Multi-domain Solutions to the Burgers Equation

4.2 Variational Methods

The essential aspects of variational spectral domain decomposition methods can be gleaned from an examination of their application to the one-dimensional Helmholtz problem,

$$-\frac{d^2 u}{dx^2} + \lambda u = f, \qquad in\ \Omega,$$

$$u = 0, \qquad on\ \partial\Omega. \tag{151}$$

An equivalent, variational formulation of this problem is that u is the solution to

$$\int_\Omega (\frac{du}{dx}\frac{dv}{dx} + \lambda uv)dx = \int_\Omega fvdx, \qquad for\ all\ v \in E, \tag{152}$$

where E is the space of all functions which vanish on $\partial\Omega$ and which, together with their first derivatives, are square integrable over Ω. On each subdomain the approximation is a polynomial of degree less than or equal to N_s.

In a variational domain decomposition method the space of trial (and test) functions X_N is given by

$$X_N = \{v \in C^0(\overline{\Omega}) \mid v_s \in P_N, for\ s = 1,2\ and\ v = 0\ on\ \partial\Omega\}, \tag{153}$$

where v_s denotes the restriction of v to Ω_s. The space, X_N, is therefore composed of continuous functions which are piecewise polynomials defined on the decomposition of the physical domain, Ω. Eq. (152) is simply replaced by its discrete counterpart in which the test and trial spaces are X_N rather than the full space E:

$$\int_\Omega (\frac{du^N}{dx}\frac{dv}{dx} + \lambda u^N v)dx = \langle f, v\rangle, \qquad for\ all\ v \in X_N, \tag{154}$$

where $\langle f, v\rangle$ is a convenient approximation to $\int_\Omega fvdx$, usually obtained by numerical quadrature.

This is a variational formulation with trial and test functions which are continuous across element (or subdomain) boundaries. Flux continuity, i.e., continuity of the normal derivative, at element interfaces, is not satisfied for fixed N, but only as part of the convergence process, i.e., when N tends to infinity.

In the spectral-element method, the discrete approximation satisfies (154), and, on each subdomain (here called element), u^N has a polynomial expansion of high degree. In the spectral-element method, the basis in Ω_s is formed by Lagrange interpolants using the Legendre collocation points.

After a global node numbering and an assembly of the elemental equations, the algebraic representation of (154) takes the form,

$$CU = BF. \tag{155}$$

The unknown vector, U, contains the values of the discrete solution at all grid-points except the end points of the interval. The matrices, C and B, are positive-definite, symmetric and banded, the bandwidth being determined by the largest N_s.

When a two-dimensional problem is considered, the Cartesian products of the Legendre-Lobatto nodes are used on each subdomain. Then Lagrangian interpolants at these points can be represented using tensor products of Legendre polynomials. Spectral-element methods have also been developed which use isoparametric elements. More details are provided in CHQZ, Sec. 13.3.2.

CHQZ, Sec. 13.5.2 demonstrate an equivalence between certain patching methods and variational methods. The equivalent patching condition in variational methods follows from appropriate quadrature rules. It now appears clear that patching methods yield the best results when the interface condition is derived variationally. Funaro (1988b) presents some recent results on this subject.

The articles by Maday and Patera (1989) and Fischer, Ho, Karniadakis, Ronquist, and Patera (1988) provide recent summaries of the latest developments. Among these is the construction of nonconforming spectral-element methods by Maday, Mavriplis, and Patera (1988).

Early implementations of the spectral-element method favored direct methods, in particular, the static condensation method, for the solution of the implicit system represented by (155). More recently, iterative methods have been employed. See the discussion in Sec. 5.2.5 and the article by Fischer et al (1988).

4.3 Alternating Schwarz Method

This class of domain decomposition methods uses overlapped domains, as illustrated in Figure 8. For elliptic problems an iterative procedure is clearly required in order for the solution on the individual subdomains to converge to the restriction of the solution on the entire domain. CHQZ, Sec. 13.4 discuss some of the techniques currently in use.

Recently, Kopriva (1989a) presented a spectral domain decomposition method on a combination of patched and overlapped grids for hyperbolic equations. Figure 11 illustrates the grid that he employed for a two-dimensional hyperbolic system, and Figure 12 demonstrates the achievement of spectral accuracy in the steady-state. Kopriva (1989a) discusses the merits of two different procedures for handling the overlapped grids.

4.4 Recent Developments

Spectral domain decomposition methods have been an extremely active area of research in the last few years. A recent collection of papers has appeared in a special issue of **Appl. Numer. Math.** (cited individually in the bibliography). Here we summarize some of the recent work which was not covered in CHQZ, Ch. 13.

Several methods have been developed in which finite-difference or finite-element methods are used in some domains, whereas spectral methods are used in others. Kopriva (1989b) coupled upwind finite-difference methods to improve resolution in the vicinity of a shock wave with spectral methods in the smooth part of the flow. Bernardi, Debit, and Maday (1987) proposed and analyzed a method which used finite-elements near a singularity and spectral methods elsewhere.

Gastaldi and Quarteroni (1989) devised a well-posed mathematical formulation of a problem in which the equation itself differed on the subdomains, in particular hyperbolic in one domain and parabolic in the other. They also analyzed a spectral domain decomposition method for this application.

5 Solution of Implicit Equations

The solution of implicit equations is an important component of many spectral algorithms. For steady problems this task is unavoidable, while spectral algorithms for many unsteady problems are only feasible if they incorporate implicit (or semi-implicit) time-discretizations. We concentrate on linear systems, assuming that nonlinear ones are attacked by standard linearization techniques.

We focus on the elliptic equation,

$$\Delta u - \lambda u = f, \tag{156}$$

where f is a function of $\mathbf{x}$ and λ is a constant. The Poisson equation corresponds to $\lambda = 0$, and an implicit time-discretization of the heat equation gives rise to such a system with $\lambda > 0$ (see Sec. 3.4 for the one-dimensional case).

If Fourier series are used for any spatial direction, then a Fourier transform in that direction eliminates the derivative and simply alters λ. For this reason, we discuss only Chebyshev approximations here. However, CHQZ, Sec. 5.1 discuss direct methods for Fourier approximations at some length, including cases involving separable, variable coefficients and mappings.

Spectral collocation approximations lead to a linear system,

$$LU = F, \tag{157}$$

where U and F are vectors consisting of the grid-point values of u, f and any boundary data, and L is a matrix (constructed as a tensor product matrix in two or more spatial dimensions). A similar linear system is obtained for Galerkin and tau approximations, but now U and F are vectors consisting of the expansion coefficients of u, f and the boundary data, and L is the appropriate matrix in transform space.

The spectral linear systems are essentially full. The matrices for one-dimensional problems are dense and the matrices for multi-dimensional problems are tensor products of these full matrices. As Deville and Mund (1989) have stressed, there are a large number of zeros in the spectral matrices for multi-dimensional problems; however, the bandwidth is nearly maximal. Gaussian elimination may, in principle, be applied, but it requires $O(N^{3d})$ operations (where addition, subtraction, multiplication, and division are counted as separate and equal operations) and $O(N^d)$ storage, where d is the dimension of the problem. (We assume, for simplicity, that the number of degrees of freedom in each spatial dimension is N.) In the first section of this chapter we discuss some direct solution techniques which can yield the solution to (157) in $O(N^d)$, $O(N^d log_2 N)$, or at worst $O(N^{d+1})$ operations with at most $O(N^d)$ additional storage. The remainder of the chapter is devoted to a discussion of iterative techniques. These require $O(N^d log_2 N)$ operations per iteration and $O(N^d)$ additional storage.

5.1 Direct Methods

Our objectives in this section are to explain the principles underlying the basic direct solution techniques. We shall call a solution "efficient" if it enables the solution to (157) to be obtained in at most $O(N^d log_2 N)$ operations. This makes the cost of solving (157) comparable, even for N large, to the cost of typical explicit spectral operations such as differentiation and the evaluation of convolution sums. In many cases, a solution cost of $O(N^{d+1})$ is still acceptable in the sense that it only overwhelms the cost of other spectral operations for values of N of 128 or so.

An important consideration is whether only a few or else a large number of solutions to (157) with different data, F, are sought. The latter case is typical of semi-implicit methods

for unsteady problems: hundreds, or even thousands of solutions to a single linear system are required. In such situations, it is reasonable to invest a substantial amount of calculations on a pre-processing stage that greatly reduces the subsequent cost of solving (157).

Let us consider (156) on a square with homogeneous Dirichlet boundary conditions. The collocation approximation to this can be written

$$AU + UB - \lambda U = F, \tag{158}$$

where U is the $(N-1)\times(N-1)$ matrix (u_{ij}) for $i,j = 1,2,\ldots,N-1$, F is defined similarly, A is the second-derivative operator (in x) in which the boundary conditions have been incorporated, and B is the transpose of the second-derivative operator (in y).

Systems of the form (158) are solvable by Schur-decomposition. An orthogonal transformation is used to reduce A to block-lower-triangular form with blocks of size at most 2. Similarly, B is reduced to block-upper-triangular form. If P and Q denote the respective orthogonal transformations, then (158) is equivalent to

$$A'U' + U'B' - \lambda U' = F', \tag{159}$$

where

$$\begin{aligned} A' &= P^T A P, \\ B' &= Q^T B Q, \\ U' &= P^T U Q, \\ F' &= P^T F Q. \end{aligned} \tag{160}$$

The solution process has 4 steps: 1) reduction of A and B to real Schur form (and determination of P and Q); 2) construction of F' via (160); 3) solution of (159) for U'; and 4) transformation of U' to U via the inverse of (160).

The first step can be accomplished via the QR algorithm in approximately $40N^3$ operations. Step 3 requires $2N^3$ operations, and steps 2 and 4 take $4N^3$ operations apiece. A single solution requires roughly $50N^3$ operations. Hence, step 1 is the most time-consuming. When the same problem must be solved repeatedly, then step (1) need only be performed once, in a pre-processing stage. The matrices A', B', P, and Q may then be stored and used as needed. In this case a complete solution takes $10N^3$ operations.

The matrix-diagonalization approach is similar to the Schur-decomposition method. The difference is that the matrices, A and B, in (158) are diagonalized rather than merely reduced to block-triangular form. An algebraic problem of the form (159) is obtained with (160) replaced by

$$\begin{aligned} A' &= P^{-1} A P = \Lambda_A, \\ B' &= Q^{-1} B Q = \Lambda_B, \\ U' &= P^{-1} U Q, \\ F' &= P^{-1} F Q, \end{aligned} \tag{161}$$

where Λ_A is the diagonal matrix with the eigenvalues of A on the diagonal. Thus, we have

$$\Lambda_A U' + U' \Lambda_B - \lambda U' = F'. \tag{162}$$

The matrices, P and Q, are not necessarily orthogonal and their columns consist of the eigenvectors of A and B, respectively.

The matrix-diagonalization scheme for (158) consists of the same 4 steps as the Schur-decomposition method, except that the first, pre-processing stage also requires that the eigenvectors and the inverse transformations be computed. This takes an additional $8N^3$ operations. Step 3 takes only $3N^2$ operations since the system is diagonal, and steps 2 and 4 require $4N^3$ operations apiece, as before.

For collocation problems requiring multiple solutions, the matrix-diagonalization method has the advantage of taking only 80% of the solution time of the Schur-decomposition method: $8N^3$ operations. Moreover, the entire solution process – steps 2, 3 and 4 – is extremely simple and can be optimized readily. In fact, most computer libraries contain assembly-language routines for all the requisite operations. The third stage of the Schur-decomposition method is more complicated.

In the case of tau approximations to (156), further gains in efficiency are possible. The discrete problem may be written in the form (158) where U is the $(N-1) \times (N-1)$ matrix, $(\hat{u}_{nm})$, consisting of the Chebyshev coefficients of u (minus those used to enforce the boundary conditions). The matrix, F, is defined similarly, and A and B are the representations in transform space of the second-derivative operator (with the boundary conditions used to eliminate the 2 highest-order coefficients in each direction).

In the case of Dirichlet (or Neumann) boundary conditions, the even and odd modes decouple. Thus, A, B, P, and Q contain alternating zero and non-zero elements. This property may be exploited to reduce the cost of both the pre-processing step (by a factor of 4) and the matrix multiplies (by a factor of 2). The cost of steps 2 through 4 is thus $4N^3$.

The cost of the solution stages may be halved again by performing the diagonalization in only 1 direction and resorting to a standard tau solution in the other. Thus, (158) is reduced to

$$AU' + U'G - \lambda U' = F', \tag{163}$$

where

$$\begin{aligned} U' &= UQ, \\ F' &= FQ, \end{aligned} \tag{164}$$

instead of (161). The system (163) decouples into $N-1$ systems of the form (113). Each of these may be reduced to a system like (115) and solved accordingly in $16N$ operations. The cost of the solution process is essentially halved, to $2N^3$ operations, since the number of matrix multiplies is cut in two.

In these algorithms, as indeed with matrix computations in general, the accumulation of round-off error is a concern. Haidvogel and Zang (1979) reported the loss of 3 to 4 digits (for N between 16 and 64) with the Schur-decomposition method. These were recovered through iterative improvement. Since the computation of eigenvectors can be a sensitive process, double-precision is advisable for the pre-processing stage of the matrix-diagonalization method.

Both methods can be generalized. The use of Neumann or Robin boundary conditions is straightforward. However, with Robin boundary conditions the even and odd modes do not de-couple; hence, some of the economies of the tau method are lost. These methods can also be applied to separable equations of the form,

$$\frac{\partial}{\partial x}\Big(a(x)\frac{\partial u}{\partial x}\Big) + \frac{\partial}{\partial y}\Big(b(y)\frac{\partial u}{\partial y}\Big) - \lambda u = f. \tag{165}$$

As it happens, these methods are more attractive in three-dimensional problems than in two-dimensional ones. Suppose that the number of degrees of freedom in each direction is N. The

pre-processing cost is some large multiple of N^3. In two dimensions, the solution cost is a small multiple of N^3, and typical explicit spectral calculations take $O(N^2 log_2 N)$ operations. Thus, the pre-processing cost is substantially larger than the cost of a single solution. In three dimensions, the pre-processing cost remains $O(N^3)$, the solution cost is a small multiple of N^4 and typical explicit spectral calculations require $O(N^3 log_2 N)$ operations. Thus, the pre-processing cost may even be smaller than the cost of the solution phase. Similarly, the extra memory required for P and its inverse is proportionately smaller in three dimensions than in two.

5.2 Iterative Methods

The direct methods discussion above have limited application compared to iterative methods. Invariably, iterative solution schemes are applied to spectral methods employing collocation.

5.2.1 Richardson Iteration

The fundamentals of iterative methods are perhaps easiest to grasp for the simple model problem,

$$-\frac{d^2u}{dx^2} = f, \tag{166}$$

on $(0, 2\pi)$ with periodic boundary conditions. The Fourier collocation approximation to the left-hand side of (166) at the collocation points is

$$\sum_{p=-\frac{N}{2}+1}^{\frac{N}{2}-1} p^2 \tilde{u}_p e^{ipx_j}. \tag{167}$$

The spectral approximation to (166) may be expressed as

$$LU = F, \tag{168}$$

where $U = (u_0, u_1, \ldots, u_{N-1})$, $F = (f_0, f_1, \ldots, f_{N-1})$, and L represents the Fourier spectral approximation to $-d^2/dx^2$.

A Richardson's iterative scheme for solving (168) is

$$V^{n+1} = V^n + \omega(F - LV^n), \tag{169}$$

where V^n is the n^{th} iterate and ω is a relaxation parameter. The eigenfunctions of L are

$$\psi_j(p) = e^{2\pi jp/N}, \qquad p = -\frac{N}{2}, \ldots, \frac{N}{2} - 1, \tag{170}$$

with the corresponding eigenvalues,

$$\lambda_p = p^2, \qquad p = -\frac{N}{2}, \ldots, \frac{N}{2} - 1, \tag{171}$$

The index p has a natural interpretation as the frequency of the eigenfunction.

The error at any stage of the iterative process is $V^n - U$; it can be resolved into an expansion in terms of the eigenvectors of L. The error obeys the relation,

$$(V^{n+1} - U) = G(V^n - U), \tag{172}$$

where the iteration matrix of the Richardson scheme is given by

$$G = I - \omega L. \tag{173}$$

The iterative scheme is convergent if the spectral radius ρ of G is less than 1.

Each iteration reduces the p^{th} error component to $\nu(\lambda_p)$ times its previous value, where

$$\nu(\lambda) = 1 - \omega\lambda. \tag{174}$$

(This is referred to as the damping factor.) The optimal choice of ω results from minimizing $\mid \nu(\lambda) \mid$ for $\lambda \in [\lambda_{min}, \lambda_{max}]$, where $\lambda_{min} = 1$ and $\lambda_{max} = \frac{N^2}{4}$. (One need not worry about the $p = 0$ eigenfunction since it corresponds to the mean level of the solution, which is at one's disposal for this problem.) The optimal relaxation parameter for this procedure is

$$\omega = \frac{2}{\lambda_{max} + \lambda_{min}}. \tag{175}$$

It produces the spectral radius,

$$\rho = \frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}} = \frac{\kappa - 1}{\kappa + 1}, \tag{176}$$

where

$$\kappa = \frac{\lambda_{max}}{\lambda_{min}} \tag{177}$$

is referred to as the spectral condition number. Unfortunately, $\kappa \simeq \frac{N^2}{4}$ and $\rho \simeq 1 - \frac{8}{N^2}$, which implies that $O(N^2)$ iterations are required to achieve convergence.

