JOURNAL OF COMPUTATIONAL PHYSICS 69, 112–142 (1987)

Spectral Methods Using Rational Basis Functions on an Infinite Interval

JOHN P. BOYD

*Department of Atmospheric & Oceanic Science,
University of Michigan,
2455 Hayward Avenue, Ann Arbor, Michigan 48109*

Received November 25, 1985; revised June 10, 1986

By using the map $y = L \cot(t)$ where L is a constant, differential equations on the interval $y \in [-\infty, \infty]$ can be transformed into $t \in [0, \pi]$ and solved by an ordinary Fourier series. In this article, earlier work by Grosch and Orszag (*J. Comput. Phys.* **25**, 273 (1977)), Cain, Ferziger, and Reynolds (*J. Comput. Phys.* **56**, 272 (1984)), and Boyd (*J. Comput. Phys.* **25**, 43 (1982); **57**, 454 (1985); *SIAM J. Numer. Anal.* (1987)) is extended in several ways. First, the series of orthogonal rational functions converge on the exterior of bipolar coordinate surfaces in the complex y-plane. Second, Galerkin's method will convert differential equations with polynomial or rational coefficients into banded matrix problems. Third, with orthogonal rational functions it is possible to obtain exponential convergence even for $u(y)$ that asymptote to a constant although this behavior would wreck alternatives such as Hermite or sinc expansions. Fourth, boundary conditions are usually "natural" rather than "essential" in the sense that the singularities of the differential equation will force the numerical solution to have the correct behavior at infinity even if no constraints are imposed on the basis functions. Fifth, mapping a finite interval to an infinite one and then applying the rational Chebyshev functions gives an exponentially convergent method for functions with bounded endpoint singularities. These concepts are illustrated by five numerical examples. © 1987 Academic Press, Inc.

1. INTRODUCTION

Grosch and Orszag [1] showed that differential equations on a semi-infinite interval can be solved very effectively by mapping the interval into $[-1, 1]$ using an algebraic function for the map and then expanding the unknown as a series of Chebyshev polynomials. Boyd [2] generalized their technique by giving a map appropriate for $y \in [-\infty, \infty]$ and using the method of steepest descent to derive simple, explicit criteria for choosing the optimum value of the map parameter L. One interesting conclusion was that L is not merely a function of the singularities and asymptotic behavior (as $|y| \to \infty$) of the solution $u(y)$, but also depends upon N, the number of Chebyshev polynomials retained in the truncated series for $u(y)$: the L which is best for $N = 10$ may be much smaller than the best choice for $N = 40$. Cain, Ferziger, and Reynolds [10] independently devised the same mapping for $y \in [-\infty, \infty]$.

0021-9991/87 $3.00

TABLE I

The Mapped Chebyshev Polynomials: $TB_n(y)$ $[L=1]$

n	$TB_n(y)$
[Symmetric about $y=0$]	
0	1
2	$(y^2-1)/(y^2+1)$
4	$(y^4-6y^2+1)/(y^2+1)^2$
6	$(y^6-15y^4+15y^2-1)/(y^2+1)^3$
8	$(y^8-28y^6+70y^4-28y^2+1)/(y^2+1)^4$
10	$(y^{10}-45y^8+210y^6-210y^4+45y^2-1)/(y^2+1)^5$
[Antisymmetric about $y=0$]	
1	$y/(y^2+1)^{1/2}$
3	$y(y^2-3)/(y^2+1)^{3/2}$
5	$y(y^4-10y^2+5)/(y^2+1)^{5/2}$
7	$y(y^6-21y^4+35y^2-7)/(y^2+1)^{7/2}$
9	$y(y^8-36y^6+126y^4-84y^2+9)/(y^2+1)^{9/2}$

The earlier articles, however, left much undone. Grosch and Orszag [1] gave only numerical examples with almost no theory; Boyd [2] is all theory and does not actually solve any differential equations (although this is remedied for one unusual class of problem in Boyd [4]). In this new report, we improve both the algorithms and the theoretical foundations for spectral methods on an infinite interval.

The orthogonal rational functions, Table I, we employ as a basis set are merely mapped Chebyshev polynomials, which in turn are but the transformed cosines of a Fourier series. Throughout this work, we shall employ the convention that $y \in [-\infty, \infty]$ is the original, unmapped coordinate, $x \in [-1, 1]$ is the argument of the Chebyshev polynomials, and t (for trigonometric!) is the argument of the cosine functions where $t \in [0, \pi]$. Then with the maps

$$y = Lx/(1-x^2)^{1/2} \tag{1.1}$$

$$= L\cot(t), \tag{1.2}$$

$$x = \cos(t), \tag{1.3}$$

where L is a constant, our basis functions, which we denote by $TB_n(y)$ as in [4], can be given the equivalent representations

$$TB_n(y) \equiv T_n(x) = T_n[y/(L^2+y^2)^{1/2}], \qquad n = 0, 1, 2, \tag{1.4}$$

$$\equiv \cos(nt) = \cos[n \operatorname{arccot}(y/L)], \tag{1.5}$$

where the $T_n(x)$ are the usual Chebyshev polynomials [5].[1] The new basis fuctions

[1] Note that one must use a nonstandard branch of the arccot(y)—by adding π to the result of the compiler's arccot for $y<0$—to avoid a discontinuity in the inverse map from y to t at $y=0$.

$TB_n(y)$ are rational functions of y for n even and are rational functions divided by a square root for the functions antisymmetric about $y=0$, which have n odd. In a mild abuse of terminology, we will call the $TB_n(y)$ the "rational Chebyshev functions" as in [3] even though, strictly speaking, they are rational only for even degree. They satisfy the orthogonality relation (for $L=1$),

$$\int_{-\infty}^{\infty} \frac{TB_m(y)\,TB_n(y)}{(1+y^2)}\,dy = \begin{cases} \pi/2, & m=n>0 \\ \pi, & m=n=0 \\ 0, & m \neq n. \end{cases} \tag{1.6}$$

However, blindly applying Galerkin's method using (1.6) is usually inefficient. The mappings (1.1) offer both the opportunity of simplying the programming and also of understanding the special problems posed by differential equations on an infinite interval.

Section 2 discusses the transformations in the complex plane: the strip of convergence for a Fourier series in becomes the exterior of a *bipolar* coordinate surface in the complex y-plane! The topic of Section 3 is the mechanics of using the $TB_n(y)$: it is usually simplest to convert the problem to the trigonometric coordinate t and then apply Fourier methods. Section 4 exploits this "trigonometric" methodology to show that spectral methods often lead to sparse, banded matrices when the rational Chebyshev functions are the basis set. Section 5 analyzes functions $u(y)$ that decay *algebraically* or asymptote to a constant as $|y| \to \infty$: exponentially convergent series are still possible provided the Fourier series in t is properly matched to the asymptotic behavior of the function. Section 6 illuminates the distinction between "natural" and "essential" boundary conditions; in contrast to Chebyshev methods on a finite interval where the boundary conditions usually must be imposed *explicitly* on the numerical solution, the boundary conditions can often be ignored when the domain of integration is infinite. Section 7 offers five numerical examples. The final section is a summary and prospectus.

2. Regions of Convergence

For ordinary power series, it is well known that (i) the domain of convergence is bounded by a *circle* in the complex y plane, (ii) the function $u(y)$ is *singular* somewhere on the convergence-limiting circle, and (iii) the power series coefficients a_n decrease like the terms of a *geometric* series, that is,

$$a_n \sim (1/\delta)^n A(n) \qquad \text{[Geometric Convergence]}, \tag{2.1}$$

where the "exponent of convergence" δ is equal to the radius of convergence of the power series, i.e., is equal to the absolute value of the location of that singularity of $u(y)$ which is nearest the origin. $A(n)$ denotes an algebraic (as opposed to exponential) function of n which is independent of the location of the convergence-limiting singularity but depends on its type: $A(n)$ is a constant when the singularity is a simple pole, a constant divided by n when the singularity is a logarithm, and so on.

These familiar results have direct parallels for Fourier, Chebyshev, and rational Chebyshev series. For Fourier series, the review [5] notes the following: (i) the domain of convergence is a strip centered on the real t-axis, (ii) $u(t)$ is singular somewhere on one or both of the lines of constant $\mathrm{Im}(t)$ bounding the strip, and (iii) the coefficients have "geometric convergence" as defined by (2.1) with

$$|\delta| = \exp(|\mathrm{Im}(t_{\mathrm{sing}})|) \qquad \text{(Fourier)}, \tag{2.2}$$

where t_{sing} is the location of the convergence-limiting singularity, that is to say, the position of the singularity nearest the real t-axis. The contours of constant "exponent of convergence" $\delta(t_{\mathrm{sing}})$ are simply straight lines parallel to the real axis.

We can translate these Fourier results to their Chebyshev analogues by analyzing the mapping (1.3): $x = \cos(t)$. We find that the straight lines parallel to the real t-axis are mapped into *ellipses* with foci at $x = \pm 1$. It then follows that (i) the domain of Chebyshev convergence is bounded by an ellipse, (ii) $u(x)$ is singular somewhere on the ellipse of convergence, (iii) the coefficients decrease geometrically as in (2.1) with $|\delta|$ being the same for all functions that have convergence-limiting singularities on the same ellipse.