5.2.2 Preconditioning

The primary cause of the inefficiency of the straightforward Richardson method is the rapid increase with N of the spectral condition number. This can be alleviated by "preconditioning" the problem, in effect solving

$$(H^{-1}ML)U = (H^{-1}M)F \tag{178}$$

rather than (168). A preconditioned version of (169) is

$$V^{n+1} = V^n + \omega H^{-1}M(F - LV^n). \tag{179}$$

In practice the inverse of the preconditioning matrix, H, is never explicitly required; instead one solves

$$H(V^{n+1} - V^n) = \omega M(F - LV^n), \tag{180}$$

The effective iteration matrix is

$$G = I - \omega H^{-1}ML. \tag{181}$$

One obvious requirement for H is that (180) can be solved inexpensively, i.e., in fewer operations than are required to evaluate LV; it is also desirable for MV to be inexpensive to evaluate. The second requirement on the preconditioning matrices is that $H^{-1}M$ be a good approximation to L^{-1}, i.e., that the spectral condition number of $H^{-1}ML$ be small.

A finite-difference approximation to L is one useful choice for the preconditioner. In this case H represents a finite-difference approximation to the left-hand side of (166) and M is the identity matrix. A finite-element discretization is another, and usually superior, choice. In practical

terms the principal difference between finite-difference and finite-element preconditioning is that the latter includes a weighting of the residuals. Here H is the appropriate stiffness matrix and M is the mass matrix. For one-dimensional problems, only $O(N)$ operations are required to solve (180) and MV takes only $O(N)$ operations. The evaluation of LV takes $O(N\ log_2 N)$ operations. Hence, the preconditioning work is a minor part of the cost of a single iteration.

Specific formulas for finite-difference preconditioners are fairly straightforward and are discussed in CHQZ, Sec. 5.2. Canuto and Pietra (1987) and Deville and Mund (1989) both provide thorough discussions of the appropriate formulation of the finite-element preconditioners.

Simple estimates can be made of the effectiveness of various preconditioners for the model problem (178) since the eigenfunctions of the preconditioned problem have the form e^{ikx_j}. Let an eigenvalue of the preconditioned operator, $H^{-1}ML$, be denoted by Λ. For example, a simple calculation indicates that the eigenvalues arising from second-order finite-differences are

$$\Lambda_p = (p\Delta x/2)^2/sin^2(p\Delta x/2), \qquad p = -\frac{N}{2}, \ldots, \frac{N}{2} - 1. \tag{182}$$

Clearly, $\Lambda_{min} = 1$, and $\Lambda_{max} = \pi^2/4$. Table 7 summarizes the properties of second-order finite-difference (FD2), fourth-order finite-difference (FD4), and linear finite-element (FE1) preconditioning. It turns out that these simple estimates are also good predictors of the performance of these preconditioning techniques for Chebyshev approximations to non-periodic problems. Note that the finite-element preconditioning achieves the smallest condition number, just 1.44.

Preconditioner	Λ_{min}	Λ_{max}	ρ
FD2	1.00	2.47	0.424
FD4	1.00	1.85	0.298
FE1	.693	1.00	0.181

Table 7: Properties of Preconditionings for the Model Problem

Ronquist and Patera (1987b), in the context of the spectral element method, have used scaled Jacobi preconditioning, i.e., $H = \alpha D$, where D is the matrix containing the diagonal elements of L, and α is a scaling parameter chosen so that the spectral radius of H matches that of L. The primary application of this has been in the context of multigrid methods and it is discussed further in Section 5.2.5.

Central-difference preconditioners work well for problems which are dominated by even-order derivatives. When significant odd-order derivatives are present, however, other preconditioners must be employed. The Fourier first-derivative operator, preconditioned by second-order central differences, has the eigenvalues

$$\Lambda_p = (p\Delta x)/sin(p\Delta x), \qquad p = -\frac{N}{2}, \ldots, \frac{N}{2} - 1. \tag{183}$$

Now, $\Lambda_{min} = 1$, but $\Lambda_{max} = \infty$. No iterative scheme can succeed for this spectrum. Two alternatives are to employ a staggered grid, or to apply the preconditioning on a grid with more points, typically 50% more, than the basic grid. The former alternative is discussed in CHQZ, Sec. 5.2.2.

In the latter case, effective preconditioning of second derivatives in a mixed-order problem is retained. Figure 13, taken from Streett and Macaraeg (1986), shows how the eigenvalue

spectrum of the preconditioned Chebyshev first-derivative operator is compressed in shifting the central-difference preconditioning grid from the original one to ones which are 20% and 50% finer.

5.2.3 Conventional Iterative Methods

Much better iterative methods than the elementary Richardson scheme discussed above are available. Both non-stationary Richardson and Chebyshev acceleration of Richardson iteration are superior: in effect the term κ in (176) is replaced by $\sqrt{\kappa}$. Variational iterative schemes such as minimum residual and conjugate gradient are other possibilities. They have the advantage of being parameter-free. See CHQZ, Sec. 5.3 for a thorough discussion.

5.2.4 Multi-dimensional Problems

Consider now the generalized Helmholtz problem,

$$\nabla \cdot (a(\mathbf{x})\nabla u) - \lambda u = f, \tag{184}$$

in two-dimensions. The discrete approximation to (184) may also be written in the form (168), with obvious notation. Assuming that N collocation points are employed in each direction, the evaluation of LV takes $O(N^2 log_2 N)$ operations, and the term MV costs $O(N^2)$ operations. The direct solution of (180) takes $O(N^3)$ operations by a banded matrix solver after the $O(N^4)$ cost of factoring the matrix has been paid. In this case the cost of a single iteration is dominated by the preconditioning. The situation is even worse in three dimensions, in which the cost of evaluating LV is only $O(N^3 log_2 N)$, the cost of solving (180) is $O(N^4)$, while the factorization expense climbs to $O(N^5)$.

Of course, the finite-difference or finite-element preconditioner may be inverted approximately by an iterative scheme, preferably conjugate gradient or multigrid. The essential approach is the iterative solution of the preconditioned system (180). Experience indicates that the iterative solution of the preconditioned system need not be carried out to a high degree of convergence; typically a reduction in the residual of the preconditioned problem by two to three orders of magnitude suffices for convergence of the overall iteration. Multigrid solution of the finite-difference preconditioned system is particularly effective.

Alternatively, in two and three dimensions the economics may favor trading off less effective iterative schemes (accepting larger spectral radii) in favor of less expensive preconditioners. Applications to date in this area resort to starting with a finite-difference approximation and using incomplete-LU decompositions, ADI relaxation, or alternating line relaxation (see CHQZ, Sec. 5.4).

ADI and alternating line relaxation are well-suited to vector and shared-memory parallel processors; incomplete-LU decompositions are not. Of course, these methods would surely benefit from replacing the finite-difference approximation with a finite-element one. An interesting alternative, which has apparently not yet been tried for spectral methods, is to employ the polynomial preconditioning strategy of Dubois, Greenbaum, and Rodrique (1979). Their approach is to write $H = D(I - G)$, where D is diagonal, and to use a polynomial in G as an approximation to $(I - G)^{-1}$. In this case the trivial inversion of D and the approximate inversion of $(I - G)$ both vectorize and parallelize well. This approach does require $||G|| < 1$ for convergence. Wong (1989) showed that for solving fourth-order discretizations, an effective strategy is to use the second-order finite-difference approximation to determine G. This use of the second-order finite-difference approximation will not succeed as a preconditioner for spectral discretizations,

since the condition number of $H^{-1}ML$ is greater than 2. (It must be less than 2 to satisfy the requirement on G.) However, finite-element approximations should work well (see Table 7).

5.2.5 Multigrid Methods

Iterative schemes for spectral collocation equations can be accelerated dramatically by applying multigrid concepts. Briefly put, multigrid methods take advantage of a property shared by a wide variety of relaxation schemes - potential efficient reduction of the high-frequency error components, but unavoidable slow reduction of the low-frequency components. This slow convergence is the outcome of balancing the damping of the lowest-frequency eigenfunction with that of the highest-frequency one in the minimax problem described after (174). The multigrid approach takes advantage of the fact that the low-frequency modes ($|p| < N/4$ for a coarse grid with only half as many points as the fine grid) can be represented just as well on coarser grids. It settles for balancing the middle-frequency eigenfunction ($|p| = N/4$) with the highest-frequency one ($|p| = N/2$), and hence damps effectively only those modes which cannot be resolved on coarser grids. In (175) and (176), λ_{min} is replaced with $\lambda_{mid} = \lambda(\frac{N}{4})$. The optimal relaxation parameter in this context is

$$\omega_{MG} = \frac{2}{\lambda_{max} + \lambda_{mid}}. \tag{185}$$

The multigrid smoothing factor,

$$\mu_{MG} = \frac{\lambda_{max} - \lambda_{mid}}{\lambda_{max} + \lambda_{mid}}, \tag{186}$$

measures the damping rate of the high-frequency modes. In this example, $\mu_{MG} = 0.60$, independent of N. The price of this effective damping of the high-frequency errors is that the low-frequency errors are hardly damped at all. Table 8 compares the single-grid and multigrid damping factors (see (174)) for $N = 64$. However, on a grid with $N/2$ collocation points, the modes for $| p | \in [\frac{N}{8}, \frac{N}{4}]$ are now the high-frequency ones. They get damped on this grid. Still coarser grids can be used until relaxations are so cheap that one can afford to damp all the remaining modes, or even to solve the discrete equations exactly. For the case illustrated in Table 8, the high-frequency error reduction in the multigrid context is roughly 250 times as fast as the single-grid reduction for $N = 64$.

As in most multigrid methods the smoothing rates are distinctly slower as the dimensionality of the problem increases. For this simple model problem, the smoothing rate is 0.33 in one dimension, 0.60 in two dimensions, and 0.78 in three dimensions.

Let us consider just the interplay between two grids. A general, nonlinear fine-grid problem can be written

$$L^f(U^f) = F^f. \tag{187}$$

The shift to the coarse grid occurs after the fine-grid approximation, V^f, has been sufficiently smoothed by the relaxation process, i.e., after the high-frequency content of the error, $V^f - U^f$, has been sufficiently reduced. The related coarse-grid problem is

$$L^c(U^c) = F^c, \tag{188}$$

where

$$F^c = R[F^f - L^f(V^f)] + L^c(RV^f). \tag{189}$$

The restriction operator, R, interpolates a function from the fine grid to the coarse grid. The coarse-grid operator and solution are denoted by L^c and U^c, respectively. After an adequate

p	Single-Grid	Multigrid
1	.9980	.9984
2	.9922	.9938
4	.9688	.9750
8	.8751	.9000
12	.7190	.7750
16	.5005	.6000
20	.2195	.3750
24	.1239	.1000
28	.5298	.2250
32	.9980	.6000

Table 8: Damping Factors for N = 64

approximation, V^c, to the coarse-grid problem has been obtained, the fine-grid approximation is corrected via

$$V^f \leftarrow V^f + P(V^c - RV^f). \tag{190}$$

The prolongation operator, P, interpolates a function from the coarse grid to the fine grid. A common strategy for shifting between grids is illustrated in Figure 14. This is called a V-cycle. Usually a coarse grid has only half as many points (in each coordinate direction) as the fine grid.

A complete multigrid algorithm requires specific choices of the interpolation operators, the coarse-grid operators, the relaxation schemes, and a strategy for shifting between grids. Generally, the effective spectral radius, which measures the average reduction in residual per fine-grid relaxation, is somewhat slower than the smoothing factor, μ_{MG}, due to errors introduced in the interpolation stages. These issues are discussed at length in CHQZ, Sec. 5.5 for both Fourier and Chebyshev multigrid methods.

The isotropic, fully periodic Poisson problem ($\lambda = 0$) can be solved extremely efficiently by spectral multigrid methods. The discretization of problem (184) is called isotropic if

$$\frac{a}{\Delta x^2} = \frac{a}{\Delta y^2} = \frac{a}{\Delta z^2}. \tag{191}$$

Using a residual weighting technique, the effective spectral radius can be as low as 0.11 in two dimensions and 0.18 in three dimensions. (Note the slower convergence with increased dimensionality.) This is achieved with purely explicit relaxation. Preconditioning is entirely unnecessary. Details are given in Brandt, Fulton, and Taylor (1985), Erlebacher, Zang, and Hussaini (1987), and Heinrichs (1988b).

It is important to recognize that these effective spectral radii presume that the coefficient $a(\mathbf{x})$ is reasonably well-behaved. As Erlebacher, Zang, and Hussaini (1987) demonstrated, if this coefficient is highly oscillatory, then the convergence deteriorates, and if the coefficient is basically random, the spectral multigrid method fails. Of course, similar behavior occurs for finite-difference multigrid methods for such problems. Nevertheless, numerical examples have been presented of successful spectral multigrid solutions of the isotropic problem with fairly strongly variable coefficients.

If the fully periodic problem is anisotropic, or if the coefficient λ is non-zero, then a different strategy must be employed. The best procedure demonstrated to date in two dimensions is to

employ a finite-difference preconditioner together with alternating line relaxation. Effective convergence rates between 0.3 and 0.4 have been reported by Brandt, Fulton, and Taylor (1985) and Heinrichs (1988b). Note that the cost of this type of preconditioner is insignificant compared with the cost of the spectral residual evaluation. As a preconditioner for a single-grid iterative scheme, alternating line relaxation performs poorly, since the eigenvalues associated with the smooth components of the solution tend rapidly to zero with N. Combining alternative line relaxation with the spectral multigrid scheme provides both an effective (and nearly grid-independent) convergence rate with an iterative scheme costing only $O(N^2 log_2 N)$ operations per iteration.

For problems with at least one non-periodic direction, preconditioning is again essential. Heinrichs (1988a) has demonstrated that in two dimensions alternating line relaxation produces the same effective spectral radius (between 0.3 and 0.4) that it achieves for the fully periodic, anisotropic problem. He also observed that for Dirichlet boundary conditions, minimum residual relaxation was more robust than Richardson relaxation. For Neumann boundary conditions, however, nonstationary Richardson must be used for minimum residual fails to converge.

Spectral multigrid methods for the Legendre-based spectral-element method have recently been developed (Ronquist 1988, Ronquist and Patera 1987b). To date they have been based on the Jacobi preconditioning discussed in Section 5.2.2. In one dimension the effective spectral radius for the simple Poisson problem is 0.7, independent of both K, the number of elements, and of N, the degree of approximation in each element. In two dimensions, the multigrid condition number is again independent of K but it grows as N. (In contrast, this apparently grows no faster than $N^{1/8}$ for the spectral multigrid methods utilizing more sophisticated preconditioning.) The best two-dimensional spectral radii that have yet been achieved are above 0.9, and these have been for N no larger than 12. To date, this method has not been tested for variable coefficients anywhere near as extreme as those for which conventional spectral multigrid methods have succeeded.

The essential problem is that as N grows the aspect ratio of the grid formed by the collocation points increases. This is akin to having an anisotropic problem on a uniform grid. For finite-difference and finite-element multigrid, as well as for the conventional spectral multigrid methods as discussed above, the only known effective relaxation procedures for two-dimensional anisotropic problems involve incomplete-LU decomposition or line relaxation. It is our opinion that the proper route to acceptable convergence rates for two-dimensional spectral-element multigrid involves replacing the diagonal preconditioning used to date with more effective preconditioners.

Maday and Munoz (1989) have furnished a rigorous proof of the convergence of this spectral multigrid method. However, it does presume sufficient smoothness for the coefficient a, so that the potential failure of the method to converge for highly oscillatory coefficients is not addressed. (Note that even if the coefficient is well-resolved on the fine grid, it may be too poorly resolved on a much coarser grid for convergence.)

As yet, spectral multigrid methods have not been applied to three-dimensional anisotropic problems. However, past experience with finite-difference approximations to such problems suggests that in three dimensions alternating line relaxation alone will not suffice, regardless of whether one is dealing with spectral or spectral-element multigrid; instead, alternating plane relaxation is needed. The article by Dendy (1987) contains a summary of some recent experience with finite-difference multigrid methods for anisotropic three-dimensional problems. The planar relaxation itself can be accomplished by alternating line relaxation within a multigrid context or by preconditioned conjugate gradient iteration.

We also recommend that finite-difference preconditioning for spectral multigrid methods be

replaced by finite-element preconditioning. Table 7 makes it clear that superior convergence rates will be achieved. In fact, we project that ultimately the combination of alternating line (or, in three-dimensions, plane) relaxation with finite-element preconditioning will produce effective spectral radii as good as 0.25, even for anisotropic three-dimensional problems, but with reasonably well-behaved coefficients.

6 Euler Applications

Spectral methods have met with mixed success for problems in inviscid flow. One form of the Euler equations is

$$\begin{aligned}
&\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \\
&\frac{\partial (\rho \mathbf{u})}{\partial t} + \nabla \cdot (\mathbf{u}\mathbf{u}) + \nabla p = 0, \\
&\frac{\partial p}{\partial t} + \mathbf{u} \cdot \nabla p + \gamma p \nabla \cdot \mathbf{u} = 0,
\end{aligned} \tag{192}$$

where ρ is the density, $\mathbf{u}$ the velocity, p the pressure, and γ the ratio of specific heats (an ideal gas law is presumed here). The overwhelming difficulty is that this is a nonlinear hyperbolic system, and it admits non-smooth solutions – shock waves and contact discontinuities. A priori, one expects spectral methods to perform poorly for such problems. Two strategies for dealing with shock-wave discontinuities have been employed – shock-capturing and shock-fitting.