The ellipses of constant δ are also the contours of constant coordinate μ for a system of elliptical coordinates in the complex x-plane. This is not accidental; the curves that bound domains of convergence for many other series are also coordinate surfaces in the appropriate system. For example, the bounding circles for power series are also contours of constant radius in polar coordinates and the bounding straight lines for Fourier series are also contours of constant y in Cartesian coordinates. A well-known theorem of complex variable theory assures us that the contours of constant real part and the contours of constant imaginary part of a function $f(y)$ will always be orthogonal. Thus, any mapping $t = f(y)$ that transforms a Fourier cosine series into something else—Chebyshev series, $TB_n(y)$, or whatever—will implicitly create a system of orthogonal coordinates in the y plane such that one set of coordinate surfaces will also coincide with surfaces of constant δ.

This is very useful because we can determine the rate of convergence of a series merely by expressing the location of the singularities in the appropriate coordinate system. For a Chebyshev series, for example, the magnitude of the quasiradial elliptical coordinate completely determines the exponent of convergence δ in (2.1).

The coordinate system that plays the analogous role for the $TB_n(y)$ is a rather unusual system known as "bipolar" coordinates; Morse and Feshbach [6] give a good description.

THEOREM 1. *The region of convergence for a series of the functions $TB_n(y)$ defined by*

$$TB_n(L \cot t) \equiv \cos(n\, t) \tag{2.3}$$

is the exterior *of the surfaces of constant* $t_{\text{im}}(y)$ *in the complex y-plane where* t_{im} *is one of the bipolar coordinates* (t_r, t_{im}) *defined by*

$$t_r + i t_{\text{im}} = \operatorname{arccot}(y/L). \tag{2.4}$$

These surfaces are pairs of circles in the complex y-plane, one above the real y-axis and the other below, which enclose either $y = iL$ *or* $y = -iL$. *If these circles are of finite diameter, then the series has "geometric convergence" in the sense of* [7] *and the coefficients are asymptotically (as* $n \to \infty$*):*

$$a_n \sim (1/\delta)^n A(n) \tag{2.1}$$

with

$$|\delta| = \exp(|t_{\text{im}}|), \tag{2.5}$$

where $A(n)$ *is some algebraic (i.e., slowly varying) function of n. The function* $u(y)$ *being expanded is singular only* within *and* on *the disks enclosed by the bipolar surface.*

When the function $u(y)$ *is* singular *at* $y = \infty$, *then the series converges* only *on the* real y-axis *and the rate of convergence is "subgeometric" in the sense defined by Boyd* [7], *that is, the best that we can hope for is*

$$a_n \sim \exp(-qn^r) A(n) \qquad \text{[Subgeometric Convergence]}, \tag{2.6}$$

where $A(n)$ *is again an algebraic function of n and the "index of exponential convergence" r is* < 1. *The* $TB_n(y)$ *series diverges for all* $\operatorname{Im}(y) \neq 0$.

Proof. Starting from the formula [8],

$$\begin{aligned} y_r + i y_{\text{im}} &\equiv L \cot(t) \\ &= L[\sin(2t_r) - i \sinh(2t_{\text{im}})] \\ &\quad /[\cosh(2t_{\text{im}}) - \cos(2t_r)] \end{aligned} \tag{2.7}$$

one can show via elementary trigonometric identities that the surfaces of constant t_{im} are pairs of circles in the complex $y_r - y_{\text{im}}$ plane with centers at

$$c = \pm iL \operatorname{cotanh}(2t_{\text{im}}) \tag{2.8}$$

and radius

$$R = |L/\sinh(2t_{\text{im}})|. \tag{2.9}$$

Note that as $|t_{\text{im}}| \to 0$, $R \to \infty$, and $c \to \pm i\infty$. In other words, the region of divergence bounded by these circles fills more and more of the complex y-plane in this limit, squeezing the region of convergence (which is the exterior of these circles) closer and closer to the real y-axis until the series converges only on the axis itself. Q.E.D.

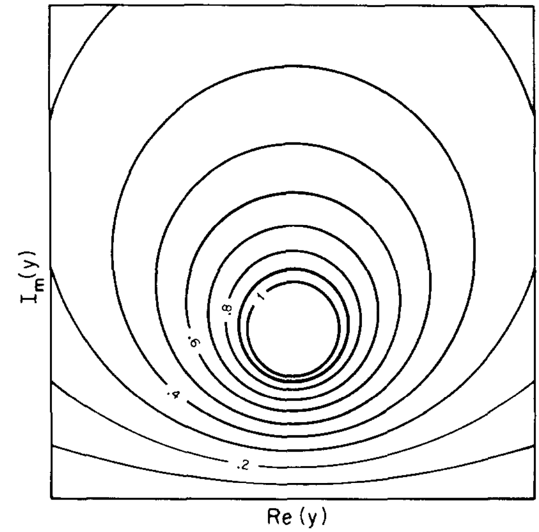

FIG. 1. Contours of constant t_{im}, the bipolar coordinate defined by (2.4), in the upper half of the complex y-plane for $\text{Re}(y) \in [-1.5, 1.5]$, $\text{Im}(y) \in [0, 3]$. All are *circles* enclosing the point $y = i$. The centers and radii of the circles are those for map parameter$L = 1$, but the circular *shape* of the contours is the same for all L. The circles in the lower half-plane are not shown because they are merely the mirror image of those for $\text{Im}(y) > 0$. These circles are contours of "equiconvergence" as explained in the text.

We omit the high school algebra needed to derive (2.8) and (2.9) because we offer Fig. 1 as a graphical proof: it shows the contours of constant t_{im} in the upper half-plane. Figure 2 is a schematic that depicts the regions of convergence for Fourier, Chebyshev, and rational Chebyshev functions and the mappings that transform the boundaries of one into another.

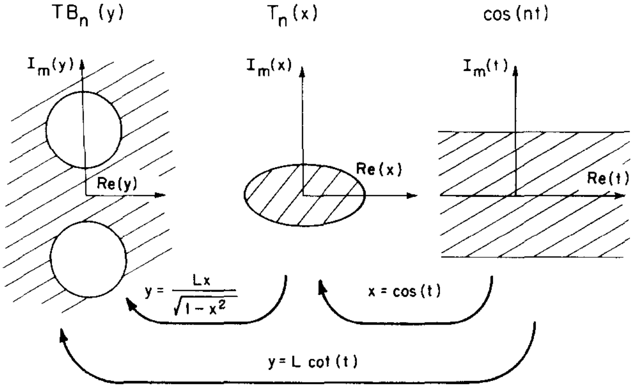

FIG. 2. Schematic illustrating the mappings which transform any of the three basis sets—Fourier cosine series, Chebyshev polynomials T_n, and rational Chebyshev functions TB_n—into one of the others. The cross-hatched regions indicate the domain of convergence for each series.

It may seem strange that the *interior* of an ellipse or strip is transformed into the *exterior* of a bipolar coordinate surface. However, the line segment $x \in [-1, 1]$, which lies wholly within the Chebyshev ellipse of convergence, is transformed by the map (1.1) into the whole y-axis. The image of the boundary ellipse is therefore broken into two curves (which turn out to be circles) separated by the real y-axis. The area between the ellipse and the interval $[-1, 1]$ is turned inside out to become the area between the circles and the y axis.

Most solutions to differential equations on $y \in [-\infty, \infty]$ are singular at infinity, so *subgeometric* convergence is *usual.* This is rather disappointing, but as explained in [7] and [9], the alternatives of series of Hermite functions or sinc functions usually also give subgeometric convergence with $r = \frac{1}{2}$. Expansions on an infinite interval are (obviously!) harder than on a finite line segment, and subgeometric rather than geometric convergence is the usual price. The simple method of determining the asymptotic series coefficients by locating the *position* of the convergence-limiting singularity provides no useful information when that singularity is at infinity, so Boyd [2] was forced to use the method of steepest descent.

Nonetheless, the analysis of the bipolar contours of convergence is important for two reasons. First, it is directly applicable to those solutions which are holomorphic for all real y including infinity. Second, it illustrates an important advantage of $TB_n(y)$ expansions over competing basis sets. Hermite series and sinc expansions converge exponentially only when the function being expanded *decays exponentially* as $|y| \to \infty$. However, Theorem 1 shows that we have not merely exponential convergence but *geometric* convergence for a function like

$$u(y) = 1/(1 + y^2). \tag{2.10}$$

It follows that series of $TB_n(y)$ are extremely effective for solutions which decay slowly—as an algebraic power of y, for example—as $|y| \to \infty$. Section 5 will amplify this point.

3. Mechanics of Coordinate Conversion

Since we are solving a problem on an infinite interval in y, it is appropriate to list the first few basis functions in Table I and to illustrate them in Fig. 3. However, in the author's experience, the easiest way to solve differential equations is to change coordinates to the trigonometric argument y and then apply an ordinary Fourier series.

Since

$$y = L \cot(t), \tag{3.1}$$

where L is the arbitrary, constant map parameter, it follows that

$$u_y = \{1/[L \cot(t)]_t\}\, u_t \tag{3.2}$$

$$= -\{\sin^2(t)/L\}\, u_t. \tag{3.3}$$

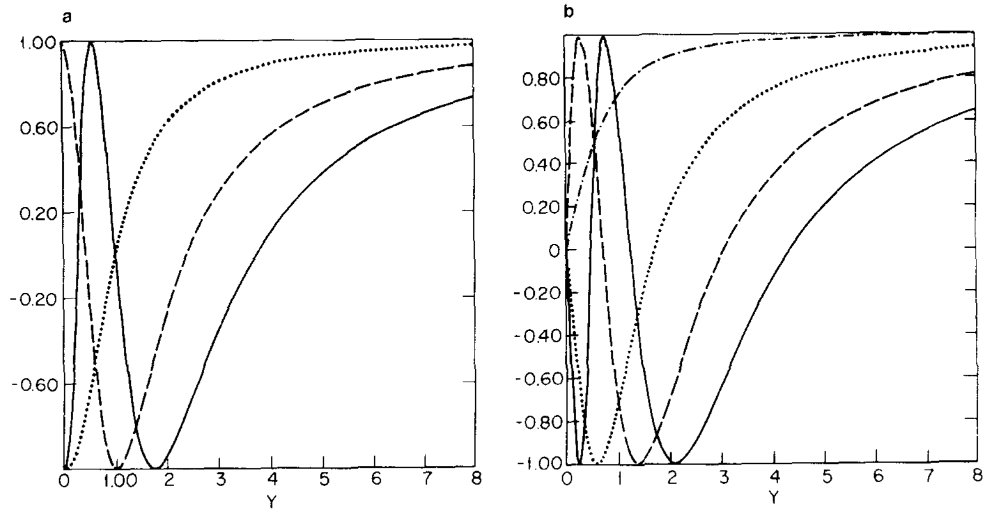

FIG. 3. (a) Plot of the first four basis functions that are symmetric about $y=0$: $TB_0(y)$ [upper y-axis; $TB_0 \equiv 1$], $TB_2(y)$ [dotted], $TB_4(y)$ [dashed], and $TB_6(y)$ [solid]. (b) Graph of the lowest four antisymmetric functions: $TB_1(y)$ [dot–dash], $TB_3(y)$ [dotted], $TB_5(y)$ [dashed], and $TB_7(y)$ [solid]. The functions are plotted only for $y \geqslant 0$ because $TB_{2n}(-y) = TB_{2n}(y)$ and $TB_{2n+1}(-y) = -TB_{2n+1}(y)$ for all n.

By iterating this derivative rule, it is possible to convert differential equations of any order from y to t. Table II gives the conversion formulas for derivatives of up to sixth order.

Examples

$$u_y + p(y)\, u = f(y) \tag{3.4a}$$

becomes

$$\sin^2(t)\, u_t - Lp[L\cot(t)]\, u = -Lf(L\cot(t)]. \tag{3.4b}$$

Similarly,

$$u_{yy} + p(y)\, u = f(y) \tag{3.5a}$$

becomes

$$\sin^4(t)\, u_{tt} + 2\cos(t)\sin^3(t)\, u_t + L^2 p[y(t)] = L^2 f(y[t]). \tag{3.5b}$$

In the next section, we show that this conversion to trigonometric form is very convenient for analyzing the sparsity of the matrices generated by Galerkin's method. Before doing so, however, we should note that it is equally possible to leave the differential equation in terms of y and use the transformation rules of

TABLE II

Transformations of Derivatives for the Mapping $y = L\cot(t)$ which Converts a Rational-Chebyshev Series in $TB_n(y)$ into a Fourier Cosine Series in $\cos(nt)$[a]

$$u_y = [-\sin^2(t)/L]\, u_t,$$
$$u_{yy} = [\sin^3(t)/L^2]\{\sin(t)\, u_{tt} + 2\cos(t)\, u_t\}$$

u_{yyy}

u_t	u_{tt}	u_{ttt}
$8\sin^2$	$-6\cos\sin$	$-\sin^2$
-6		
	$[\times \sin^4(t)/L^3]$	

u_{4y}

u_t	u_{tt}	u_{ttt}	u_{4t}
$24\cos$	$36\sin$	$12\cos\sin^2$	$\sin^3$
$-48\cos\sin^2$	$-44\sin^3$		
		$[\times \sin^5/L^4]$	

u_{5y}

u_t	u_{tt}	u_{ttt}	u_{4t}	u_{5t}
$-384\sin^4$	$400\cos\sin^3$	$140\sin^4$	$-20\cos\sin^3$	$-\sin^4$
$+480\sin^2$	$-240\cos\sin$	$-120\sin^2$		
-120			$[\times \sin^6/L^5]$	

u_{6y}

u_t	u_{tt}	u_{ttt}	u_{4t}	u_{5t}	u_{6t}
$3840\cos\sin^4$	$1800\sin$	$1200\cos\sin^2$	$300\sin^3$	$30\cos\sin^4$	$\sin^5$
$-3840\cos\sin^2$	$-6000\sin^3$	$-1800\cos\sin^4$	$-340\sin^5$		
$+\ 720\cos$	$+4384\sin^5$			$[\times \sin^7/L^6]$	

Note. The factor $[\times \sin^q/L^p]$ denotes that all entries in the table must be multiplied by this common factor of $\sin(t)$ raised to the qth power, divided by the pth power of the map factor.

[a] L is a constant: the "map parameter".

Table II, in combination with arccot-inverse mapping, to evaluate $TB_n(y)$ and its derivatives. For example, (3.3) tells us that

$$d\,TB_n/dy = -\{\sin^2[\operatorname{arccot}(y/L)]/L\}(-n)\sin[n \operatorname{arccot}(y/L)]. \tag{3.6}$$

From a mathematical viewpoint, it is quite irrelevant whether we perform the transformation from y to t by hand as in (3.5) and then apply Fourier series, or whether we use the transformation to evaluate the $TB_n(y)$ and all its derivatives using only the compiler-supplied routines for trigonometric functions.[2] To a programmer,

[2] To repeat the warning given earlier, we must add π to the result of the compiler function ACOT(Y) for negative y so that $t(y)$ is continuous at $y = 0$.

however, burying the transformation in the subroutines for evaluating the $TB_n(y)$ may be preferable since it allows writing software that shields the user from the need to understand the trigonometric transformation.

4. Galerkin's Method and Banded Matrices

The numerical examples in Section 7 were computed using the "collocation" or "pseudospectral" method. Its major drawback is that the matrices which are the discretization of the differential equation are always full matrices, that is to say, most of the matrix elements are nonzero. When the equation has complicated or transcendental coefficients, full matrices are unavoidable. However, it is well known that when the coefficients of the differential equation are powers of y, Galerkin's method, also known as the "spectral" algorithm, will give matrices with banded structure, that is, all but a handful of matrix elements in each row will be identically 0.

The classic Chebyshev example is

$$\nu u_{xx} = f(x), \tag{4.1}$$

where ν is a constant. As discussed in [5, pp. 119–120], it is possible to manipulate the algebraic equations that result from Galerkin's method into a tridiagonal matrix (three nonzero elements in each row). Second-order finite difference also generate a tridiagonal matrix, but for a given number of degrees of freedom, the Chebyshev solution is much more accurate.

This trick is extremely useful in semi-implicit time-stepping algorithms, especially in fluid mechanics. To avoid a very short time step, the viscous terms must be treated implicitly, but this requires solving an equation like (4.1) at *every* time step where $f(x)$ represents advective, pressure, and forcing terms at the previous time level. If we had to invert a *full* matrix at every time step, the semi-implicit algorithm would be prohibitively expensive. However, because band matrices can be inverted in $O(N)$ operations where N is the number of terms in the Chebyshev series, Galerkin solution of (4.1) allows the semi-implicit algorithm to be only slightly more expensive—per time step—than a fully explicit procedure. The latter is then much more costly per run because it requires a time step an order of magnitude smaller than the semi-implicit algorithm.

The rational Chebyshev functions also share this property of "bandedness," that is, of creating banded matrices when the coefficients of the differential equation are powers of y. It is easiest to illustrate the idea through a specific example:

$$u_{yy} + (\lambda - y^2)\, u = 0, \tag{4.2}$$

where λ is the eigenvalue. The equivalent equation in the trigonometric variable t is

$$[\sin^6(t)/L^2]\, u_{tt} + [2\cos(t)\sin^5(t)/L^2]\, u_t + [\lambda \sin^2(t) - L^2\cos^2(t)]\, u = 0. \tag{4.3}$$

To solve (4.3) via either the spectral or pseudospectral method, we assume

$$u \approx \sum_{j=0}^{N} a_j \cos(jt) \tag{4.4}$$

and substitute this into (4.4). In Galerkin's method, the matrix representation of the differential equation

$$Mu = f, \tag{4.5}$$

where M is the differential operator is

$$\bar{H}\bar{A} = \bar{F}, \tag{4.6}$$

where $\bar{A}$ is the column vector whose elements are the Fourier coefficients, a_j, and

$$\bar{F}_k = (\cos(kt), f(t)), \tag{4.7}$$

$$\bar{H}_{kj} = (\cos(kt), M\cos(jt)), \tag{4.8}$$

where the inner product is

$$(f(t), g(t)) \equiv \int_0^{\pi} f(t)\, g(t)\, dt. \tag{4.9}$$

Since the basis functions are orthogonal under the inner product, it follows that the matrix $\bar{H}$ will be banded if and only if the result of applying the differential operator M to $\cos(jt)$ can be expressed as a *finite* sum of cosines. However, familiar identities such as

$$\cos(a)\cos(b) = [\cos(a+b) + \cos(a-b)]/2, \tag{4.10a}$$

$$\sin(a)\cos(b) = [\sin(a+b) + \sin(a-b)]/2. \tag{4.10b}$$

justify the following.