6.1 Shock-capturing

In this approach the discontinuity is located in the interior of the domain. Finite-difference methods require some form of explicit or implicit artificial viscosity to handle the shock. Shock-capturing spectral methods have resorted to a combination of explicit artificial viscosity and sophisticated filtering techniques.

Filtering methods for spectral methods have been investigated for linear model problems, and CHQZ, Secs. 8.3 and 12.1.4 summarize this work. Some recent contributions to artificial viscosities for spectral approximations to the periodic conservation laws have been made by Tadmor (1989) and by Maday and Tadmor (1989). Cai, Gottlieb, and Shu (1989) have employed discontinuity subtraction techniques to construct spectral schemes with bounded total variation. In all three cases exponential convergence for smooth solutions has been proven. Of course, the real challenge is to produce exponential convergence in smooth regions for discontinuous solutions. Thus far, numerical tests have been disappointing.

It is as true today as it was 4 years ago (Hussaini, Kopriva, Salas, and Zang (1985a)) that although spectral shock-capturing methods have succeeded in producing solutions to the Euler equations, they have yet to demonstrate that they are superior or even comparable in efficiency to modern upwind finite-difference, finite-element, and finite-volume approaches.

For the simpler inviscid problem of compressible potential flow, however, the dependent variable itself is continuous. In this case spectral shock-capturing methods (combined with a spectral multigrid solution algorithm) have proven most efficacious. See CHQZ, Sec. 8.5.1.

An interesting alternative is the use of a domain decomposition method in which the region surrounding the shock is treated by an upwind finite-difference scheme and the remaining part of the flow by a Chebyshev spectral method. Kopriva (1989b) has demonstrated that this approach

produces accuracy in the spectral domains that is superior to the accuracy achievable by a pure upwind scheme.

6.2 Shock-fitting

In the spectral shock-fitting approach the burden of resolving the discontinuity is removed by making the shock one of the boundaries of the domain. The numerical method, then, computes the shape and motion of the shock front as part of the solution process. This technique is described in CHQZ, Sec. 8.6.

Recently, Kopriva, Hussaini, and Zang (in preparation) in a sequel to Hussaini, Kopriva, Salas, and Zang (1985b), have improved the shock-fitting method used in their earlier work. Improvements have been made in both the shock-fitting itself and in the boundary conditions. They have re-visited the blunt-body problem that was reported in Hussaini, Kopriva, Salas, and Zang (1985b). The boundary conditions on the body were changed to a compatibility condition, and the complete (rather than approximate) compatibility condition was applied at the shock. As a result of these improvements, not only is a true steady-state achievable (to within machine precision), but also spectral accuracy is achieved, even without periodic filtering of the solution. Figure 15 displays the decay of the error in the stagnation pressure (which is known exactly) as the grid is resolved for Mach 4 flow past a circular cylinder. See this forthcoming work for the details.

7 Navier-Stokes Applications

The convection form of the incompressible Navier-Stokes equations is

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} - \nu \Delta \mathbf{u} + \nabla p = \mathbf{f}, \tag{193}$$

$$\nabla \cdot \mathbf{u} = 0. \tag{194}$$

The velocity, $\mathbf{u} = (u_x, u_y, u_z)$, the static pressure is denoted by p, kinematic viscosity by ν, and a possible source term is denoted by $\mathbf{f}$.

A simpler problem, and one which exhibits most of the numerical difficulties is the Stokes problem,

$$\frac{\partial \mathbf{u}}{\partial t} - \nu \Delta \mathbf{u} + \nabla p = \mathbf{f}, \tag{195}$$

$$\nabla \cdot \mathbf{u} = 0. \tag{196}$$

The nonlinear term is no longer explicitly present. Spectral discretizations of the Navier-Stokes equations usually include an explicit treatment of the nonlinear term, and one may consider that it has been absorbed into $\mathbf{f}$. For most purposes the basic algorithms can be stated for the Stokes problem.

The numerical analysis of spectral methods for the steady Stokes problem,

$$-\nu \Delta \mathbf{u} + \nabla p = \mathbf{f}, \tag{197}$$

$$\nabla \cdot \mathbf{u} = 0, \tag{198}$$

has been the first step in an analysis of the unsteady Navier-Stokes equations. The second step is the analysis of the steady Navier-Stokes equations. The inclusion of the nonlinear terms in the analysis is usually accomplished by resorting to an implicit function theorem (see CHQZ, Sec. 11.3). Analysis of fully-discrete approximations to the unsteady problems is much more difficult, and results are only now beginning to appear (Sacchi-Landriani and Vandeven 1989).

7.1 Fourier Methods for Fully Periodic Problems

Homogeneous, isotropic turbulence is perhaps the one fluid dynamical problem for which strictly periodic boundary conditions in all spatial directions are justifiable. Hence, Fourier spectral methods are ideally suited for this class of problems. Moreover, since the nonlinearities of the Navier Stokes equations are at worst quadratic, Fourier Galerkin methods are the most natural and efficient spectral techniques for this problem.

7.1.1 Galerkin Methods

The primitive variables are expanded in a three-dimensional Fourier series:

$$\mathbf{u}(\mathbf{x},t) = \sum_{\mathbf{k}} \hat{\mathbf{u}}(t) e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{199}$$

$$p(\mathbf{x},t) = \sum_{\mathbf{k}} \hat{p}(t) e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{200}$$

where $\mathbf{k} = (k_x, k_y, k_z)$, and each component ranges between $-\frac{N}{2}$ and $\frac{N}{2} - 1$. The semi-discrete Galerkin approximation to (193) - (194) is

$$\frac{d\hat{\mathbf{u}}_{\mathbf{k}}}{dt} + \nu k^2 \hat{\mathbf{u}}_{\mathbf{k}} + i\mathbf{k}\hat{p}_{\mathbf{k}} = \hat{\mathbf{r}}_{\mathbf{k}}, \tag{201}$$

$$i\mathbf{k}\cdot\hat{\mathbf{u}}_{\mathbf{k}} = 0, \tag{202}$$

where $k^2 = |\mathbf{k}|^2$, and

$$\hat{\mathbf{r}}_{\mathbf{k}} = -\widehat{(\mathbf{u}\cdot\nabla\mathbf{u})}_{\mathbf{k}} + \mathbf{f}_{\mathbf{k}}. \tag{203}$$

The incompressibility condition implies that the pressure term in (201) is

$$\hat{p}_{\mathbf{k}} = -i\,\mathbf{k}\frac{(\mathbf{k}\cdot\hat{\mathbf{r}}_{\mathbf{k}})}{k^2}. \tag{204}$$

The only complication in this Orszag-Patterson algorithm is the evaluation of the nonlinear term, $\widehat{(\mathbf{u}\cdot\nabla\mathbf{u})}_{\mathbf{k}}$. This is done most efficiently by transform methods and costs $O(N^3 log_2 N)$ operations. The one-dimensional version of transform methods was discussed in Section 3.6. If the naive (aliased) transform method is applied, then the algorithm is a pseudospectral one. If the de-aliased transform method is used, then the algorithm is a true spectral Galerkin method.

The Galerkin method can be implemented with as few as 8 three-dimensional FFT's per step. A linear coordinate transformation has been developed that permits simulation of flows with constant strain, shear, and rotation within the confines of periodic boundary conditions. See CHQZ, Sec. 7.2 for more details.

7.1.2 Collocation Methods and the Skew-symmetric Form

Fourier collocation approximations to this problem are also possible. For these approximations use of the appropriate form of the Navier-Stokes equations is crucial.

The standard convection form is given by (193), and the divergence form differs from it by the replacement of the nonlinear term by $\nabla\cdot(\mathbf{u}\,\mathbf{u})$. The rotation form of the momentum equation, given by

$$\frac{\partial\mathbf{u}}{\partial t} + \omega\times\mathbf{u} - \nu\Delta\mathbf{u} + \nabla(\frac{1}{2}|\mathbf{u}|^2 + p) = \mathbf{f}, \tag{205}$$

where the vorticity,

$$\omega = \nabla \times \mathbf{u}, \tag{206}$$

has long been the preferred form. However, recently Zang (1989), following the work of Horiuti (1987), has demonstrated that the skew-symmetric form, given by

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{1}{2}\mathbf{u} \cdot \nabla \mathbf{u} + \frac{1}{2}\nabla \cdot (\mathbf{u}\mathbf{u}) - \nu \Delta \mathbf{u} + \nabla p = \mathbf{f}, \tag{207}$$

is superior. The term skew-symmetric is used because the operator $\frac{1}{2}\mathbf{v} \cdot \nabla \mathbf{u} + \frac{1}{2}\nabla \cdot (\mathbf{v}\mathbf{u})$ (for fixed $\mathbf{v}$ satisfying $\nabla \cdot \mathbf{v} = 0$) is skew-symmetric.

As discussed in Chapter 2, for Fourier Galerkin methods the projection (spatial discretization) and differentiation operators commute, but for collocation methods they do not. This means that for the semi-discrete momentum equation, the Fourier Galerkin convection, divergence, rotation, and skew-symmetric forms are algebraically identical. On the other hand, for collocation methods none of these forms are algebraically identical.

Consider now the conservation properties for the ideal inviscid ($\nu = 0$), fully periodic case. The ideal Navier-Stokes equations conserve linear momentum, $\int \mathbf{u} dV$, and kinetic energy, $\int \frac{1}{2}|\mathbf{u}|^2 dV$. For the semi-discrete ideal equations, a Galerkin method conserves the discrete counterparts of both quantities. However, for a collocation method the semi-discrete conservation properties depend upon the exact form of the momentum equation which is discretized. In particular, for the rotation and skew-symmetric forms, both quantities are conserved, for the divergence form only linear momentum is conserved, and for the convection form neither quantity is conserved.

The conservation of kinetic energy is virtually mandatory for a simulation to be numerically stable in time. Thus, for quite practical reasons use of the convection or divergence forms is ruled out. Although either of these forms alone is numerically unstable, a method which uses the convection and divergence forms on alternate time steps appears to be well-behaved. We shall refer to this version as the alternating form.

Only 6 derivatives are required for the evaluation of the nonlinear terms for a collocation approximation in the rotation form, whereas 18 derivatives are needed for the skew-symmetric form. The convection, divergence, and alternating forms take 9 derivatives. (By invoking the incompressibility constraint, the number of derivatives for the convection and skew-symmetric forms can be reduced by 1.) Thus, the rotation form takes the least work for the evaluation of the nonlinear terms, the alternating form slightly more, and the skew-symmetric form appreciably more. Recall, however, that there is additional work required for the pressure and viscous terms.

Zang (1989) illustrated the difference between the Fourier collocation methods using the rotation and skew-symmetric (actually the alternating) forms, and compared them with the Galerkin method on a homogeneous turbulence problem involving a mean flow with uniform shear. The initial condition consisted of a random divergence-free velocity field with a prescribed three-dimensional energy spectrum. The difference between the algorithms was dramatically evident in the temporal evolution of the turbulence intensity,

$$q = \langle u_x^2 + u_y^2 + u_z^2 \rangle^{\frac{1}{2}}, \tag{208}$$

where $\langle \cdot \rangle$ denotes a spatial average. Figure 16 displays the results for the rotation, alternating, and Galerkin forms of the nonlinear term for the test problem on 64^3, 96^3, and 128^3 grids. The rotation form on the 64^3 grid proved incapable of sustaining the turbulence at the proper level. The alternating form also suffered a slight loss of turbulence intensity, but a much milder loss than the rotation form. The same trend is apparent on the 96^3 grid, but the actual differences are

less. The rotation form is still unsatisfactory, but now the alternating form is quite acceptable. On the 128^3 grid, even the rotation form now appears adequate. Clearly, all the (stable) forms of the momentum equation converge to the same result. However, the absolute error level is much larger for the (collocation) rotation form.

In the presence of physical boundaries, the ideal conservation laws still hold with the boundary fluxes taken into account. However, for the requisite expansions in Chebyshev polynomials the Galerkin projection and differentiation operators do not commute, and a semi-discrete conservation law, even accounting for the boundary fluxes, does not hold precisely (see CHQZ, Sec. 4.5). Nevertheless, it holds to a sufficient degree of precision to warrant a preference for the conservation form of the semi-discrete equations. Zang (1989) demonstrated that the advantages of the skew-symmetric form extend to methods employing Chebyshev polynomials. Ronquist (1988) showed that it has advantages as well in spectral-element methods. In both these applications the CPU time is dominated by the solution of the implicit equations, so that the use of the skew-symmetric form in place of the rotation form adds a negligible cost to the calculation.

7.2 Fourier-Chebyshev Methods for Partially Periodic Problems

No-slip boundary conditions are required at walls. In directions normal to the walls, then, the flow is non-periodic and Chebyshev expansions must be used. Customarily, the diffusion term in non-periodic directions is treated implicitly to overcome an otherwise prohibitive stability restriction on the time-step. In this section we consider problems with walls in only one direction and periodic boundary conditions in the other directions. Some important applications are to plane and circular Poiseuille flow, to Taylor-Couette flow with periodic conditions in the axial direction, and to the parallel boundary layer.

The primary difficulty of algorithms for wall-bounded incompressible flows is the simultaneous enforcement of the incompressibility constraint and the no-slip boundary condition. Three approaches are in common use: splitting methods, coupled methods, and Petrov-Galerkin methods.

The incompressibility constraint is most easily but least rigorously satisfied in splitting methods. Let the spatial domain be denoted by Ω and the walls by $\partial\Omega$. Modern splitting methods take the form of an advection-diffusion step,

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}\cdot\nabla\mathbf{u} - \nu\Delta\mathbf{u} = \mathbf{f}, \qquad in\ \Omega, \tag{209}$$

$$\mathbf{u} = \mathbf{g}, \qquad on\ \partial\Omega, \tag{210}$$

followed by a pressure correction step,

$$\frac{\partial \mathbf{u}}{\partial t} + \nabla p = 0, \qquad in\ \Omega, \tag{211}$$

$$\nabla\cdot\mathbf{u} = 0, \qquad in\ \Omega, \tag{212}$$

$$\mathbf{u}\cdot\hat{n} = 0, \qquad on\ \partial\Omega, \tag{213}$$

$$(\frac{\partial \mathbf{u}}{\partial t} + \nabla p)\cdot\hat{\tau} = 0, \qquad on\ \partial\Omega, \tag{214}$$

where $\hat{n}$ and $\hat{\tau}$ denote unit vectors normal to and tangential to the wall, respectively. (No-slip boundary conditions on the tangential velocity may not be imposed in the pressure correction step; doing so produces an ill-posed problem.) At the end of the full step the velocity will be divergence-free, but there will be a slip velocity at the walls.

If the intermediate boundary conditions are simply $\mathbf{g} = 0$, then this slip velocity will be $O(\Delta t)$. The boundary conditions (210) on the intermediate step can be chosen to reduce the final slip velocity to $O(\Delta t^2)$, via

$$\begin{aligned} \mathbf{g} \cdot \hat{n} &= 0, \\ \mathbf{g} \cdot \hat{\tau} &= \Delta t \, \nabla p \cdot \hat{\tau}, \end{aligned} \tag{215}$$

or to $O(\Delta t^3)$ via

$$\begin{aligned} \mathbf{g} \cdot \hat{n} &= 0, \\ \mathbf{g} \cdot \hat{\tau} &= (\Delta t \, \nabla p + \Delta t^2 \, \frac{\partial}{\partial t} \nabla p) \cdot \hat{\tau}. \end{aligned} \tag{216}$$

(These formulas presume that a backward Euler time-discretization is applied to (211) and (214).)

The big advantage of these splitting techniques is that they require the solution of only Poisson equations (for the pressure) or Helmholtz equations (from a Crank-Nicolson discretization of the viscous term). These positive-definite, scalar equations are much easier to solve numerically than the indefinite, coupled equations that arise in unsplit methods. The implicit equations are diagonal in the periodic directions after a Fourier transform is applied, so that if there is only a single non-periodic direction, then one must solve just a set of independent one-dimensional problems.

If no-slip boundary conditions are imposed on the intermediate step, i.e. $\mathbf{g} = 0$, instead of using (215) or (216), then the $O(\Delta t)$ slip velocity corresponds to splitting errors which are $O(1)$ near the boundary for the normal pressure gradient and diffusion terms. They appear to cause no serious errors in the channel flow problem. However, the boundary errors produce serious inaccuracies in Taylor-Couette flow, which is shear, rather than pressure-gradient, driven. No such difficulties have yet surfaced when appropriate intermediate boundary conditions are employed. See CHQZ, Sec. 7.3.2 for further discussion.

One way to avoid the splitting errors is to integrate the incompressible Navier-Stokes equations in a single step that couples the divergence-free constraint with the momentum equations. The numerical difficulty of coupled methods is that one must invert a larger set of equations (it involves the pressure as well as the three velocity components), which is indefinite. The influence-matrix technique enables this solution to be achieved in the constant viscosity case by a series of scalar Poisson and Helmholtz solutions. This is discussed in great detail for channel flow in CHQZ, Secs. 7.3.1 and 11.3.3, and a recent general formulation has been provided by Tuckerman (1989a). Ehrenstein (1988) has derived the eigenvalue spectrum of the influence-matrix method for the two-dimensional streamfunction-vorticity formulation of the Stokes equations, and he has used these results to analyze the stability of the fully discrete approximation. Lacroix, Peyret, and Pulicani (1987) have used the influence-matrix method in conjunction with domain decomposition in streamfunction-vorticity variables. The Uzawa scheme, which is discussed in Sec. 7.3, is another option for avoiding splitting errors.