THEOREM 2. *Let α and β be nonnegative integers and let β be even. Then*

$$\cos^{\alpha}(t)\sin^{\beta}(t)\cos(jt) = \sum_{m=-(\alpha+\beta)}^{\alpha+\beta} b_{j+m}\cos[(j+m)\,t], \tag{4.11}$$

that is, the product of $\cos(jt)$ *with a symmetric trigonometric polynomial of degree* $(\alpha+\beta)$ *can be written as the sum of at most* $2(\alpha+\beta)+1$ *terms. Although not necessary for our immediate needs, the theorem is also true when* $\cos(jt)$ *is replaced by* $\sin(jt)$. *If β is odd so that the polynomial is antisymmetric about $t=0$, then* (4.11) *still applies with the replacement of cosines by sines on the* r.h.s.

Proof. Omitted because it is a trivial consequence of repeatedly applying (4.10) and similar identities.

The theorem, although so elementary that it could be easily proved by a high school studient, is very powerful. Comparing (4.11) with (4.3), we see that since $\cos(jt)$ must be multiplied by powers of $\sin(t)$ and $\cos(t)$ of sixth degree, it follows that there can be no more than thirteen nonzero elements in each row of the matrix. For this particular equation, the coefficients are symmetric about $y=0$ ($t=\pi/2$), so the matrix rows for n even decouple from those for n odd and there are in fact only seven nonzero elements in each row or column as follows:

$$H_{jj} = 32L^2\lambda - 32L^4 - 20j^2, \tag{4.12a}$$

$$H_{j\pm 2,j} = 15j^2 \pm 10j - 16L^4 - 16L^2\lambda, \tag{4.12b}$$

$$H_{j\pm 4,j} = \mp 8j - 6j^2, \tag{4.12c}$$

$$H_{j\pm 6,j} = j^2 \pm 2j, \tag{4.12d}$$

where we have divided out a common factor of $\pi/(128L^2)$. The identities show that elements with negative array indices should be interperted as terms added to the element whose row index has the same absolute value, i.e., $H_{-2,j}$ should be added to $H_{2,j}$ and so on.

For some problems, it is necessary to modify the Galerkin's matrix by altering a couple of rows to impose the boundary conditions. In Section 6 and in the numerical examples of Section 7, however, we will see that for most problems on an infinite domain, the boundary conditions are "natural" rather than essential, and the Galerkin's matrix is unaltered. When boundary conditions must be imposed, matrix "bordering" techniques [11] allow one to still exploit the sparsity of the matrix.

This delightful property of the cotangent transformation was first noted by Cain, Ferziger, and Reynolds [10]: the derivative transformation formulas in Table II are, after multiplication by a common factor, trigonometric polynomials. They observe quite correctly that this property minimizes aliasing error. However, aliasing will be exponentially small for any smooth mapping if N is sufficiently large to resolve $u(y)$. It is probably more significant that the sparsity of the Galerkin's matrix will reduce the cost by an order of magnitude than that it will simultaneously make the error a little smaller.

This illustration, (4.3), was chosen with malice aforethought to be the parabolic cylinder equation; the exact eigenfunctions are simply the Hermite functions, which are a long-standing competitor to the $TB_n(y)$ for solving problems on an infinite interval. The Hermite–Galerkin equivalent of (4.12) is a diagonal matrix with only a single nonzero element in each row or column. When the coefficients of the differential equation are polynomials in y, Hermite expansions also give Galerkin matrices which are banded. For example, when (4.3) is generalized by adding in a term in y^4 to the coefficient of u, the Hermite–Galerkin matrix is pentadiagonal with five nonzero elements in each row. The corresponding $TB_n(y)$ matrix has only nine elements in each row. We conclude that rational Chebyshev series will give banded matrices whenever Hermite functions or other competitors can do the same, but the bandwidth may be either wider or smaller, depending on the differential equation.

For the "Yoshida jet" discussed in Section 7—an inhomogeneous parabolic cylinder equation—there is no contest, however. The Hermite series converges very, very slowly while the $TB_n(y)$ expansion converges exponentially fast. The rational Chebyshev series is clearly the best method for this problem because it simultaneously offers both bandedness and rapid convergence.

5. Series for Functions Which Decay Algebraically as $|y| \to \infty$

Hermite and sinc expansions are the main competitors to $TB_n(y)$ series for solving problems on an infinite interval. As explained in Boyd [7] and Stenger [12], however, Hermite and sinc series converge very poorly when the function $u(y)$ decays slowly as $|y| \to \infty$.[3] The coefficients decrease as algebraic functions of n when the asymptotic behavior of $u(y)$ is algebraic decay with $|y|$. In contrast, when the function decays exponentially fast for large $|y|$, so too do the coefficients a_n.

As we have already seen in Section 2, however, rational Chebyshev expansions are different: it is possible for the $TB_n(y)$ series to converge exponentially—perhaps even geometrically—for a broad class of functions with slow asymptotic decay including some that asymptote to a nonzero constant (no decay!). We have already shown that *geometric* convergence will be the norm when all the singularities of the function are confined to a *finite* portion of the complex y-plane which does not include *any* part of the *real y-axis including* ∞.

Unfortunately, these conditions are quite restrictive. So simple an example as

$$u(y) \equiv y \tag{5.1}$$

is excluded because it is singular at ∞. (Proof: change variable to $\zeta = 1/y$ and inspect $u(\zeta)$ at $\zeta = 0$, which shows that (5.1) has a simple pole at $y = \infty$.) Fortunately, there is an important class of functions which are *weakly* singular at ∞, and yet still exhibit *subgeometric* exponential convergence.

Theorem 3 (*Convergence of the Fourier series for an algebraically decaying or asymptotically constant function*).

If a function $u(y)$ has the inverse power expansion

$$u \sim c_0 + c_1/y + c_2/y^2 + \cdots \tag{5.2}$$

as $y \to \infty$ and a similar series as $y \to -\infty$, then the coefficients of its representation as a Fourier series in the new coordinate t where $y = L \cot(t)$,

$$u(y) = \sum_{n=0}^{\infty} a_n \cos(nt) + \sum_{n=1}^{\infty} b_n \sin(nt), \tag{5.3}$$

[3] Summability methods can cure this slow convergence for Hermite series as shown by Boyd and Moore [21].

will have exponential *convergence in the sense that the* $|a_n|$ *and* $|b_n|$ *decrease with n faster than any finite inverse power of n.*

We may distinguish two cases. First, if the series (5.2) *is* convergent *and* $u(y)$ *is free from singularities for all real y, then the series* (5.3) *converges* geometrically. *Second, if the series* (5.2) *is* asymptotic *but divergent, then* $u(y)$ *is* singular *at infinity, but only weakly in the sense that all derivatives at infinity are bounded.* (*In mathematical jargon,* $u(y)$ *is* C^{∞} *but not* C^{ω} *on the real axis.*) *The coefficients* a_n *decrease at a* subgeometric *rate and the series in* (5.3) *is convergent only when y is real.*

Proof. One can prove that the coefficients of a Fourier series decrease at least as rapidly as n^{-k} by applying k integration-by-parts to the usual integrals that define the coefficients as explained in [5]. It is not necessary that $u(t)$ be holomorphic everywhere on $[0, \pi]$; it is sufficient that (i) the first k derivatives exist and (ii) certain boundary conditions, discussed later, are satisfied. If the integration-by-parts procedure can be applied an arbitrary number of times, then the convergence is exponential. Some references use "infinite order" as a synonym for "exponential" to denote that the convergence is faster than n^{-k} for any finite power of k.

The coefficients of (5.2) are the derivatives of $u(1/y)$ evaluated at $y=\infty$, so the existence of the series implies the boundedness of these derivatives to all orders. To apply the integration-by-parts argument to prove the theorem, we need merely show that the boundedness of the derivatives with respect to $1/y$ implies the boundedness of t-derivatives when we make the change of variable $y = L\cot(t)$ to convert (5.3) into a series in $\cos(nt)$. However, if we substitute the convergent Laurent series for $\cot(t)$ into the inverse power series in y, we obtain

$$u(t) \sim d_0 + d_1 t + d_2 t^2 + \cdots. \tag{5.4}$$

Since the leading term in $\cot(t)$ is $1/t$, $y^{-(k+1)} \sim O(t^{k+1})$ for any k. It follows that d_k is a linear combination of $c_0, c_1, \ldots, c_k$. Since these c_n are all bounded, it follows that d_k will also be finite. This in turn implies that the kth derivative of $u(t)$ at $t=0$ is bounded; the same argument can be applied at the other endpoint, $t=\pi$, so that one can integrate by part k times. Extending this argument to arbitrarily large k proves the theorem. Note that we have not proven that the a_n have the asymptotic form represented by (2.6) or determined the "exponential index of convergence" r; this must be done case-by-case using steepest descent methods [2]. What we have shown is simply that coefficients must decrease faster than any finite inverse power of k. Q.E.D.

There is still a subtle problem: even if all the necessary integrals exist, the integration-by-parts can be blocked if the boundary terms do not all vanish. It is this difficulty—which arises only for *algebraically* decaying $u(y)$—which requires a *general* Fourier series in Theorem 3 even though a *cosine* series is always sufficient for functions that decay *exponentially* as $|y| \to \infty$. When $u(y)$ satisfies certain symmetry properties, it is possible to simplify the general Fourier basis set as described by the following.

THEOREM 4 (Symmetry relations). (i) *If the asymptotic expansion of* $u(y)$ *as in* (5.2) *contains only* even *powers of* y [*or if* $u(y)$ *decays exponentially fast with* $|y|$], *then only* cosine *functions are needed to expand* $u(y[t])$ *as a Fourier series. If the asymptotic series in* $1/y$ *contains only* odd *powers of* y, *then a Fourier* sine *expansion is necessary and sufficient.*

(ii) *If the function* $u(y)$ *is* symmetric *about* $y=0$, *that is,* $u(y)=u(-y)$, *then only the* even cosines ($\cos[2nt]$, $n=0, 1,\ldots,$) *and* odd sines ($\sin[(2n+1)t]$, $n=0, 1,\ldots,$) *are needed. If* $u(y)$ *is* antisymmetric *about the origin, that is,* $u(y)=-u(-y)$, *then all the coefficients in the general Fourier series* (5.3) *for* $u(y[t])$ *will vanish except for those multiplying the* odd cosines *and the* even sines, *i.e.,* $\cos([2n+1]\,t)$ *and* $\sin(2nt)$ *for* $n=0, 1,\ldots,$ *in both cases.*

Proof. It is easy to show [8] that all cosines are symmetric about $t=0$ and that the even cosines are symmetric about $t=\pi/2$ [equivalent to $y=0$] while the odd cosines are antisymmetric about $t=\pi/2$. Similarly, the sines are all antisymmetric about $t=0$, but the odd sines are symmetric about $t=\pi/2$ while the even sines are antisymmetric about this point. These symmetries are important because one can easily prove, directly from the integrals that define the coefficients of a Fourier series, that the coefficients must vanish unless the basis function shares the symmetry with $u(y)$. For example,

$$b_{2n} \equiv (1/\pi) \int_{-\pi}^{\pi} \sin(2nt)\, u(y[t])\, dt \tag{5.5}$$

is zero for all n if $u(y[t])$ is symmetric about $t=0$ because the integrand is antisymmetric about $t=0$ (thanks to the sine) so that the portion of the integral for $t>0$ will automatically cancel that for $t<0$.

The first part can be proved by applying the mapping: even powers of $1/y$ translate into a power series about $t=0$ which also contains only even powers, implying that $u(y[t])$ must be symmetric about $t=0$. The second part of the theorem then follows immediately by recalling that $y=0$ maps into $t=\pi/2$. Q.E.D.

In physics, Theorem 4 would be described as a collection of "parity selection rules." Table III illustrates these rules by giving a simple example of each the four classes of functions which satisfy symmetry rules of both types. However, (i) and (ii) are completely independent: a function whose asymptotic expansion contains only even powers of $1/y$ can be expanded as a cosine series even if it has no definite parity with respect to $y=0$. The only consequence of the lack of symmetry about the origin is that we must keep all the cosines for such a function.

Of course, the Fourier expansion is such a powerful tool that it is possible to expand a function even in terms of a basis set with a different symmetry. For example,

$$\sin(2t) = -(8/\pi) \sum_{n=0}^{\infty} \{1/[(2n+3)(2n-1)]\} \cos([2n+1]\,t). \tag{5.6}$$

TABLE III

Examples of Functions which Asymptote to a Constant or Decay Algebraically with y

$u(y[t])$	$u(y)$	Asymptotic form	Parity with respect to $y=0$	Code	Basis set
$\cos(t)$	$y/(1+y^2)^{1/2}$	$\sim y/\lvert y\rvert + O(1/y^2)$	$u(y) = -u(-y)$	$A\&E$	$TB_{2n-1}(y)$
$\cos(2t)$	$(y^2-1)/(y^2+1)$	$\sim 1 + O(1/y^2)$	$u(y) = u(-y)$	$S\&E$	$TB_{2n}(y)$
$\sin(t)$	$1/(1+y^2)^{1/2}$	$\sim 1/\lvert y\rvert + O(1/\lvert y\rvert^3)$	$u(y) = u(-y)$	S&O	$SB_{2n+1}(y)$
$\sin(2t)$	$2y/(1+y^2)$	$\sim 2/y + O(1/y^3)$	$u(y) = -u(-y)$	A&O	$SB_{2n}(y)$

Note. These illustrate the four classes of functions that have both symmetry with respect to $y=0$ (denoted by S for symmetric and A for antisymmetric in the column labeled "Code") and also have asymptotic expansions which contain only even or only odd powers of $1/y$ (indicated by E or O in the "Code" column). The third and fourth columns give the mathematical forms of these symmetries. The rightmost column indicates the restricted basis set that is sufficient to represent all $u(y)$ that fall into this symmetry class.

However, in contrast to the exponential convergence predicted by Theorem 3, the [illegible] n is the usual price for ignoring the symmetry relationships. It is to avoid this, and recover exponential convergence, that we should be prepared to use a *general* Fourier series, as opposed to a cosine expansion, for functions which decay slowly with y.

The addition of sines to our trigonometric basis set implies that in terms of the original coordinate y ($\varepsilon\ [-\infty, \infty]$), we must augment the $TB_n(y)$ by a second set of basis functions that we shall denote $SB_n(y)$. For the sake of mathematical completeness, we shall give an explicit expression for these new functions. Since the Chebyshev polynomials of the second kind satisfy the identity

$$U_n(\cos t) = \sin[(n+1)\,t]/\sin(t) \tag{5.7}$$

and since one can easily show that

$$\sin(t) = L/(L^2+y^2)^{1/2} \tag{5.8}$$

from the definition of the mapping, $y = L\cot(t)$, it follows that the new basis functions are

$$SB_{n+1}(y) \equiv \{L/(L^2+y^2)^{1/2}\}\ U_n[y/(L^2+y^2)^{1/2}] \tag{5.9a}$$

$$= (1-x^2)^{1/2}\ U_n(x) \tag{5.9b}$$

$$= \sin([n+1]\,t). \tag{5.9c}$$

The recommended procedure, however, is still to change to the trigonometric variable t and then apply Fourier series.

Christov [13] and Higgins [14] defined rational basis sets which are equivalent to the $SB_n(y)$ for even n as proved in Boyd [3]. Unfortunately, functions of y which are equivalent to the odd sines in t are not included in their expansions, but their work and [3] develop general methods for manipulating the basis functions directly without using the machinery of the change of variable plus trigonometric identities. The latter approach, however, seems easier.

The class of functions covered by Theorems 3 and 4 considerably broadens the applicability of spectral methods. However, there are still many classes of functions which cannot be efficiently represented by rational Chebyshev expansions because the asymptotic expansion for $|y| \to \infty$ either (i) involves *fractional* powers of y or (ii) multiplication by a transcendental function. For example,

$$J_0(y) \sim [2/(\pi y)]^{1/2} \cos(y - \pi/4), \qquad |y| \to \infty \tag{5.10}$$

illustrates both difficulties. Fractional powers can sometimes be eliminated by dividing out a common factor or by changing variables. The cosine in (5.10), however, shows that the Bessel function has an infinite number of roots on the real axis. It is not possible to mimic such behavior in a uniform fashion over the whole interval unless the function decays exponentially as $|y| \to \infty$ so that the oscillations for large y can simply be ignored without the penalty of a large absolute error.

It follows that it is unreasonable to expect exponential convergence for an example like (5.10) for *any* series of discrete basis functions on $y \in [-\infty, \infty]$. It is possible to obtain a "global" approximation for (5.10) as shown by Boyd [22], but only by computing *two* series, one for the amplitude and one for the phase, in a manner that mimics the asymptotic WKB approximation for this function. The $TB_n(y)$ are effective for as large a class of functions as is possible for any single basis set.

At the risk of a bit of redundancy, we summarize the theorems and conclusions above with the following.

THEOREM 5 (Expansion in orthogonal rational functions). (A) *The $TB_n(y)$ will give exponentially rapid convergence for all functions $u(y)$ that are free of singularities for real y (except perhaps at ∞) and have one of the following properties*:

(i) *$u(y)$ decays faster than $1/|y|^k$ for arbitrarily large k as $|y| \to \infty$ ["exponential decay"] or*

(ii) *$u(y)$ has an asymptotic power series for $|y| \gg 1$ that contains only EVEN, nonnegative powers of $1/y$.*

(B) *The $SB_n(y)$ will give exponentially rapid for all $u(y)$ that are holomorphic on the analytic y-axis and either*:

(i) *has exponential decay with $|y|$, or*

(ii) *$u(y)$ has an asymptotic power series with only* ODD *positive powers of $1/y$.*

(C) *The $TB_n(y)$ and $SB_n(y)$ together give exponential convergence when*

(i) *$u(y)$ decays exponentially with $|y|$, or*

(ii) *has an asymptotic power series that contains only* NONNEGATIVE, INTEGRAL *powers of $1/y$.*

Both basis sets have the property that the even-numbered basis functions are symmetric about $y=0$ while the odd basis functions are antisymmetric, so the basis set can be halved if $u(y)$ has definite parity with respect to the origin.

When $u(y)$ decreases exponentially fast as $|y| \to \infty$, the $TB_n(y)$ are always sufficient.