Many of the numerical problems caused by the incompressibility constraint can be avoided by an expansion in functions which are divergence-free. Such schemes are characterized as Petrov-Galerkin methods since both the trial and test functions satisfy the boundary conditions, but they are different sets of functions. Thus far, spectral Petrov-Galerkin methods have been applied to pipe flow, plane channel flow, curved channel flow, and the parallel boundary layer. The efficiency of these methods is dependent upon the bandwidth of the matrices which arise from the implicit treatment of the viscous terms. In the examples cited above, the bandwidth is quite small, roughly of order 10. This requirement has dictated the use of special Jacobi

polynomials rather than Chebyshev ones in pipe and boundary layer flow. As a consequence, transform methods are not applicable in the non-periodic direction. Moreover, in even slightly more general cases, the matrices can be completely full. A description of this method is available in CHQZ, Sec. 7.3.3, and its rigorous numerical analysis (for the steady problem) is presented by Pasquarelli, Quarteroni, and Sacchi-Landriani (1987).

7.3 Chebyshev Methods for Fully Non-periodic Problems

Spectral methods for problems which are non-periodic in more than one direction tend to be more expensive than the methods discussed previously in this chapter due to the added expense of solving implicit equations with multiple non-periodic directions. (Recall the discussion in Chapter 5.) For simplicity the discussion in this section will focus on two-dimensional problems.

The first consideration in formulating such a method is the choice of the collocation mesh. The various options are illustrated in Figure 17. A non-staggered mesh, in which both velocity components as well as the pressure are defined at the Gauss-Lobatto points, given by (55), is of course the most straightforward to implement. However, such a discretization admits spurious pressure modes, and generally requires some form of pressure boundary condition. Spurious modes for the pressure are characterized as those non-constant pressures which have vanishing x-derivatives at all the interior nodes for u and vanishing y-derivatives at all the interior nodes for v. Such modes have no effect upon the velocity since the interior velocity nodes are where the momentum equations are enforced.

Schemes with velocities collocated on the usual Gauss-Lobatto points, but with the pressure defined on the Gauss points, given by (59), are said to use a half-staggered mesh. Although this discretization circumvents the requirement for pressure boundary conditions, it does admit one spurious mode, and has seen little use.

In a fully-staggered mesh, the pressure is collocated on the Gauss points, whereas the velocity components use a mesh which is a tensor product of the usual Gauss-Lobatto points and the extended Gauss points. The latter consist of the Gauss points (59) plus the boundary points -1 and $+1$. The u velocity component is defined for the Gauss-Lobatto points in x and the extended Gauss points in y, and the v component is defined for the reverse combination. CHQZ, Sec. 7.4.1 discuss one method for carrying out differentiation on the extended Gauss points. Although cumbersome to implement, the fully-staggered mesh has the advantage of eliminating spurious modes.

Spurious modes have been of more concern to theorists than to practitioners of numerical methods for incompressible flow. They are distasteful to theorists for they prevent the usual type of error bounds to be placed on the approximate pressure. Nevertheless, the admission of spurious modes need not disqualify a numerical scheme. All that's required is that they be filtered from the pressure. Morchoisne and his colleagues have identified the spurious modes for both two- and three-dimensional problems and have described how they may be filtered from the solution. Bernardi, Canuto, Maday, and Metivet (1988) provide a recent discussion of both the practical details of the filtering procedure and the rigorous numerical analysis of spectral methods on non-staggered grids. Additional discussion is available in CHQZ, Secs. 7.4.1 and 11.3.1.

Splitting methods, particularly on a non-staggered mesh, are straightforward extensions of the algorithm discussed in Sec. 7.2. The equations (211) -(214) for the pressure correction step may be combined to form a Poisson equation for the pressure,

$$\Delta p = \frac{1}{\Delta t} \nabla \cdot \mathbf{u}, \qquad in\ \Omega, \tag{217}$$

where $\mathbf{u}$ is understood to be the velocity resulting from the first, advection-diffusion step. (Although we have written this Poisson equation in continuous form, it should actually be constructed directly from the discrete forms of (211) - (214). CHQZ, Sec. 7.3.2 describe this process in detail in the context of a partially periodic problem.) Once the pressure is found from (217), the velocity is updated via (211), (213), and (214). For the non-staggered mesh, (217) requires a pressure boundary condition. Streett and Hussaini (1989) have used the homogeneous Neumann condition. This is justified on the grounds of consistency with the normal component of (211) at the boundary.

Coupled methods, of course, avoid splitting errors. The influence-matrix technique is one approach to solving the resulting implicit equations. See Tuckerman (1989a) for a thorough discussion of influence-matrix methods for both tau and collocation discretizations of the primitive variable formulation, and see Ehrenstein and Peyret (1987) for such methods in streamfunction-vorticity variables. Influence-matrix methods are subject to an ambiguity concerning the treatment of the points at the corners of the domain. This issue is discussed in the above two references.

An alternative coupled approach has been applied to spectral and spectral domain decomposition methods recently and is particularly efficient for the steady problem. The discrete version of (197) - (198) may be written

$$-\mathcal{L}Q + \mathcal{G}P = F, \tag{218}$$

$$\mathcal{D}Q = 0, \tag{219}$$

where $\mathcal{L}$ denotes ν times the discrete vector Laplacian, $\mathcal{G}$ the discrete gradient operator, $\mathcal{D}$ the discrete divergence operator, and $Q = (U,V)$. Solving (218) for Q and substituting into (219) yields

$$(\mathcal{D}\mathcal{L}^{-1}\mathcal{G})P = (\mathcal{D}\mathcal{L}^{-1})F. \tag{220}$$

This general approach is known as the Uzawa strategy.

The solution scheme for (220) is motivated by the observation that $\mathcal{D}$ and $\mathcal{G}$ are first-order operators, whereas $\mathcal{L}$ is second-order. The Uzawa operator, $\mathcal{D}L^{-1}\mathcal{G}$, thus is essentially zeroth-order, aside from the effect of boundary conditions which are enforced in the operator $\mathcal{L}$. As a result the eigenvalues of the Uzawa operator are tightly clustered, and the condition number of the problem is quite small. Apparently, the eigenvalues of the discrete Uzawa operator in (220) are clustered more tightly for spectral discretizations than for the finite-element discretizations for which the algorithm was originally devised.

Conjugate gradient relaxation is appropriate for the iterative solution of (220); it yields rapid convergence for problems with limited eigenvalue spread. (Actually, since the operator is not symmetric the relaxation scheme is more appropriately referred to as a truncated conjugate gradient iteration.) Of course, for each of these conjugate gradient iterations, two scalar Helmholtz problems must be inverted. The typical spectral radius observed for the conjugate gradient iteration in spectral applications of the steady Uzawa algorithm is between .05 and .15, with only an extremely weak degradation with mesh refinement.

Streett and Hussaini (1989) have employed the spectral Uzawa algorithm on a fully staggered mesh to solve for steady states in three-dimensional nonlinear Taylor-Couette flow. This problem has one periodic and two non-periodic directions. The nonlinear convection term is incorporated by a Newton iteration, with each step involving the solution of a Stokes problem. By applying a Fourier transform in the periodic direction, the Stokes problem de-couples into a set of independent two-dimensional problems. See Streett and Hussaini (1989) for the details of the use of the fully staggered mesh. The Uzawa algorithm has also been employed in numerous

spectral-element calculations of the steady, incompressible Navier-Stokes equations (Maday and Patera 1989).

Unfortunately, for the unsteady Stokes problem (195) - (196) the Uzawa algorithm is much less efficient. Eq. (218) becomes

$$-(\mathcal{L}-\lambda I)Q+\mathcal{G}P=F, \tag{221}$$

where $\lambda \sim 1/\Delta t$, and (220) generalizes to

$$[\mathcal{D}(\mathcal{L}-\lambda I)^{-1}\mathcal{G}]P=[\mathcal{D}(\mathcal{L}-\lambda I)^{-1}]F. \tag{222}$$

For moderate to small time steps ($\Delta t \ll 1$), the unsteady Uzawa operator, $\mathcal{D}(\mathcal{L}-\lambda I)^{-1}\mathcal{G}$, is essentially second-order. Consequently, its eigenvalues are widely spread; in fact, its condition number scales as N^4. In practice, the truncated conjugate gradient method fails to converge for all but the smallest problems.

Streett and Hussaini (1989) developed a multigrid technique to alleviate this difficulty. Eq. (221) is the basic fine-grid equation, which in multigrid notation is written as (187). Their relaxation scheme was conjugate gradient iteration, and standard multigrid methodology was used to take advantage of rapid smoothing on coarse grids – see Section 5.2.5. Two relaxations were typically performed at each level of the fixed V-cycle; relaxations were performed only during the grid-coarsening phase of the cycle. (In terms of Figure 14, $N_d = 2$ and $N_u = 0$.) Natural spectral interpolation operators were used for restriction and prolongation. On the coarsest mesh, a direct solution of the Uzawa operator was used; the operator was generated, inverted and stored in a preprocessing step.

The multigrid/conjugate gradient algorithm is quite robust. For solution of the steady problem, the multigrid algorithm requires about the same amount of machine time as the single-grid/conjugate gradient algorithm; the improvement in convergence rate just offsets the overhead of the algorithm. For the unsteady problem, a small degradation occurs in the convergence rate of the multigrid method with decreasing Δt relative to the steady case, but the machine time to reach a fixed residual level is roughly constant over a wide range of time-steps. (Recall that the single-grid scheme is not a viable option at all for the unsteady problem.) This is because the initial residual levels are lower for the case of smaller Δt (with the solution at the previous time-step used for the initial guess).

The staggered-mesh multigrid/conjugate gradient coupled scheme is considerably more expensive per time-step than the non-staggered splitting algorithm. On equivalent grids, the coupled scheme takes about 50-100 times longer per time-step than the split scheme. The split scheme is much more efficient per time-step because it requires only the solution of 11 scalar uncoupled Helmholtz/Poisson equations, and these can be solved very rapidly by the matrix-diagonalization method (see Sec. 5.1). However, the coupled scheme is somewhat more accurate than the split scheme, due to the splitting errors and imposed pressure boundary conditions of the latter. To reduce the magnitude of these errors in the split scheme, the time-step chosen must be between 10 and 100 times smaller than that required purely by accuracy of the temporal discretization. Thus neither scheme is clearly superior; comparison of these schemes for transition simulations is ongoing.

One spectral domain decomposition application of particular interest incorporates a non-reflecting outflow boundary treatment (Streett and Macaraeg 1989). The lack of viable non-reflecting outflow boundary conditions for the Navier-Stokes equations, ones capable of benignly passing large unsteady disturbances out of the computational domain, i.e., without disrupting the flow upstream of the outflow boundary, has limited the progress that large-scale numerical simulations have made in investigations of transition in shear flows. While the knowledge gained

through temporal simulations has been significant (see Zang and Hussaini 1987 for a recent review), the streamwise periodicity assumptions underlying these simulations limit their usefulness as precise, quantitative, predictive tools for spatially-developing flows. Moreover, the neglect of non-parallel effects in temporal simulations prevents them from addressing such fundamentally spatial effects as receptivity and roughness.

The outflow boundary treatment of Streett and Macaraeg (1989) first breaks the computational domain into two regions, the "primary" domain, which is the region of physical interest, and a "buffer" domain which is appended onto the downstream boundary of the primary domain. This scheme is motivated by the recognition that, for incompressible flow, the spatial ellipticity of the Navier-Stokes equations has two sources – the viscous terms and the pressure field. The upstream influence influence of streamwise diffusion is ameliorated by multiplying the streamwise viscous term in the buffer domain by a function which smoothly reduces this term to zero by the end of this domain. The viscous term remains unmodified in the primary domain. Additionally, the convective velocity in the nonlinear advection term, which in the primary domain is the sum of mean and disturbance velocities, is smoothly modified in the buffer domain so that only the (strictly out-going) mean velocity appears at the buffer domain outflow boundary. The upstream influence of the pressure field is suppressed by attenuating the forcing function of the pressure Poisson equation in the buffer domain. A homogeneous Neumann boundary condition is utilized at outflow. The interface between the primary domain and the buffer domain is treated in a standard domain decomposition manner.

Although a matrix-diagonalization direct solver can be used to invert the discrete Helmholtz equations resulting from the usual viscous terms, inversion of the modified viscous operator in the buffer domain apparently must be performed iteratively. (The eigenvalues of the modified viscous operator are complex and some of them are nearly zero.) Preconditioned minimum residual iteration is used, with the appropriate second-order central finite-difference operator used as a preconditioner. This operator is inverted using LSOR. Both iteration schemes converge rapidly, due in part to the scalar term of $O(1/\Delta t)$ from the discrete time derivative yielding overwhelming diagonal dominance of the system.

Disturbances from inconsistencies in pressure boundary conditions are generally confined to narrow regions near the boundary; in this application, the region of disturbance falls well inside the buffer domain.

This outflow boundary condition has been verified against the predictions of spatial, linear stability theory for plane Poiseuille flow in a channel with the streamwise length of the primary domain 20 times the channel half-height; the buffer domain is the same length. The Reynolds number for the simulation is 1500 based on channel half-height. The numerical simulation employed 33 points in the streamwise direction for the primary domain and 21 points in the normal direction. At the inflow boundary a periodic disturbance is imposed which is the least stable eigenmode of the corresponding linear stability problem; the maximum velocity perturbation of this inflow condition is roughly 0.01% of the mean center-line velocity. The temporal period of the disturbance is about 19.25. Figures 18 and 19 illustrate the performance of the outflow technique. Figure 18 displays contours of disturbance vorticity at equally-spaced time intervals. The leading wave of the disturbance clearly transmits from the primary domain into the buffer domain without noticeable upstream disruption. Additionally, the leading wave has reached the end of the buffer domain by the last frame of the sequence. Figure 19 shows a plot of surface disturbance vorticity, taken at about 24 cycles after the start of the simulation. The discretization used for the primary domain in this computation was 45×33. The results of the simulation are compared in the figure against the "exact" result from linear theory. As can be seen, the velocity of propagation and the amplitude decay are predicted well, except perhaps near the endpoints of the plot, or the corners of the domain.

7.4 Compressible Problems

Spectral numerical simulations of compressible, viscous flow have begun to emerge. Erlebacher and Hussaini (1989) studied transition on a flat plate boundary layer in supersonic flow. They used a straightforward, fully explicit scheme. Gauthier (1988) and Dang and Loisel (1987) have developed semi-implicit spectral collocation algorithms for compressible flow and have applied them to two- and three-dimensional problems in compressible convection. Simulation of homogeneous compressible turbulence (CHQZ. Sec. 8.4) are becoming quite common.

Another method currently being examined for the solution of the time-dependent compressible Navier-Stokes equations is the fully-implicit technique of Streett and Macaraeg (in preparation). The main idea here is to formulate the time advancement in a backward-Euler form, then to solve the resulting equations iteratively. The equation set we wish to solve at any time level is

$$\beta_1 U^{n+1} + L(U^{n+1}) = \sum_{k=0}^{p} \beta_k U^{n-k}, \tag{223}$$

where the β_k's are coefficients of the $(p+1)^{th}$-order backward time difference. Currently, second- to fifth-order differencing is used. Eq. (223) is solved by first Newton-linearizing about a provisional value, U^*:

$$[\beta_1 U + A(U^*)]V^m - \sum_{k=0}^{p} \beta_k U^{n-k} = R^m, \tag{224}$$

then driving R^m to zero using preconditioned iteration. The provisional value U^* is updated several times to reduce linearization errors. When this process converges, the value V^m is taken as U^{n+1}. Since the operator, A, is dominated by first derivatives, the preconditioning is carried out using central finite-differences on a refined mesh, as suggested in Section 5.2.2. The multi-dimensional preconditioner is inverted iteratively, using an ADI scheme related to the well-known Beam and Warming algorithm; fourth-order artificial viscosity is used in the preconditioner to remove odd-even decoupling of the central differences, and the ADI scheme uses pentadiagonal solvers to include the artificial viscosity terms. Early tests indicate that the preconditioned iteration has a spectral radius of about .2 on reasonable meshes, and that the overall time-stepping scheme is stable to Courant numbers of several hundred.

8 Other Applications

Spectral methods have been developed and applied to many other problems in viscous flow. This chapter will summarize some of the recent developments.

CHQZ, Sec. 6.3 describe the use of spectral methods for solving the incompressible boundary-layer equations. This technique has been extended to the compressible boundary-layer equations, and has been used in the work of Macaraeg, Streett, and Hussaini (1988) in furnishing highly accurate self-similar mean flows for flat plate boundary layers and for free shear layers, both in supersonic flow.

The linear stability equations of incompressible flow have been a very fruitful application for spectral methods (CHQZ, Sec. 6.4). Some interesting algorithmic issues are raised by the Orr-Sommerfeld equation, for it is a fourth-order equation and requires two boundary conditions at each endpoint. Bernardi and Maday (1988) have analyzed the alternative collocation schemes for fourth-order problems and have deduced which are the optimal formulations. They have studied domain decomposition versions as well. Phillips and Karageorghis (1989) present some results for a spectral domain decomposition method for a two-dimensional fourth-order problem.