(D) *Both basis sets* fail *(in the sense that their series converge algebraically rather than exponentially with N) if $u(y)$ decays as an* algebraic *(rather than exponential) function of $1/|y|$* and *if in addition either:*

(i) *the asymptotic approximation contains* nonintegral *powers of y or*

(ii) *the function has an infinite number of roots on the real y-axis.*

6. Natural Versus Essential Boundary Conditions

Most texts on the finite element method make a careful distinction between "essential" boundary conditions, which must be imposed on each individual basis function, and "natural" boundary conditions, which do not require modifying the expansion functions. When the boundary condition is "natural," the differential equation itself will *force* the boundary condition to be satisfied by the sum of the basis functions, at least approximately, even though this constraint is not explicitly imposed on the numerical solution in any way.

The same distinction also arises when solving differential equations through spectral or pseudospectral methods. When the differential equation is *singular* at the endpoints of the interval, the usual boundary conditions are that the solution should be analytic at the endpoints, and it is not necessary to constrain the basis functions. The pseudospectral algorithm will automatically pick out that solution which is smooth and analytic.

Boyd [15] gives a good discussion of "natural" boundary conditions in spherical geometry. The merging of the coordinate surfaces at the north and south poles implies that differential equations in spherical coordinates are always singular at the poles; but because these singularities are artifacts of the coordinate system, not the physics, the solution is well behaved and holomorphic at the poles. If one expands a function of latitude and longtitude as

$$u(\lambda, \theta) = \sum_{m=-\infty}^{\infty} F_m(\theta)\, e^{\mathrm{i}m\lambda} \tag{6.1}$$

one can show that $F_m(\theta)$ must have an $|m|$th order root at both poles. The spherical harmonics individually satisfy this root condition, which is one reason why they are an obvious and efficient basis set for spherical problems. However, it is quite unnecessary to impose these constraints on the basis. Boyd [15] shows that one can represent the lowest $m=49$ spherical harmonic using twenty unconstrained cosine functions of the proper symmetry. Although the spherical harmonic has 98 zeroes (all at the poles) and the trigonometric polynomial approximating it cannot have more than 38 roots, the approximation is nonetheless accurate to 8 decimal places. For this case, we have not only failed to impose boundary conditions on the basis set, but we have used a series truncated so ruthlessly that it cannot possibly satisfy all the constraints imposed by the differential equation on the exact solution. It does not matter: the blind use of a small number of trigonometric functions still gives exceptional accuracy.

The situation for an infinite interval, $y \in [-\infty, \infty]$, is very similar. Indeed, the digression into spherical coordinates is quite relevant because problems in colatitude θ may be recast into differential equations on an infinite interval by making the change of variable

$$\tanh y \equiv \cos(\theta) \qquad \text{[Mercator coordinate]} \tag{6.2}$$

which is popular in cartography. The boundary conditions remain "natural" conditions even after the shift to the Mercator coordinate. In the next section, we shall give some numerical examples of Legendre and associated Legendre functions of $\tanh(y)$ which were calculated using $TB_n(y)$ as the basis without any explicit imposition of boundary conditions on the numerical solution. Just as on the sphere, the singularities of the differential equation force the pseudospectral method to choose the correct solution.

One would like to be able to state this as a theorem: since the coefficients of a differential equation on an unbounded interval are usually singular at infinity, the boundary conditions at infinity can always be ignored in setting up the spectral or pseudospectral matrix. Unfortunately, this is probably not a universal truth, but merely a statement that is true in most practical applications.

In the first place, we note that although most differential equations are singular at infinity, it is easy to contrive examples which are not. For instance, Morse and Feshbach [6, p. 537] showed that

$$u_{yy} + [2/(y-i)]\, u_y - [1/(y-i)^4]\, u = f(y) \tag{6.3}$$

has the general solution

$$u = A \exp[1/(y-i)] + B \exp[-1/(y-i)] + u_p(y), \tag{6.4}$$

where $u_p(y)$ denotes the particular integral. The homogeneous solution will be holomorphic and finite for all real y including ∞. Unless we specifically impose boundary conditions at infinity, there is no way to determine the arbitrary constants A and B. We can add arbitrary multiples of either homogeneous solution to u without destroying the analyticity or boundedness of $u(y)$.

Equation (6.3) is admittedly artificial, but Stenger [12], Boyd [9], Lund and Riley [16], and Bowers and Lund [17] have discussed physical examples that require "essential" boundary conditions at infinity. These equations are actually defined on a *finite* interval, and therefore are seemingly irrelevant to the unbounded domain which is the theme of this paper, but all have bounded endpoint singularities which cause very slow convergence if $u(x)$ is approximated as an ordinary Chebyshev series. The remedy, first proposed by Stenger [12], is to use an exponential mapping to transform the finite interval to an infinite domain and then apply Hermite, sinc, or $TB_n(y)$ expansions. One important class of such problems are partial differential equations with singularities at the corners of a rectangular domain: the boundary conditions on the walls of the domain must still be explicitly imposed even after the domain is stretched to infinity. Example Four of Section 7 is exceptional in that the boundary conditions are natural because the solution vanishes at the singularities.

The author has not been able to prove or to find in the literature a general theorem for differential equations whose coefficients are singular at infinity. Some singular differential equations have solutions that cannot be obtained without altering the basis set.

The differential equation for the spherical harmonics is such an example: but only on the *sphere*. One can show that when m is *odd*, the associated Legendre functions have square root singularities at the poles in the variables $x \equiv \cos\theta$; in other words, *both* solutions of the differential equation in x are singular at the endpoints. The remedy is to define new basis functions ϕ_n by multiplying each of the old set by this square root factor:

$$\phi_n(x) \equiv (1-x^2)^{1/2}\, T_n(x) \qquad [\text{in } x] \tag{6.5a}$$

or equivalently in colatitude θ,

$$\phi_n(\theta) \equiv \sin(\theta)\cos(n\theta) = \tfrac{1}{2}\{\sin[(n+1)\,\theta] + \sin[(n-1)\,\theta]\}. \tag{6.5b}$$

On $y \in [-\infty, \infty]$, however, the transformed versions of the spherical harmonics have natural boundary conditions for all m: the mapping seems to greatly weaken the singularities. Thus, even this case is not a counterexample on the *infinite* interval.

Of course, if a solution is known to vanish rapidly at infinity, one may impose endpoint zeros on the basis functions as an option. In theory, the result is better accuracy with a given number of degrees of freedom since the basis is more closely matched to the solution. In practice, one must be careful. An exponentially decaying function has in effect an infinite order zero at both $\pm\infty$, so one can impose zeros of arbitrarily high order on the basis functions. If this is carried to an extreme, however, the basis functions will be ill-conditioned because they are indistinguishable from one another—all zero to within machine precision—over a large portion of $y \in [-\infty, \infty]$. Boyd [15] describes this problem for equations on the sphere. In consequence, this article will discuss only unmodified rational Chebyshev functions.

In the next section, we give five examples (including one solution that decays algebraically with y) for which an unconstrained basis set is a success. Because boundary conditions at infinity are "natural" in *most* applications, we usually do not have to modify the spectral or pseudospectral matrix by using some rows to impose explicit boundary conditions.

7. Numerical Examples

The numerical illustrations have been chosen to reiterate the main points of this work. Since one theme is the importance of symmetry, all the solutions are either symmetric or antisymmetric with respect to $y=0$. All references to N give the number of basis functions with the same symmetry as that of the solution; we need not include the other basis fuctions because their computed coefficients would be zero. To avoid a discussion of time-stepping methods, all five examples are boundary value problems in one dimension. To avoid the distractions of the Newton–Kantorovich method [18], all are linear.

Although we described Galerkin's method in Section 4, these sample problems were solved using the alternative of the "pseudospectral" or "collocation" method [5, 18, 19] because this is simpler to program. The truncated series of $TB_n(y)$ is substituted into the differential equation. The undetermined series coefficients a_n are obtained by demanding that the residual be zero at each of N points ("collocation" or "grid" points). This converts the boundary value problem into an $N \times N$ matrix equation which is solved by Gaussian elimination. As explained in [5] and [19], the pseudospectral method gives accuracy comparable to Galerkin's if the collocation points are chosen to be those of the corresponding Gaussian quadrature. In our case, this means the points should be evenly spaced in the trigonometric argument t (which implies an uneven grid in y). It is important, however, to choose that grid which includes only *interior* points on $t \in [0, \pi]$ because the differential equation is usually singular at the endpoints.

When the differential equation is of definite parity $[u(y)=u(-y)$ or $u(y)=-u(-y)]$, the basis should be halved to included only those $TB_n(y)$ of the same symmetry as the solution, and the collocation points should be evenly spaced on $t \in [\pi/2, \pi]$. The collocation conditions on the other half interval will then automatically be satisfied because of the symmetry of the differential equation and of the functions in the reduced basis set.

As explained in Section 6, it is sometimes necessary to impose explicit boundary conditions at infinity. For most problems on an infinite interval, however, the boundary conditions are "natural." To emphasize this, all the problems were solved using an unconstrained series of basis functions.

When we apply the collocation method to solve a differential equation, there are two sources of error. The "truncation" error is the sum of all the neglected higher coefficients of the infinite series for $u(y)$,

$$E_T(N) \equiv \sum_{n=N+1}^{\infty} a_n TB_n(y) \qquad \text{[truncation error]}. \tag{7.1}$$

In addition, the neglect of the higher coefficients causes the computed coefficients, $\bar{a}_n$, to differ from those of the infinite series. (Note that we obtain the $\bar{a}_n$ not by evaluating inner product integrals but rather by solving a *truncated* matrix equation.) This "discretization" error is

$$E_D(N) \equiv \sum_{n=0}^{N} (a_n - \bar{a}_n)\, TB_n(y) \qquad \text{[discretization error]}. \tag{7.2}$$

In most of the graphs below, we shall merely consider the sum of these two errors. Empirically, experience with problems on finite intervals has shown that the two are of roughly the same order of magnitude. For the Yoshida jet, we present a graph to show that the same seems to be true for equations on an infinite interval: the discretization error and truncation error are roughly equal, so estimates of the latter can be simply be doubled (as in [2]) to obtain a crude estimate of the total error.

EXAMPLE 1.

$$u_{yy} - y^2 u = -\exp(-0.5y^2). \tag{7.3}$$

The exact solution is

$$u(y) \equiv \exp(-0.5y^2) \tag{7.4}$$

Boyd [2] describes the role of the map parameter L for efficiency and offers simple, analytic formulas for estimating the optimum L for certain classes of problems. Note that exactly half of the (unevenly spaced) grid points will lie on $y \in [-L, L]$ and the rest satisfy $L \leqslant |y| < \infty$. Figure 4 illustrates the sensitivity of the numerical solution to various L. Even with only four symmetric basis functions, we obtain moderate accuracy for all $L \in [1, 5]$.

Thus, it is silly to attempt to make heroic efforts to optimize L; any choice that is of the same order-of-magnitude as the best L will give good results. When one needs to make many calculations for different parameter values, one can estimate L through a single set of preliminary runs with fixed N and fixed parameters and variable L by inspecting the rate at which the computed series coefficients decrease. Although the exact solution is usually not known, one can make a single, expensive calculation with high N to compare with the set of runs with small N and varying L to obtain a better estimate of L_{optimum}. The error varies slowly with L in the vicinity of the most efficient L as evident in the flat minima in Fig. 4, so a crude estimate of the map parameter is sufficient.

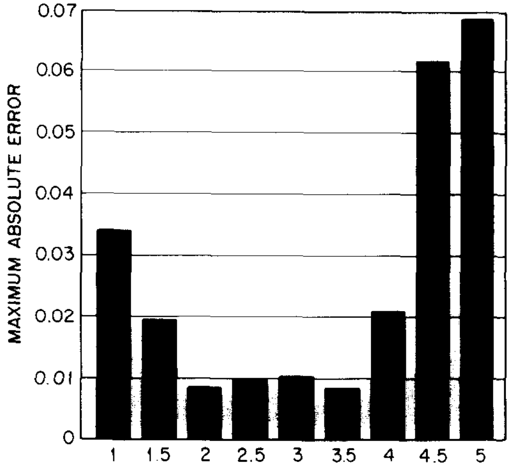

FIG. 4. The maximum pointwise errors on $y \in [-\infty, \infty]$ for Example 1 [exact solution: $u = \exp(-0.5y^2)$] for various values of the map parameter L. The truncation included only the lowest four symmetric basis functions: (TB_0, TB_2, TB_4, and TB_6).

Figure 5 compares the four-term and five-term collocation solution of (7.3) with the exact solution. The accuracy is impressive; the six-term approximation is graphed, too, but it is invisible because the difference between it and the exact solution is smaller than the thickness of the curves.

EXAMPLE 2 (North–south current of the steady part of the "Yoshida jet" in equatorial oceanography [20].)

$$v_{yy} - y^2 v = y. \tag{7.5}$$

Exact solution: infinite series of Hermite functions given in [21].

Because the solution of (7.3) decays exponentially as $|y| \to \infty$, Hermite function expansions and sinc series also converge exponentially fast for Example 1. (The Hermite series of a Gaussian is just one term!) The solution to (7.5), however, has the asymptotic expansion [19],

$$v \sim -(1/y)\{1 + 2/y^4 + 60/y^8 + \cdots\}, \qquad |y| \to \infty \tag{7.6}$$

which shows that the Yoshida current decays *algebraically* with $|y|$ for large y. This slow decay wrecks the convergence of the Hermite and sinc series; the Hermite coefficients decrease as $O(n^{-3/4})$!

In contrast, the rational Chebyshev expansion converges at a subgeometric but exponential rate. As explained in Section 5, the asymptotic series (7.6) indicates that a series of the usual $TB_n(y)$ [which are equivalent to cosines of t when we make the change of variable $y = L\cot(t)$] converge poorly for the Yoshida flow, but the alternative functions $SB_n(y)$, which are the images of $\sin(nt)$ under the map, give

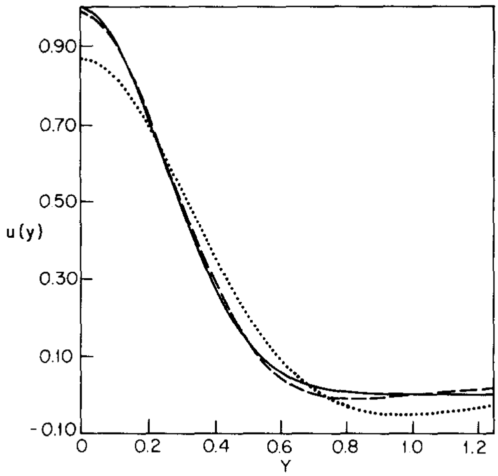

FIG. 5. Comparison of exact and approximate $u(y)$ for Example 1 (Gaussian solution) with map parameter $L=4$. Exact [solid line], 4 symmetric basis functions [dotted], and 5 symmetric basis functions [dashed] are shown. The numerical solution using 6 collocation points is also graphed, but is indistinguishable from the exact $u(y)$ to within the thickness of the curve.

excellent results. Since (7.6) contains only *odd* powers of $1/y$, the "parity selection rules" of Theorem 4 part (ii), tell us that we need only the *even* sine funtions, $\sin(2nt)$, to represent $v(y[t])$.

The results are shown in Fig. 6: the two-term approximation is not bad, and the seven-term approximation is graphically indistinguishable from the exact solution. The three-term truncation of the large N solution can be summed to give

$$v \approx -y\,(440.8017 + 15.098y^2 + 1.1412y^4)/(9+y^2)^3 \tag{7.7}$$

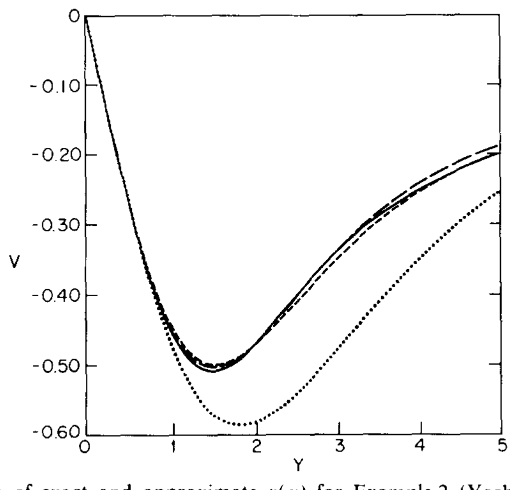

FIG. 6. Comparison of exact and approximate $v(y)$ for Example 2 (Yoshida jet) with $L=3$. The exact solution is the solid line, and the approximations with two, three, and four antisymmetric basis functions are shown as the dotted, long-dashed, and short-dashed curves, respectively.

which has a maximum absolute error of only 0.006 on $y \in [-\infty, \infty]$. Hermite functions are the "obvious" way to solve (7.5) since they are the eigenfunctions of the differential operator; the coefficients of the infinite Hermite series can be evaluated in closed form. Because the Hermite expansion converges so poorly,

lacks the additional discretization error of the pseudospectral solution computed with three collocation points. Figure 7 compares the $N = 1$ (two-term) truncation of the infinite series for v (as calculated using large N) with the results of the pseudospectral method using two collocation points. Table IV shows how the pseudospectral coefficients converge to those of the exact $v(y)$ as N increases.

EXAMPLE 3 (Legendre's differential equation transformed from $X \in [-1, 1]$ to $y \in [-\infty, \infty]$.) Letting $P_n(X)$ denote the usual Legendre polynomial,

$$u_{yy} + [n(n+1)\operatorname{sech}^2(y) + 1]\,u = P_n(\tanh[y]). \tag{7.8}$$

Exact solution: $P_n(\tanh[y])$.

As noted in Section 6, "natural" boundary conditions are the norm on the sphere; the same is true when Legendre's equation, with singularities at $X = \pm 1$, is transformed to an unbounded interval by setting $y = \tanh^{-1}(X)$. Figure 8 illustrates the result for P_{12} $(\tanh[y])$. Because this solution has a dozen roots and thus more structure than the first two examples, more terms are needed: even the sixteen-term approximation is rather crude. However, the graph for $N = 20$ is indistinguishable from the exact solution.

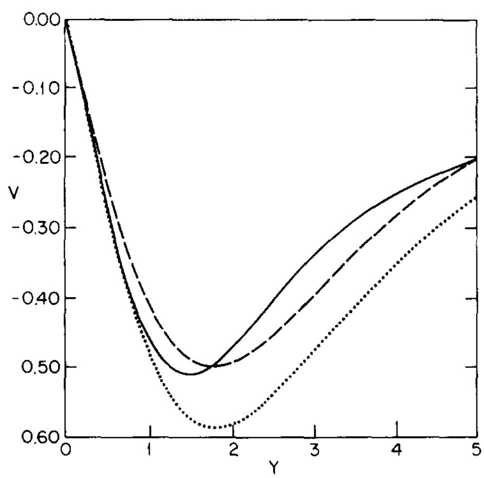

FIG. 7. The exact solution for the Yoshida jet [solid] is compared with the approximation obtained via collocation with two (positive) interpolation points [dotted] and also with the result of truncating the infinite series for $v(y)$ to just two terms [dashed]. The truncated approximation is more accurate because it lacks the "discretization error" of what we obtain by inverting the 2×2 pseudospectral matrix.