A number of researchers have developed spectral codes for solving the compressible linear stability equations. Just within our group at NASA Langley, different spectral codes have been developed by Ng (1989), Macaraeg, Streett, and Hussaini (1988), Malik(1988), and Erlebacher, Herbert, and Hussaini (in preparation). Moreover, A. Thunn of the University of Stuttgart (1988, private communication) has developed a similar spectral code. These have been applied to jet, boundary-layer, free shear-layer, and plane Couette compressible stability problems. Macaraeg, Streett, and Hussaini (1988) utilized a staggered mesh in their algorithm. This avoids the need for a pressure boundary condition. Velocities and temperature were located on the Gauss-Lobatto points where the momentum and energy equations were collocated, and the pressure was located on the Gauss points where the continuity equation was enforced. A multi-domain version of the scheme, using the integral flux balance condition, was also implemented. It proved extremely useful in improving the efficiency of the spectral method, particularly in resolving the highly oscillatory modes that exist in supersonic flow. They presented comparisons with a second-order finite-difference scheme, Malik (1988) compared with a fourth-order method, and Thunn has made comparisons with up to eighth-order finite-difference methods.

Spectral methods continue to be quite useful for investigations of secondary instability (CHQZ, Sec. 6.4). The reviews by Bayly, Orszag, and Herbert (1988) and Herbert (1988) provide a recent status report. Ng (1989) has used spectral methods to study the secondary instability of subsonic jet flows.

Problems in reacting flow continue to challenge numerical methods. Guillard and Peyret (1986) describe a spectral method for a premixed flame problem in which a prescribed mapping is used to increase the resolution of the flame front. Bayliss, Gottlieb, Matkowsky, and Minkoff (1987) have pursued adaptive grid strategies for spectral methods. Augenbaum (1988) has studied adaptive solutions of the one-dimensional wave equation and the Burgers equation. Macaraeg and Streett (1988) have used a domain decomposition method to great advantage in studying non-equilibrium chemistry effects at hypersonic speeds.

Several serious investigations have been made on the suitability of spectral methods for local memory parallel processors. These are, of course, more challenging for spectral methods than shared memory parallel processors due to the global nature of the approximation. Erlebacher, Bokhari, and Hussaini (1989) implemented on the FLEX/32 a three-dimensional explicit spectral code for simulating transition on a flat plate in supersonic flow. They achieved over 95% efficiency with as many as 16 processors. Fischer, Ho, Karniadakis, Ronquist, and Patera (1988) have run incompressible spectral-element simulations on a 16-processor Intel iPSC/2 hypercube at over 75% efficiency. Pelz (1989) has achieved over 80% efficiency for a 128^3 simulation of isotropic turbulence using a pseudospectral method on a 1024-processor NCUBE. Finally, Tomboulian, Streett, and Macaraeg (1988) have implemented a three-dimensional spectral collocation algorithm for channel flow on a 16,384-processor Connection Machine 2, and achieved performance comparable to that obtained for the same algorithm on a single-processor of a Cray 2. Of all these applications only Pelz (1989) used FFT's in his implementation. The others used matrix-multiplies to accomplish the Fourier/Chebyshev transforms and derivatives. The key to most of these implementations was the ability to perform fast matrix transposes and/or Fast Fourier Transforms in a manner that kept the calculation CPU-bound rather than communication-bound.

9 Trends

A cursory analysis of the post-1985 citations in the bibliography indicates where current research on the development and analysis of spectral methods has been concentrated. Certainly the work on spectral domain decomposition methods stands out. Perhaps it has been a matter of necessity, but there is little question that the most innovative and demanding work on domain decomposition techniques in general is now being performed for spectral methods. The bibliography also suggests that there has been a resurgence of contributions to iterative methods, particularly spectral multigrid methods. Taken together these two trends portend an even broader range of application for spectral methods in the 1990's. Moreover, recent achievements in implementing spectral methods on massively parallel computers indicate that spectral methods will keep pace with the on-going architectural revolution in computer hardware.

We have made relatively few comments in these notes on the great strides that have been made during the 1980's in the numerical analysis of spectral methods. The state-of-the-art has progressed from the ability at the start of this decade to analyze rigorously little more than linear model problems to the imminent capability of furnishing rigorous error estimates for spectral approximations to the full unsteady Navier-Stokes equations with general boundary conditions.

10 Bibliography

Augenbaum, J. M. (1988): An adaptive pseudospectral method for discontinuous problems, ICASE Report No. 88-54 (submitted to **Appl. Numer. Math.**).

Bayliss, A.; Gottlieb, D.; Matkowsky, B. J.; and Minkoff, M. (1987): An adaptive pseudo-spectral method for reaction diffusion problems, ICASE Report No. 87-67 (submitted **J. Comput. Phys.**).

Bayly, B. J.; Orszag, S. A.; and Herbert, T. (1988): Instability mechanisms in shear-flow transition, **Ann. Rev. Fluid Mech.**, Vol. 20, pp. 359-391.

Bernardi, C.; Canuto, C.; and Maday, Y. (1988): Generalized inf-sup conditions for Chebyshev spectral approximation of the Stokes problem, **SIAM J. Numer. Anal.**, Vol. 25, pp. 1237-1271.

Bernardi, C; Canuto, C.; Maday, Y.; and Metivet, B. (1988): Single-grid spectral collocation for the Navier-Stokes equations, ICASE Report No. 88-69 (submitted to **IMA J. Numer. Anal.**).

Bernardi, C.; Debit, N.; and Maday, Y. (1987): Coupling finite element and spectral methods: first results, ICASE Report No. 87-70 (submitted to **Math. Comp.**).

Bernardi, C.; and Maday, Y. (1989): Properties of some weighted Sobolev spaces, and application to spectral approximations, **SIAM J. Numer. Anal.**, in press.

Bernardi, C.; and Maday, Y. (1989): Approximation results for spectral methods with domain decomposition, **Appl. Numer. Math.**, in press.

Bernardi, C., and Maday, Y. (1988): Spectral methods for the approximation of fourth-order problems: application to the Stokes and Navier-Stokes equations, **Computers & Structures**, Vol. 30, pp. 205-216.

Bernardi, C.; Maday, Y.; and Sacchi-Landriani, G. (1989): Non-conforming matching conditions for coupling spectral and finite element methods, **Appl. Numer. Math.**, in press.

Bisseling, R. H.; and Kosloff, R. (1988): Optimal choice of grid points in multidimensional pseudospectral Fourier methods, **J. Comput. Phys.**, Vol. 76, pp. 243-262.

Blinova, E. N. (1944): Hydrodynamic theory of pressure and temperature waves and center of action of the atmosphere. Translation No. 113, Regional Control Office, Second Weather Region, Patterson Field.

Boyd, J. P. (1987): Exponentially convergent Fourier-Chebyshev quadrature schemes on bounded and infinite intervals, **J. Sci. Comput.**, Vol. 2, pp. 99-121.

Brandt, A.; Fulton, S. R.; and Taylor, G. D. (1985): Improved spectral multigrid methods for periodic elliptic problems, **J. Comput. Phys.**, Vol. 58, pp. 96-112.

Briggs, W. L.; Hart, L. B.; Sweet, R. A.; and O'Gallagher, A. (1987): Multiprocessor FFT methods, **SIAM J. Sci. Stat. Comput.**, Vol. 8, pp. s27-s42.

Cai, W.; Gottlieb, D.; and Shu, C. (1988): Non-oscillatory spectral Fourier methods for shock wave calculations, ICASE Report No. 88-37 (submitted to **Math. Comp.**).

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

Canuto, C.; and Pietra, P. (1987): Boundary and interface conditions within a finite element preconditioner for spectral methods, Report No. 553, Pavia University (submitted to **J. Comput. Phys.**).

Chan, T. F.; and Goovaerts, D. (1989): Schur complement domain decomposition algorithms for spectral methods, **Appl. Numer. Math.**, in press.

Dang, K.; and Loisel, P. (1987): Direct Simulation of Viscous Compressible Transitional Flows, ONERA Report No. 1987-23.

Davies, A. R.; and Anderssen, R. S. (1986): Improved estimates of statistical regularization parameters in Fourier differentiation and smoothing, **Numer. Math.**, Vol. 48, pp. 671-697.

Davies, A. R.; Karageorghis, A.; and Phillips, T. N. (1988): Spectral Galerkin methods for the primary two-point boundary value problem in modelling viscoelastic flows, **Int. J. Numer. Methods Eng.**, Vol. 26, pp. 647-662.

Demaret, P.; Deville, M. O.; and Schneidesch, C. (1989): Thermal convection solutions by Chebyshev pseudospectral multi-domain decomposition and finite element preconditioning, **Appl. Numer. Math.**, in press.

Dendy, J. E. (1987): Two multigrid methods for three-dimensional problems with discontinuous and anisotropic coefficients, **SIAM J. Sci. Stat. Comput.**, Vol. 8, pp. 673-685.

Deville, M.; and Mund, E. (1989): Finite element preconditioning for pseudospectral solutions of elliptic problems, **SIAM J. Sci. Stat. Comput.**, in press.

Dubiner, M. (1987): Asymptotic analysis of spectral methods, **J. Sci. Comput.**, Vol. 2, pp. 3-31.

Dubois, M.; Greenbaum, A.; and Rodrique, G. (1979): Approximating the inverse of a matrix for use in iterative algorithms on vector processors, **Computing**, Vol. 22, pp. 257-268.

Ehrenstein, U. (1988): Stability of a Chebyshev-Fourier Approximation for the Stokes Equations, Prepublication No. 198, Universite de Nice.

Ehrenstein, U.; and Peyret, R. (1987): A Chebyshev-collocation Method for the Navier-Stokes Equations with Application to Double-Diffusive Convection, Prepublication No. 169, Universite de Nice.

Eliasen, E.; Machenauer, E.; and Rasmussen, E. (1970): On a Numerical Method for Integration of the Hydrodynamical Equations with a Spectral Representation of the Horizontal Fields, Report No. 2, Department of Meteorology, Copenhagen University.

Erlebacher, G.; and Hussaini, M. Y. (1989): Nonlinear evolution of a second mode wave in supersonic boundary layers, **Appl. Numer. Math.**, in press.

Erlebacher, G.; Bokhari, S. H.; and Hussaini, M. Y. (1989): Three-dimensional compressible transition on a 20 precessor FLEX/32 Multicomputer, **AIAA J.**, in press.

Erlebacher, G.; Zang, T. A.; and Hussaini, M. Y. (1987): Spectral multigrid methods for the numerical simulation of turbulence, in **Multigrid Methods, Theory, Applications, and Supercomputing**, ed. by S. F. McCormick, pp. 177-194, Marcel Dekker, New York.

Fishelov, D. (1988a): Spectral methods for the small disturbance equation of transonic flows, **SIAM J. Sci. Stat. Comput.**, Vol. 9, pp. 232-251.

Fishelov, D. (1988b): The spectrum and the stability of the Chebyshev collocation operator for transonic flow, **Math. Comp.**, Vol. 51, pp. 559-579.

Fischer, P. F.; Ho, L. W.; Karniadakis, G. E.; Ronquist, E. M.; and Patera, A. T. (1988): Recent advances in parallel spectral element simulation of unsteady incompressible flows, **Computers & Structures**, Vol. 30, pp. 217-231.

Funaro, D. (1988a): Computing the inverse of the Chebyshev collocation derivative, **SIAM J. Sci. Stat. Comput.**, Vol. 9, pp. 1050-1057.

Funaro, D. (1988b): Domain decomposition methods for pseudospectral approximations. I. Second order equations in one dimension, **Numer. Math.**, Vol. 52, pp. 329-344.

Funaro, D.; and Gottlieb, D. (1988): A new method of imposing boundary conditions in pseudospectral approximations of hyperbolic equations, **Math. Comp.**, Vol. 51, pp. 519-613.

Gastaldi, F.; and Quarteroni, A. (1989): On the coupling of hyperbolic and parabolic systems: analytical and numerical approach, ICASE Report No. 88-55 (submitted to **Appl. Numer. Math.**).

Gauthier, S. (1988): A spectral collocation method for two-dimensional compressible convection, **J. Comput. Phys.**, Vol. 75, pp. 217-235.

Gottlieb, D.; and Orszag, S. A. (1977): **Numerical Analysis of Spectral Methods: Theory and Applications**, SIAM-CBMS, Philadelphia.

Gottlieb, D.; and Hirsh, R. S. (1988): Parallel pseudospectral domain decomposition techniques, ICASE Report No. 88-15 (submitted to **SIAM J. Sci. Stat. Comput.**).

Guillard, H.; and Peyret, R. (1986): On the Use of Spectral Methods for the Numerical Solution of Stiff Problems, Prepublication No. 108, University de Nice.

Guo, B. (1983): The convergence of the spectral scheme for solving two-dimensional vorticity equation, **J. Comp. Math.**, Vol. 1, pp. 353-362.

Guo, B. (1985a): Error estimations of the spectral method for solving K.D.V.-Burgers equation, **Acta Math. Sin.**, Vol. 28, pp. 1-15.

Guo, B. (1985b): Spectral method for solving Navier-Stokes equations, **Scientia Sin.**, Vol. 28, pp. 1139-1153.

Guo, G.; and Cao, W. (1988): The Fourier pseudospectral method with a restrain operator for the RLW equation, **J. Comput. Phys.**, Vol. 74, pp. 110-126.

Guo, B.; and Ma, H. (1987): The Fourier pseudo-spectral method for three-dimensional vorticity equations, **Acta Math. Appl. Sin.**, Vol. 3, pp. 296-309.

Guo, B.; and Ma, H. (1989): Fourier Pseudo-spectral method for Burgers equation, **Acta Math. Appl. Sin.**, in press.

Haidvogel, D. B.; and Zang, T. A. (1979): The accurate solution of Poisson's equation by expansion in Chebyshev polynomials, **J. Comput. Phys.**, Vol. 30, pp. 167-180.

Heinrichs, W. (1988a): Line relaxation for spectral multigrid methods, **J. Comput. Phys.**, Vol. 77, pp. 166-182.

Heinrichs, W. (1988b): Multigrid methods for combined finite difference and Fourier problems, **J. Comput. Phys.**, Vol. 78, pp. 424-436.

Herbert, T. (1988): Secondary instability of boundary layers, **Ann. Rev. Fluid Mech.**, Vol. 20, pp. 487-526.

Horiuti, K. (1987): Comparison of conservative and rotational forms in large eddy simulation of turbulent channel flow, **J. Comput. Phys.**, Vol. 71, pp. 343-370.

Hussaini, M. Y. (1988): Computational fluid dynamics: a personal view, in **Proc. 11th Inter. Conf. Num. Meth. in Fluid Dynamics**, ed. by D. L. Dwoyer and R. G. Voigt, Springer-Verlag, New York.

Hussaini, M. Y.; Kopriva, D. A.; and Patera, A. T. (1989): Spectral collocation methods, **Appl. Numer. Math.**, in press.

Hussaini, M. Y.; Kopriva, D. A.; Salas, M.; Zang, T. A. (1985a): Spectral methods for the Euler equations. Part 1. Fourier methods and shock-fitting, **AIAA J.**, Vol. 23, pp. 64-70.

Hussaini, M. Y.; Kopriva, D. A.; Salas, M.; Zang, T. A. (1985b): Spectral methods for the Euler equations. Part 2. Chebyshev methods and shock-fitting, **AIAA J.**, Vol. 23, pp. 234-240.

Hussaini, M. Y.; Streett, C. L.; and Zang, T. A. (1983): Spectral methods for partial differential equations, in **Trans. 1st Army Conf. Applied Mathmatics and Computing**, pp. 883-925.

Karageorghis, A.; and Phillips, T. N.(1989): Spectral collocation methods for Stokes flow in contraction geometries and unbounded domains, **J. Comput. Phys.**, in press.

Karageorghis, A.; Phillips, T. N.; and Davies, A. R. (1988): Spectral collocation methods for the primary two-point boundary value problem in modelling viscoelastic flows, **Int. J. Numer. Methods Eng.**, Vol. 26, pp. 805-813.

Karniadakis, G. E. (1989): Spectral element simulations of laminar and turbulent flows in complex geometries, **Appl. Numer. Math.**, in press.

Khorrami, M. R.; Malik, M. R.; and Ash, R. L. (1989): Application of spectral collocation techniques to the stability of swirling flows, **J. Comput. Phys.**, in press.

Kopriva, D. A. (1989a): Computation of hyperbolic equations on complicated domains with patched and overset Chebyshev grids, **SIAM J. Sci. Stat. Comput.**, Vol. 10, pp. 120-132.

Kopriva, D. A. (1989b): Domain decomposition with both spectral and finite difference methods for the accurate computation of flows with shocks, **Appl. Numer. Math.**, in press.

Lacroix, J. M.; Peyret, R.; and Pulicani, J. P. (1987): A pseudospectral multi-domain method for the Navier-Stokes equations with application to double-diffusive convection, in **Proc. 7th GAMM Conference on Numerical Methods in Fluid Dynamics**, Vieweg, Braunschweig.

Liu, K. M.; and Ortiz, E. L. (1988): Tau method approximation of differential eigenvalue problems where the spectral parameter enters nonlinearly, **J. Comput. Phys.**, Vol. 672, pp. 299-310.

Ma, H.; and Guo, B. (1986): The Fourier pseudospectral method with a restrain operator for the Korteweg-de Vries equation, **J. Comput. Phys.**, Vol. 65, pp. 120-137.

Ma, H.; and Guo, B. (1987): The Fourier pseudospectral method for two-dimensional vorticity equations, **IMA J. Numer. Anal.**, Vol. 7, pp. 47-60.

Macaraeg, M. G.; and Streett, C. L. (1986): Improvements in spectral collocation discretization through a multiple domain technique, **Appl. Numer. Math.**, Vol. 2, pp. 95-108.

Macaraeg, M. G.; and Streett, C. L. (1988): An analysis of artificial viscosity effects on reacting flows using a spectral multi-domain technique, in **Computational Fluid Dynamics**, ed. by G. de Vahl Davis and C. Fletcher, pp. 503-514, North-Holland, Amsterdam.

Macaraeg, M. G.; and Streett, C. L.(1989): Linear stability of high-speed mixing layers, **Appl. Numer. Math.**, in press.

Macaraeg, M. G.; Streett, C. L.; and Hussaini, M. Y. (1988): A Spectral Collocation Solution to the Compressible Stability Eigenvalue Problem, NASA TP-2858.

Maday, Y.; and Patera, A. T. (1989): Spectral element methods for the incompressible Navier-Stokes equations, in **State of the Art Surveys in Computational Mechanics**, ed. by A. Noor and J. T. Oden, ASME, New York.

Maday, Y.; Mavriplis, C.; and Patera, A. T. (1988): Nonconforming mortar element methods: application to spectral discretizations, ICASE Report No. 88-59 (submitted to **Proc. 2nd Int. Conf. Domain Decomp. Methods for PDE**).

Maday, Y.; and Munoz, R. (1988): Spectral element multigrid. II. Theoretical justification, ICASE Report No. 88-31 (submitted to **J. Sci. Comput.**).

Maday, Y.; Patera, A. T.; and Ronquist, E. M. (1987): A well-posed optimal spectral element approximation for the Stokes problem, ICASE Report No. 87-48 (submitted to **SIAM J. Numer. Anal.**).

Maday, Y.; and Quarteroni, A. (1987): Error analysis for spectral approximation of the Korteweg-de Vries equation, ICASE Report No. 87-36, (submitted to **MMAN Numer. Anal.**).

Maday, Y.; and Tadmor, E. (1989): Analysis of the spectral vanishing viscosity method for periodic conservation laws, **SIAM J. Numer. Anal.**, in press.

Malik, M. R. (1988): Numerical methods for hypersonic boundary layer stability, HTC Report No. 88-6, High Technology Corporation, Hampton, VA (submitted to **J. Comput. Phys.**).

Mercier, B. (1981): Analyse Numerique des Methodes Spectrales, Note CEA-N-2278, Commissariat a l'Energie Atomique Centre d'Etude de Limeil (English translation to be published as **An Introduction to the Numerical Analysis of Spectral Methods**, in Lecture Notes in Mathamatics, Springer-Verlag, 1989).

Ng, L. L. (1989): Resonances and Instabilities in Subsonic Heated Round Jets, Ph. D. Thesis, Univ. of Arizona.

Orszag, S. A. (1969): Numerical methods for the simulation of turbulence, **Physics Fluids**, Supp. II, Vol. 12, pp. 250-257.

Pasquarelli, F.; Quarteroni, A.; and Sacchi-Landriani, G. (1987): Spectral approximations of the Stokes problem by divergence-free functions, **J. Sci. Comput.**, Vol. 2, pp. 195-226.

Pelz, R. B. (1989): Hypercube algorithms for fluid flows: the Fourier pseudospectral method, preprint (submitted to **J. Comput. Phys.**).

Phillips, T. N. (1988): Singular matched eigenfunction expansions for Stokes flow around a corner, Tech. Report ABERNA 18, University College of Wales (submitted to **IMA J. Appl. Math.**).

Phillips, T. N.; and Davies, A. R. (1988): On semi-infinite spectral elements for Poisson problems with reentrant boundary singularities, **J. Comput. Appl. Math.**, Vol. 21, pp. 173-188.

Phillips, T. N.; and Karageorghis, A. (1989): Efficient direct methods for solving the spectral collocation equations for Stokes flow in rectangularly decomposable domains, **SIAM J. Sci. Stat. Comput.**, Vol. 10, pp. 89-130.

Przmieniecki, J. S. (1963): Matrix structural analysis of sub-structures, **AIAA J.**, Vol. 1, pp. 138-147.

Pulicani, J. P. (1986): A spectral multi-method for the solution of 1-D Helmholtz and Stokes type equations, Prepublication No. 122, University de Nice.

Quarteroni, A. (1989): Domain decomposition methods for systems of conservation laws: spectral collocation approximations, ICASE Report No. 89-5.

Quarteroni, A.; and Sacchi-Landriani, G. (1988): Domain decomposition preconditioners for the spectral collocation method, ICASE Report No. 88-11 (submitted to **J. Sci. Comput.**).

Quarteroni, A.; and Sacchi-Landriani, G. (1989): Parallel algorithms for the capacitance matrix method in domain decompositions, **Calcolo**, in press.

Ronquist, E. M. (1988): Optimal spectral element methods for the unsteady three-dimensional incompressible Navier-Stokes equations, Ph.D. Thesis, M.I.T.

Ronquist, E. M.; and Patera, A. T. (1987a): A Legendre spectral element method for the Stefan problem, **Int. J. Numer. Method Eng.**, Vol. 24, pp. 2273-2299.

Ronquist, E. M.; and Patera, A. T. (1987b): Spectral element multigrid. I. Formulation and numerical results, **J. Sci. Comput.**, Vol. 2, pp. 389-406.

Sacchi-Landriani, G.; and Vandeven, H. (1989): Error estimates for the spectral approximation of the nonstationary Stokes problem, **SIAM J. Numer. Anal.**, in press.

Schumann, U.; and Sweet, R. A. (1988): Fast Fourier transforms for direct solution of Poisson's equation with staggered boundary conditions, **J. Comput. Phys.**, Vol. 75, pp. 123-137.

Silberman, I. (1954): Planetary waves in the atmosphere, **J. Meteor.**, Vol. 11, pp. 27-34.

Streett, C. L. (1987): Spectral methods and their implementation to solution of aerodynamic and fluid mechanic problems, **Int. J. Numer. Meth. Fluids**, Vol. 7, pp. 1159-1189.

Streett, C. L.; and Hussaini, M. Y. (1989): A numerical simulation of the appearance of chaos in finite-length Taylor-Couette flow, **Appl. Numer. Math.**, in press.

Streett, C. L.; and Macaraeg, M. G. (1986): Preconditioning for First-order Spectral Discretizations, NASA TM 87619.

Streett, C. L.; and Macaraeg, M. G. (1989): Spectral multi-domain for large-scale fluid dynamic simulations, **Appl. Numer. Math.**, in press.

Tadmor, E. (1989): The convergence of spectral methods for nonlinear conservation laws, **SIAM J. Numer. Anal.**, Vol. 26, in press.

Tal-Ezer, H. (1988): High degree interpolation polynomial in Newton form, ICASE Report No. 88-39 (submitted to **SIAM J. Sci. Stat. Comput.**).

Temperton, C. (1983): Self-sorting mixed-radix fast Fourier transforms. **J. Comput. Phys.**, Vol. 52, pp. 1-23.

Tomboulian, S.; Streett, C. L.; and Macaraeg, M. G. (1989): Spectral solution of the incompressible Navier-Stokes equations on the Connection Machine-2, ICASE Report No. 89-1 (to appear in **Proc. of Supercomputing '88**).

Tuckerman, L. (1989a): Divergence-free velocity fields in nonperiodic geometries, **J. Comput. Phys.**, in press.

Tuckerman, L. (1989b): Transformations of matrices into banded form, **J. Comput. Phys.**, in press.

Vandeven, H. (1987): Family of Spectral Filters for Discontinuous Problems, Report No. 159, Ecole Polytechnique, Paris.

Voigt, R. G.; Gottlieb, D.; and Hussaini, M. Y. (eds.) (1984): **Spectral Methods for Partial Differential Equations**, SIAM-CBMS, Philadelphia.

Wong, Y. S. (1989): Iterative methods for the solution of elliptic difference equations on Cyber 205 computers, **Parallel Computing**, in press.

Zang, T. A. (1989): On the rotation and skew-symmetric forms for incompressible flow simulations, **Appl. Numer. Math.** in press.

Zang, T. A.; and Hussaini, M. Y. (1987): Numerical simulation of nonlinear interactions in channel and boundary-layer transition, in **Nonlinear Wave Interactions in Fluids**, AMD-Vol. 87, ed. by R. W. Miksad, T. R. Akylas, and T. Herbert, pp. 131-145, ASME, New York.

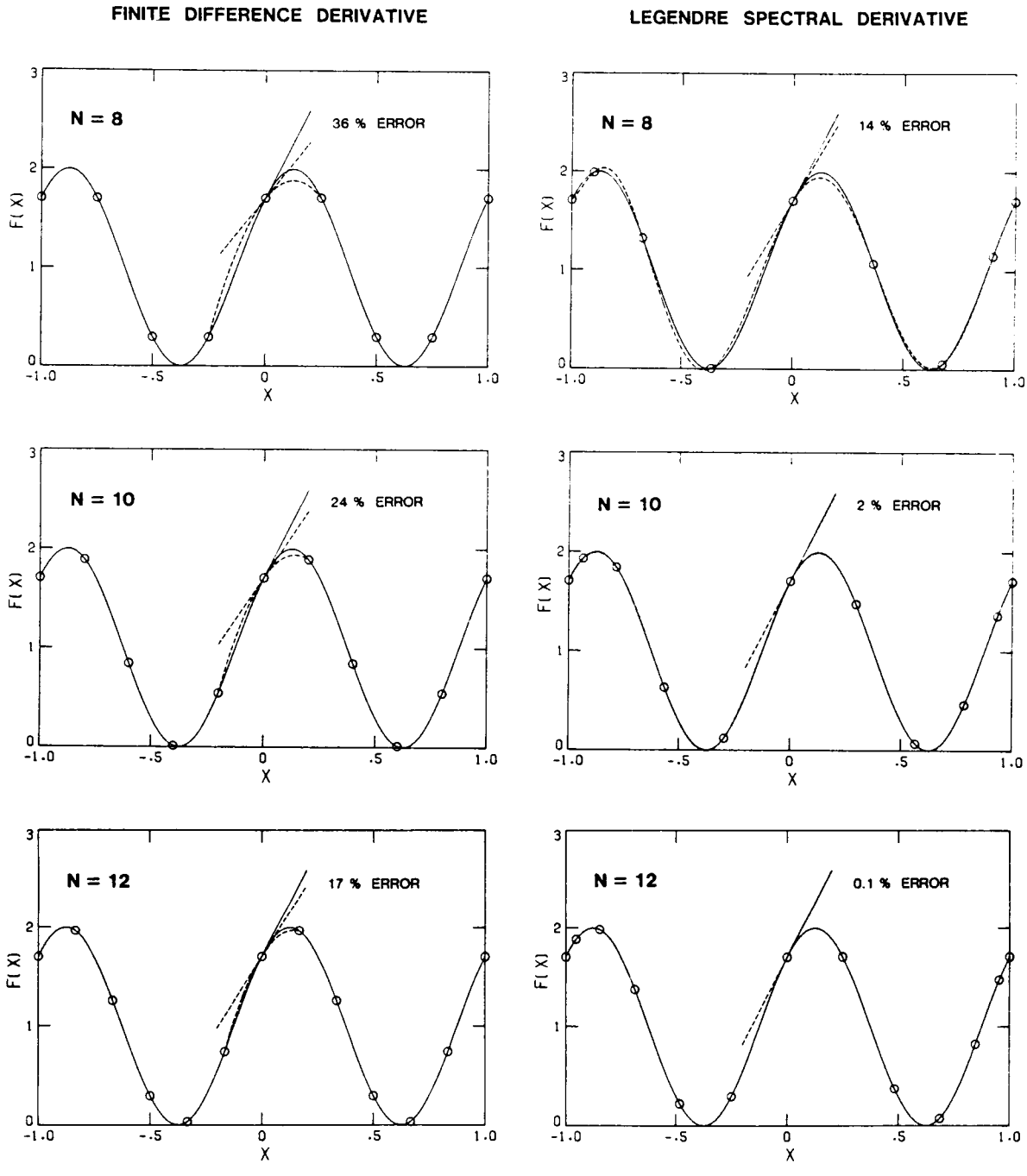

Figure 1. Comparison of finite-difference (left) and Legendre spectral (right) differentiation. The solid curves represent the exact solution, and the dashed curves are their numerical approximations. The solid lines are the exact tangents at $x = 0$, and the dashed lines the approximate tangents. The error in slope is noted, as is the number of intervals N.

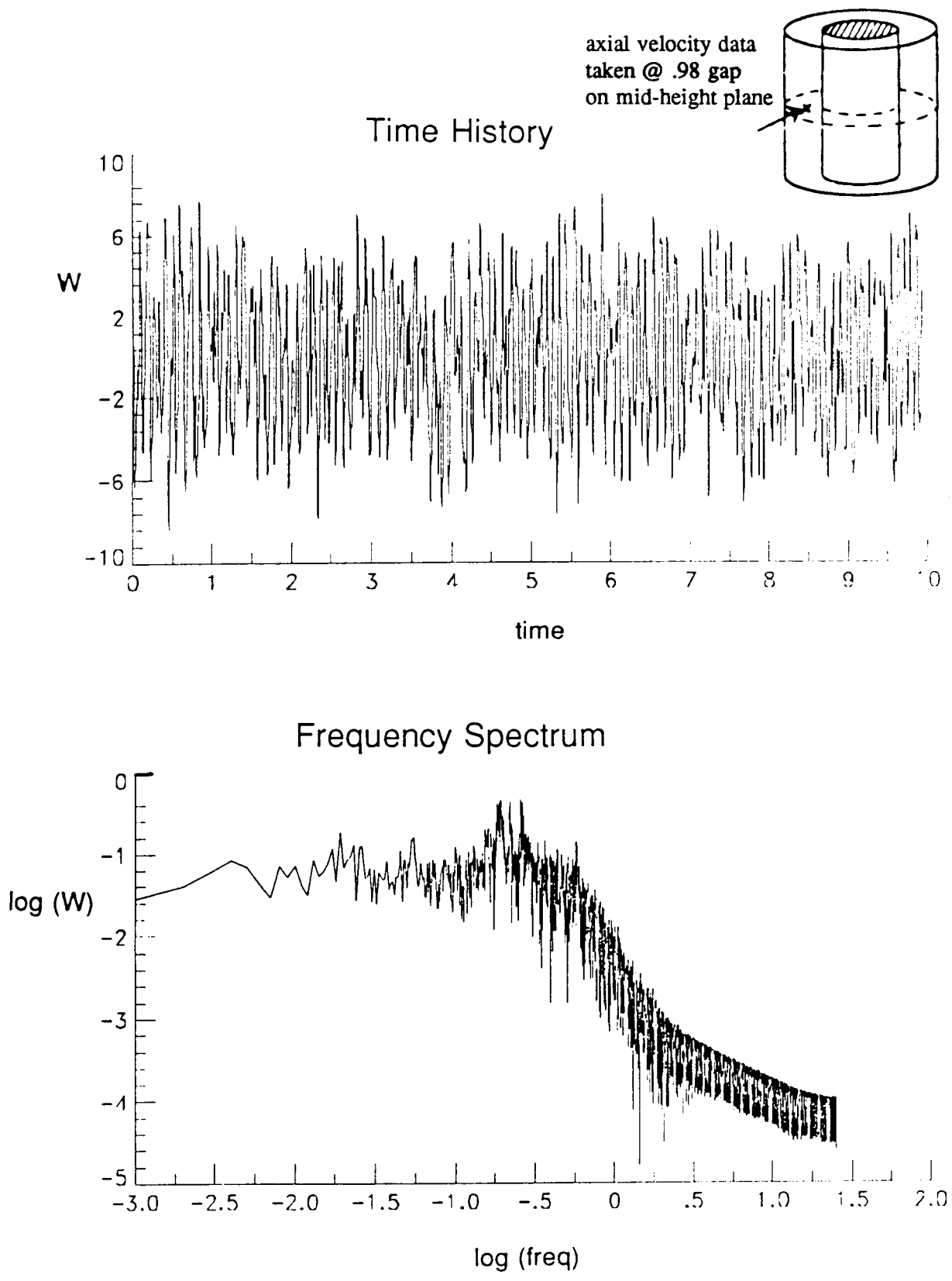

Figure 2: Spectral solution of Taylor-Couette flow. The time history and frequency spectrum of axial velocity data is taken at 98% of the gap between the cylinders and on the mid-height plane.

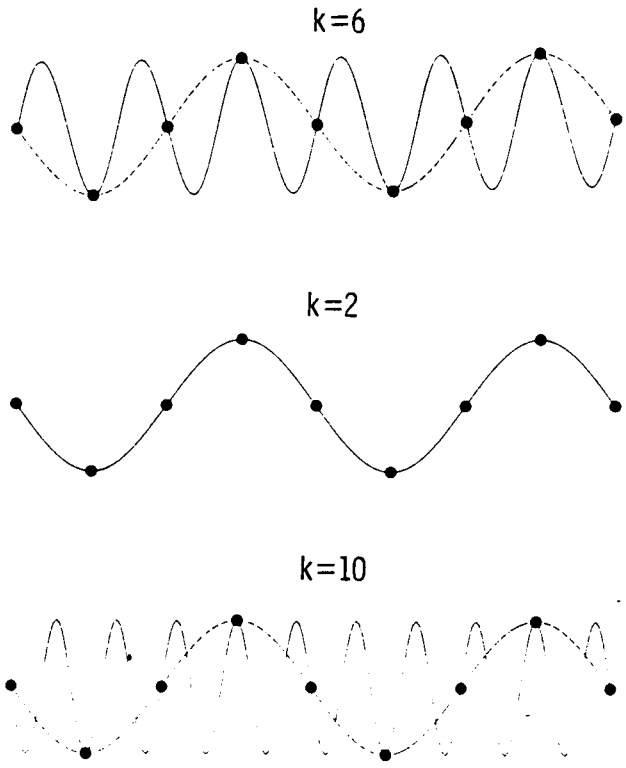

Figure 3. Three sine waves which have the same $k = -2$ interpretation on an eight point grid. The nodal values are denoted by the filled circles. The actual sine waves are denoted by the solid curves. Both the $k = 6$ and the $k = -10$ waves are misrepresented as a $k = -2$ wave (dashed curves) on the coarse grid.

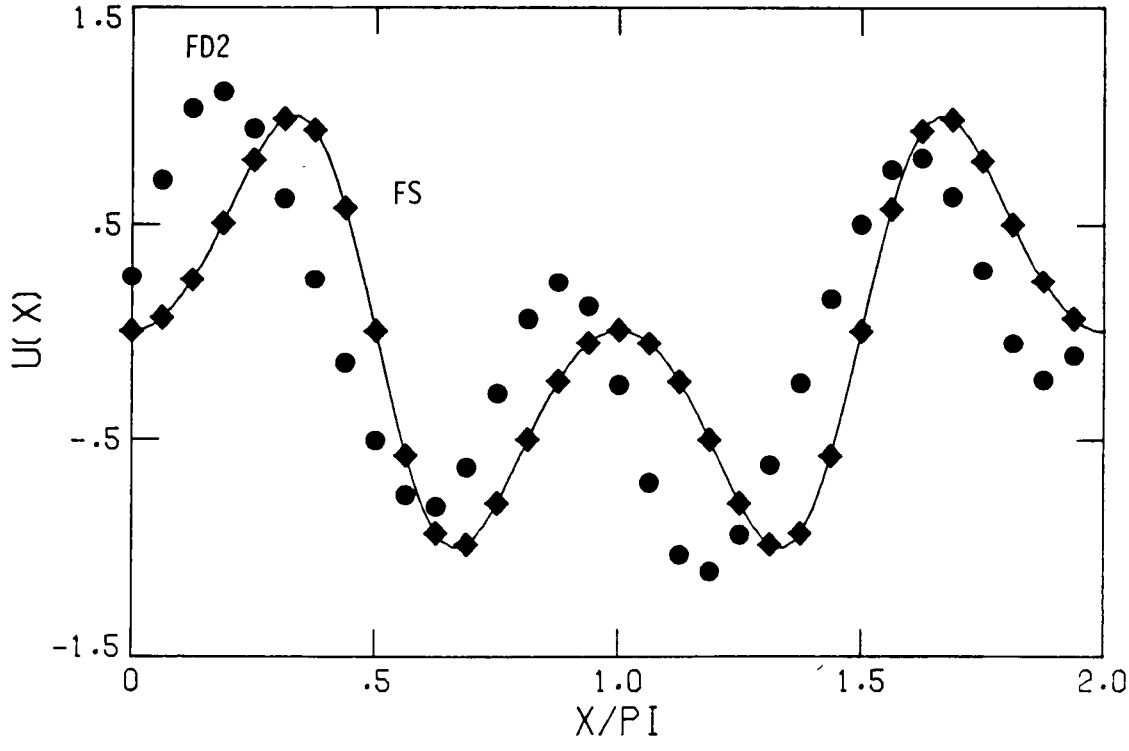

Figure 4. Finite-difference (circles) and Fourier spectral (diamonds) approximations after one period to a periodic wave equations whose exact solution is represented by the curve.

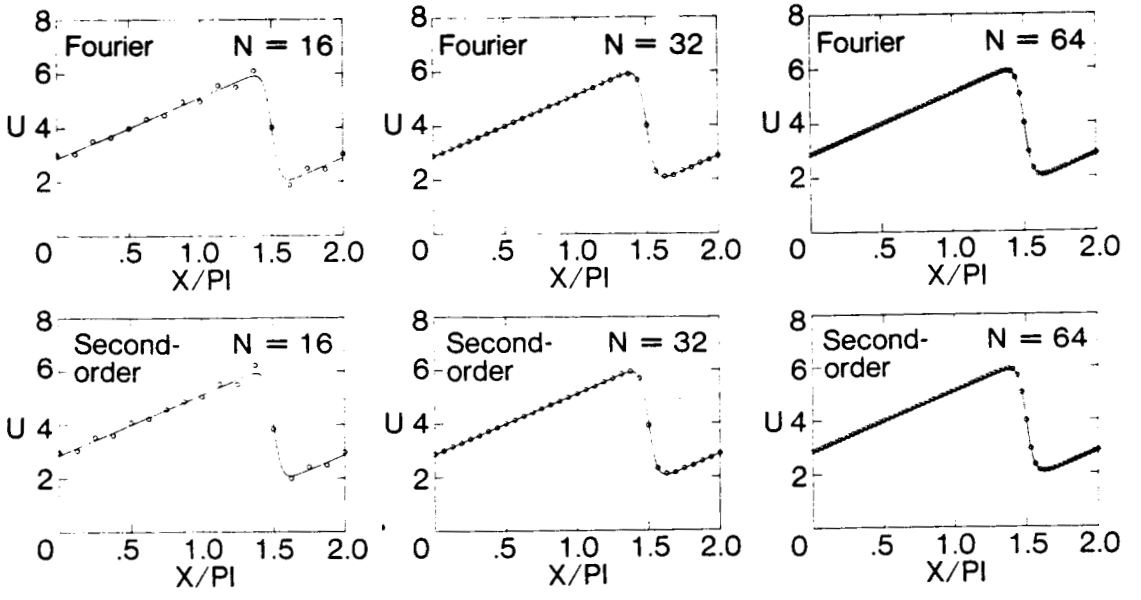

Figure 5. Fourier collocation and second-order finite-difference solutions to the periodic Burgers equation problem at $t = \pi/8$.

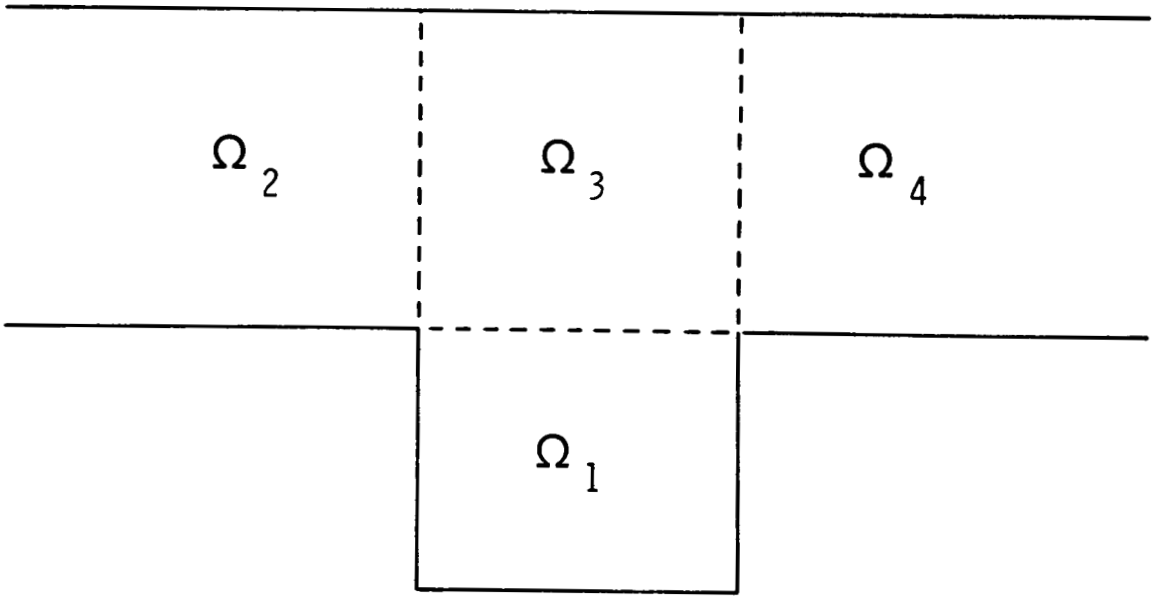

Figure 6. A non-overlapping domain decomposition for a grooved channel.

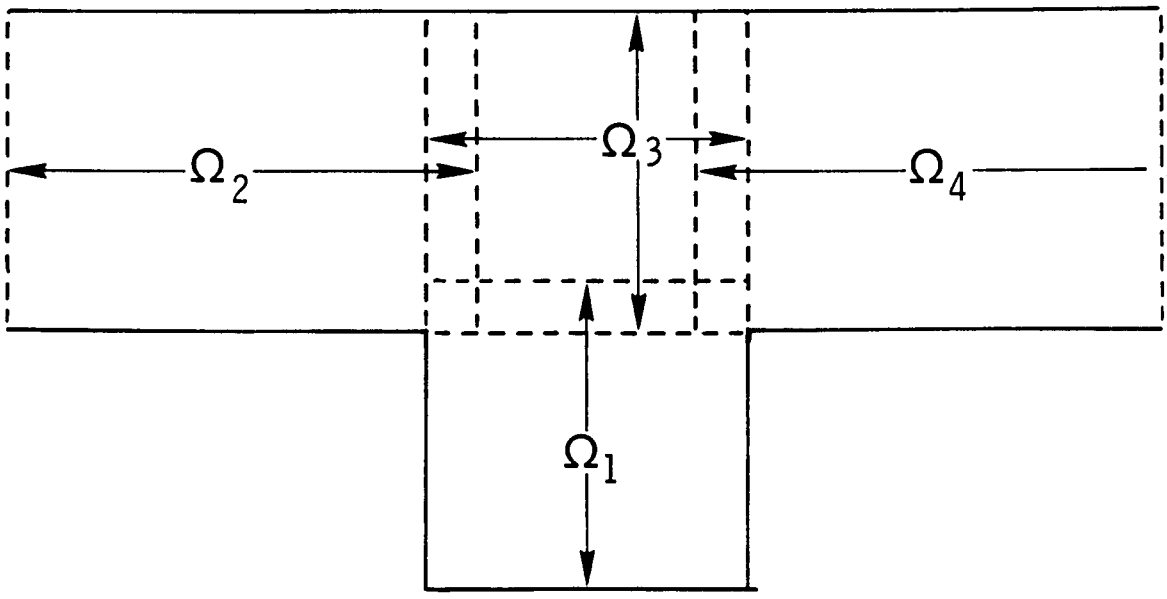

Figure 7. A overlapping domain decomposition for a grooved channel.

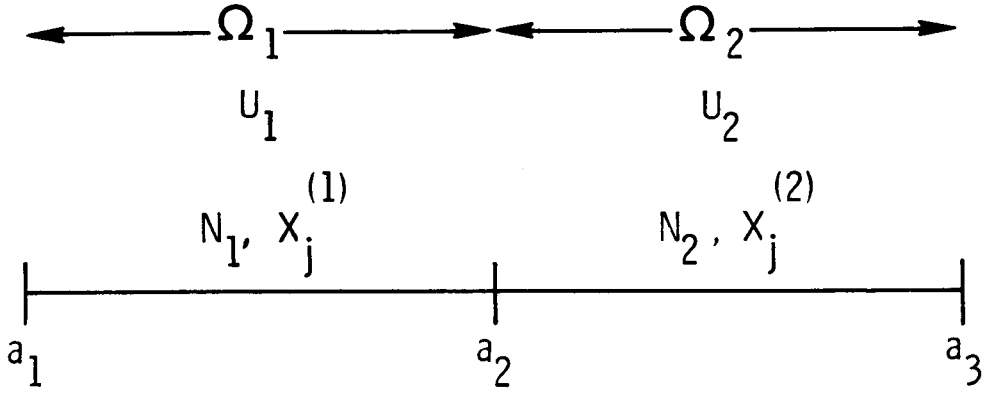

Figure 8. A one-dimensional domain decomposed into two subdomains.

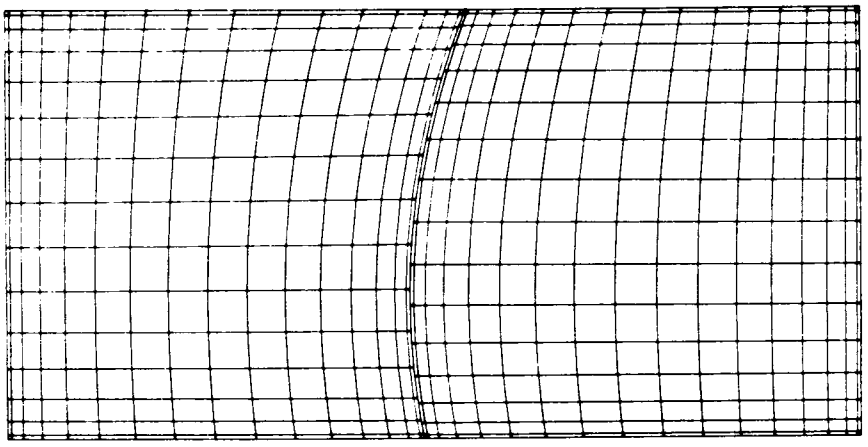

Figure 9. Two-dimensional domain decomposition with non-coincident interface collocation points.

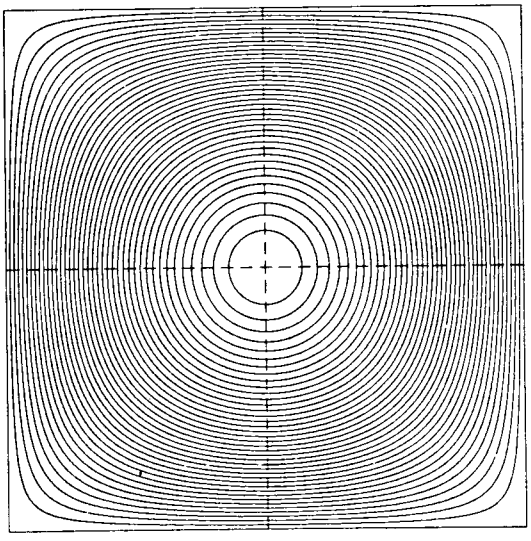

Figure 10. Two-dimensional domain decomposition with cross point. The dashed lines indicate the subdomain boundaries. The cross point is at the intersection of the dashed lines.

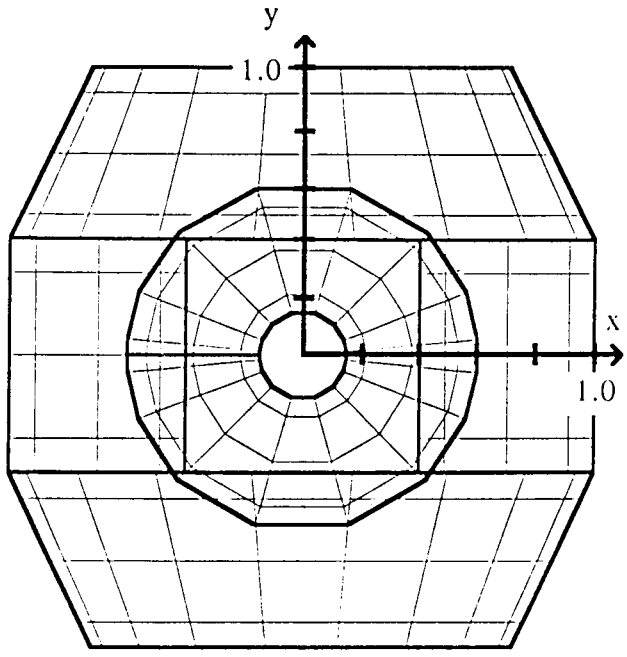

Figure 11. Patched and overset grids for a two-dimensional hyperbolic problem. Heavy lines denote subdomain boundaries and light lines indicate the Chebyshev grids with the subdomains. (Courtesy of D. A. Kopriva.)

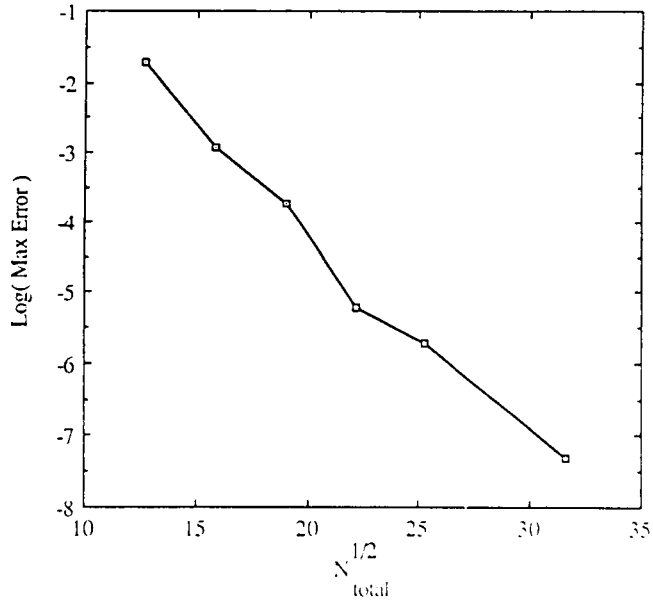

Figure 12. Convergence of spectral domain decomposition method for a hyperbolic system on the grid of Figure 11. (Courtesy of D. A. Kopriva.)

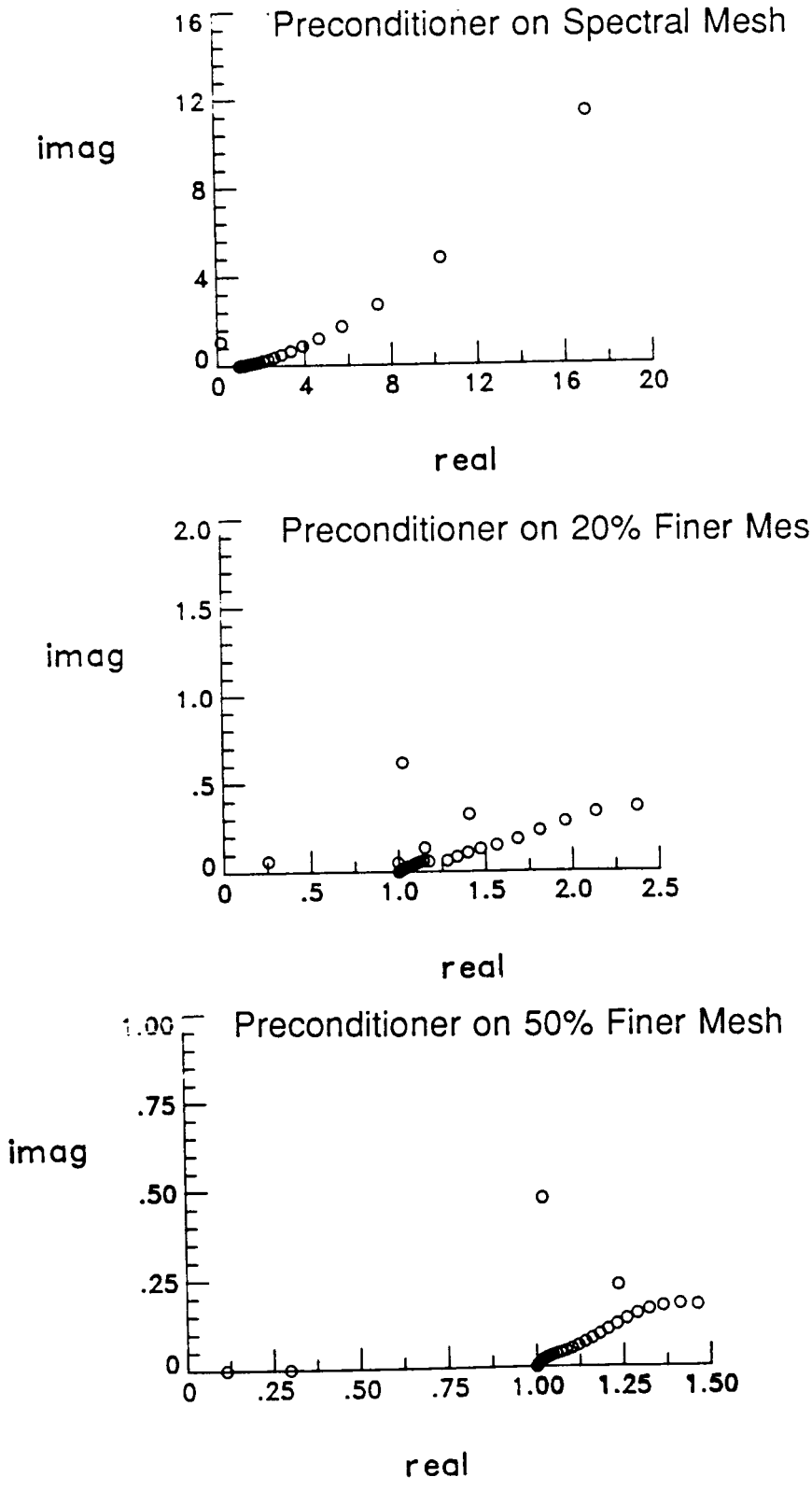

Figure 13. Eigenvalue spectrum of the Chebyshev first-derivative operator, preconditioned with central finite-differences on the spectral mesh, and on meshes refined by 20% and by 50%.

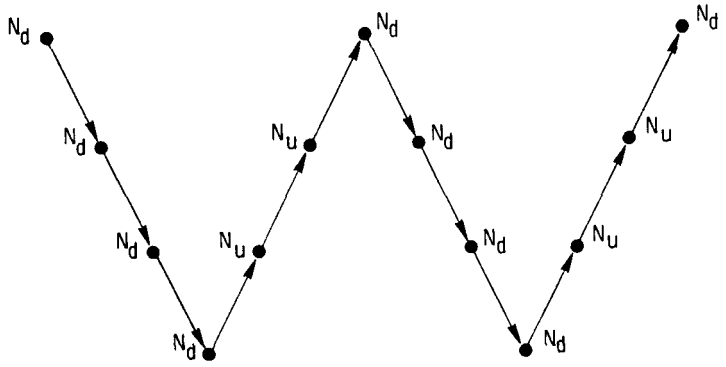

Figure 14. The multigrid V-cycle. The number of relaxations after restriction is denoted by N_d and the number of relaxations before prolongation is denoted by N_u.

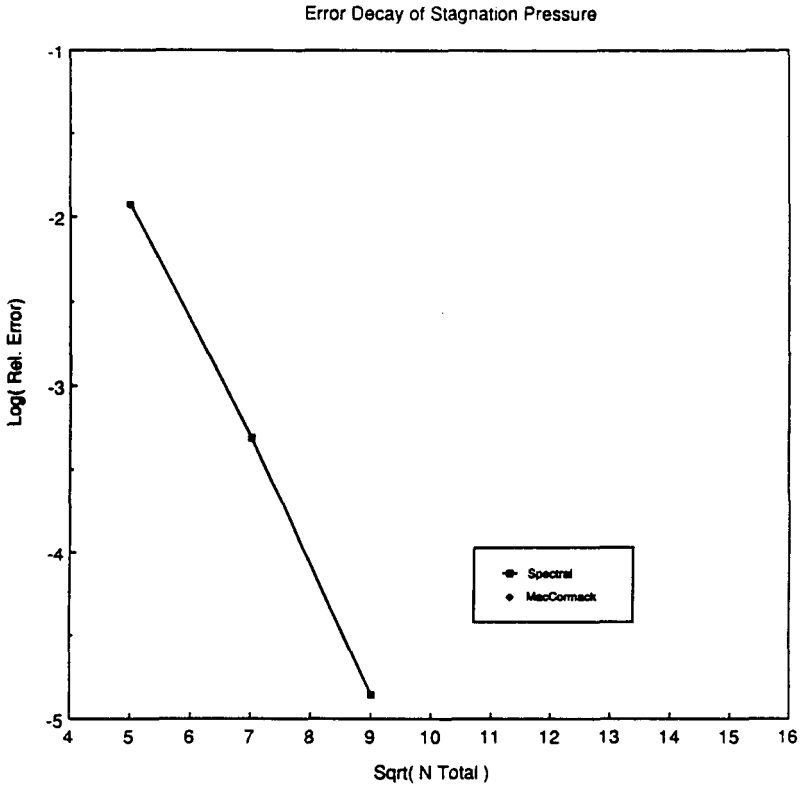

Figure 15. Convergence of spectral shock-fitting solution to the blunt-body problem.

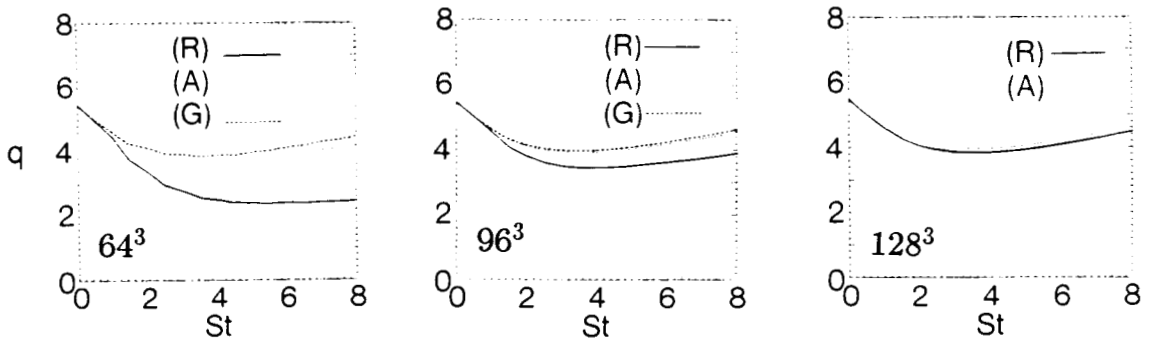

Figure 16. Rotation (R), alternating (A), and Galerkin (G) results for turbulence intensity in homogeneous turbulence on 64^3, 96^3, and 128^3 grids.

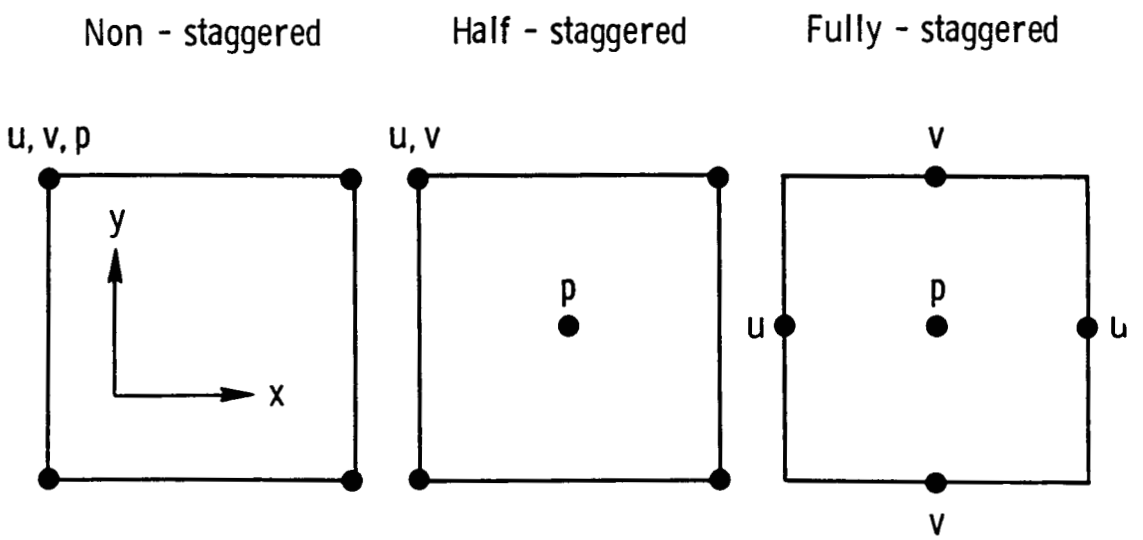

Figure 17. Alternative velocity and pressure nodes for flows with two non-periodic directions.

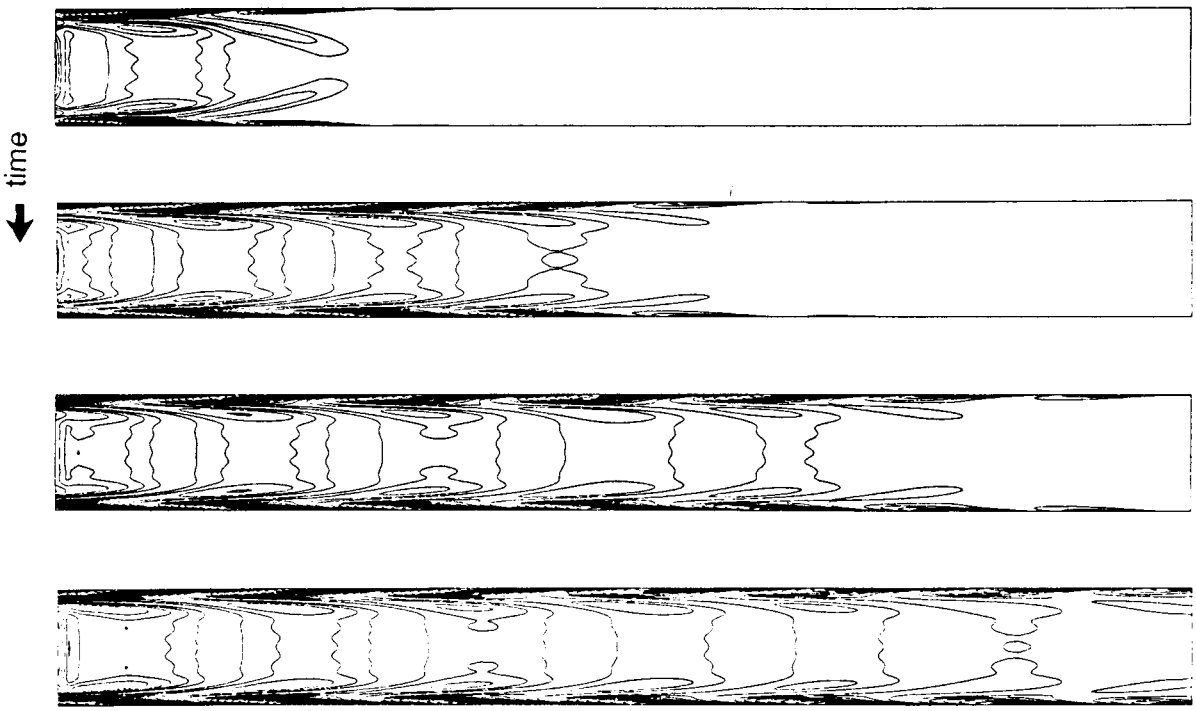

Figure 18. Contours of disturbance vorticity in the primary domain for the incompressible channel flow simulation. The solution in the buffer domain is not shown.

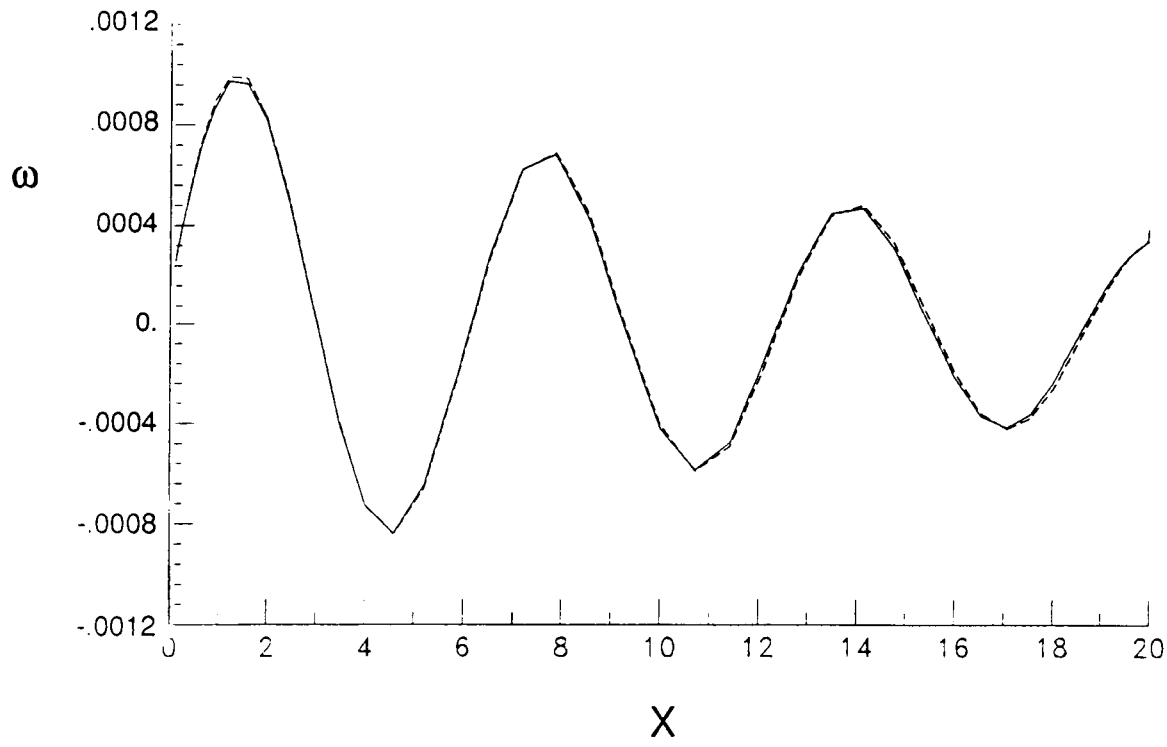

Figure 19. Surface disturbance vorticity on the lower channel wall for the incompressible channel flow simulation. The solid line is the simulation result and the dashed line is the linear theory prediction.

NASA National Aeronautics and Space Administration

Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CR-181803 ICASE Report No. 89-13		

4. Title and Subtitle

SPECTRAL METHODS FOR CFD

5. Report Date

February 1989

6. Performing Organization Code

7. Author(s)

Thomas A. Zang
Craig L. Streett
M. Yousuff Hussaini

8. Performing Organization Report No.

89-13

10. Work Unit No.

505-90-21-01

9. Performing Organization Name and Address

Institute for Computer Applications in Science and Engineering
Mail Stop 132C, NASA Langley Research Center
Hampton, VA 23665-5225

11. Contract or Grant No.

NAS1-18605

13. Type of Report and Period Covered

Contractor Report

12. Sponsoring Agency Name and Address

National Aeronautics and Space Administration
Langley Research Center
Hampton, VA 23665-5225

14. Sponsoring Agency Code

15. Supplementary Notes

Langley Technical Monitor:
Richard W. Barnwell

Final Report

von Karman Institute Lecture Survey on Computation Fluid Dynamics March 6-10, 1989

16. Abstract

One of the objectives of these notes is to provide a basic introduction to spectral methods with a particular emphasis on applications to computational fluid dynamics. Another objective is to summarize some of the most important developments in spectral methods in the last two years. The fundamentals of spectral methods for simple problems will be covered in depth, and the essential elements of several fluid dynamical applications will be sketched.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
spectral methods, fluid dynamics	Unclassified - Unlimited 64 - Numerical Analysis 34 - Fluid Mechanics & Heat Transfer

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	69	A04

NASA FORM 1626 OCT 86