TABLE IV

The Coefficients of the Spectral Series for the Yoshida Jet $v(y)$ as Computed Using Various Numbers of Collocation Points

n	2 pts.	3 pts.	4 pts.	7 pts.	21 pts.
1	−0.47769	−0.388976	−0.394428	−0.395329	−0.395324
3	0.199475	0.179639	0.176129	0.179237	0.179224
5	—	−0.050409	−0.04471	−0.051163	−0.051147
7	—	—	0.001399	0.002620	0.002651
9	—	—	—	0.003856	0.003740
11	—	—	—	−0.000541	−0.000592
13	—	—	—	−0.000452	−0.000446
15	—	—	—	—	0.000060

This is quite typical of spectral solutions: the error is large for all truncations up to some limit N that depends on the equation, and then plunges abruptly. Figures 3.7 and 3.8 of [5] give good illustrations of this. It is only after the series has begun to converge, that is, only after the error has begun its abrupt drop, that the asymptotic definitions we gave earlier—"exponential" versus "algebraic" convergence, "geometric" versus "subgeometric"—have any meaning. However, as stressed in [5] and [19], even if only a moderate error of say 5% is needed, pseudospectral methods still are much more efficient than finite difference techni-

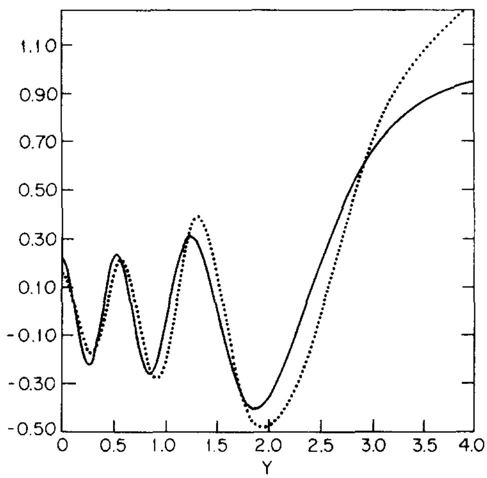

FIG. 8. Exact solution (solid) compared with the sixteen-term approximation (dotted) to Example 3: $P_{12}[\tanh[y)]$. Because this mapped Legendre polynomial is symmetric about $y = 0$, only the functions $TB_{2n}(y)$, $n = 0, 1, \ldots, 15$ were used in the approximation, and the sixteen collocation points were all at positive y. The approximation with twenty collocation points is graphically indistinguishable from the exact solution.

ques. It is quite absurd to suppose that a three grid point approximation to the Yoshida jet would yield the error of only 2.5% achieved by the three-point pseudospectral method, or that it would easily generate an *analytic* approximation like (7.7).

EXAMPLE 4 (Removal of endpoint singularities via mapping).

$$u_{yy} + 2\,\mathrm{sech}^2(y)\,u = \mathrm{sech}(y). \tag{7.9}$$

Exact solution.

$$u(y) \equiv \mathrm{sech}(y) = [1 - \tanh^2(y)]^{1/2} = P_1^1(\tanh[y]). \tag{7.10}$$

Equation (7.9) is merely a particular case of the differential equation satisfied by the associated Legendre functions $P_n^m(X)$ after the transformation $X = \tanh(y)$ has been applied. What is special about this example is that

$$P_1^1(X) \equiv (1 - X^2)^{1/2} \tag{7.11}$$

so that the solution is *singular* at the endpoints of the original integration interval in X. If we expand (7.11) as a series of Chebyshev polynomials in X, we find

$$(1 - X^2)^{1/2} = (2/\pi)\left\{1 - 2\sum_{n=1}^{\infty} [1/(4n^2 - 1)]\, T_{2n}(X)\right\}. \tag{7.12}$$

The singularity spoils the convergence of the series so that the coefficients decrease as $O(1/n^2)$: *algebraic* convergence.

Stenger [12] pointed out that the map $X = \tanh(y)$ would heal the singularity: the transformed square root function is $\mathrm{sech}(y)$, which decays exponentially fast, and can therefore be efficiently expanded in any of several basis sets. Stenger himself used sinc functions, but Boyd [9] points out that the tanh–mapping is the key, not the choice of basis set. Table V shows that the $TB_n(y)$ series also has coefficients a_n in

$$(1 - X^2)^{1/2} = \sum_{n=0}^{\infty} a_{2n}\, TB_{2n}(\mathrm{arctanh}[X]) \tag{7.13}$$

that decrease exponentially with n; $a_{2n} \sim O[\exp(-qn^{1/2})]$ for some q as shown in Boyd [2].

More important, this example shows that the mapping also generates a differential equation with smooth coefficients. The derivatives of P_1^1 with respect to X are not even bounded at the endpoints; the term-by-term derivative of the series in (7.12) is divergent. However, all the derivatives with respect to y of P_1^1 (tanh[y]) tend to 0 as $|y| \to \infty$ so that one can prove that the $TB_n(y)$ series for the derivatives converge exponentially fast with n.

Of course, this example is contrived: the proper way to deal with the square root

TABLE V

The Coefficients for Two Representations of the Function $f(X) = (1 - X^2)^{1/2}$

n	a_n	b_n
0	0.324659	0.6366
2	−0.470943	−0.4244
4	0.168566	−0.0849
6	−0.026329	−0.0364
8	0.006914	−0.0202
10	−0.002437	−0.0129
12	−0.000119	−0.0089
14	−0.000300	−0.0065
16	−0.000046	−0.0050
18	−0.000009	−0.0039
20	0.000015	−0.0032
22	0.000015	−0.0026
24	0.000011	−0.0022
26	0.000007	−0.0019
28	0.000004	−0.0016
30	0.000002	−0.0014

Note. The a_n are the coefficients of the series in $TB_n(\text{arctanh}[X])$ as computed by applying the pseudospectral method with 16 points to the differential equation. The b_n are the known exact coefficients of the expansion in $T_n(X)$. The maximum pointwise error in the Chebyshev expansion is approximately $1/(\pi N)$, which is roughly 1/50 to the order shown; the error in the rational Chebyshev approximation is about 1/300,000.

singularity of the associated Legendre functions for all odd zonal wavenumbers m is to multiply each basis function by the square root. For more complicated singularities, however, it may be impossible to multiply or divide out the branch point. Stenger's mapping of the finite interval to an infinite interval is then a real godsend (see [16] and [17]).

The classic textbook example is

$$\nabla^2 u = -1; \qquad u = 0 \quad \text{on all sides of the unit square}$$

which has weak logarithmic singularities at the corners [11] even though the equation is both linear and constant coefficient.

Example 5.

$$u_{yyyy} + 16u = 16\,\text{sech}(2y). \tag{7.14}$$

This example shows that boundary conditions at infinity can be "natural" rather than "essential" even for a *fourth* order differential equation. As shown in Fig. 9, the $TB_n(y)$ series still gives very high accuracy with a modest number of terms.

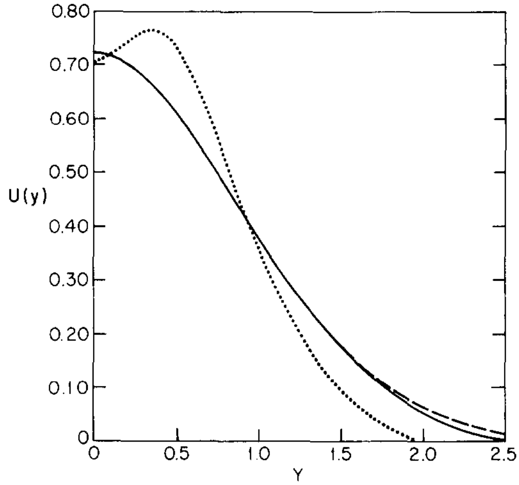

FIG. 9. Exact solution (solid) compared with the four-collocation-point (dotted) and seven-point (dashed) approximations for Example 5: fourth-order differential equation ($L = 1$).

8. Summary and Conclusions

Expansions in orthogonal rational functions have many virtues. As shown by Table II and Section 3, it is easy to evaluate their derivatives (to solve differential equations) merely by taking linear combinations of the derivatives of $\cos(nt)$. When, as is usually the case (Sect. 6), the boundary conditions are "natural" rather than "essential," it is easier to apply the pseudospectral method with these functions than with Chebyshev polynomials because it is unnecessary to use rows of the matrix to explicitly impose the boundary conditions. Galerkin's method is sometimes a very efficient alternative to the pseudospectral method for the diffusive part of a "splitting" or "fractional" steps time integration because the orthogonal rational functions yield a banded Galerkin matrix whenever the differential equation has polynomial or rational coefficients (Sect. 4). Most of these new concepts carry over, with slight modifications, to equations on a semi-infinite interval, but a full discussion is deferred to a later publication [22].

The basis set we have discussed here is so powerful that it will give an approximation whose error decreases as an exponential function of the truncation N for three broad classes of functions—with some additional tricks for the last two. The first class is that of functions which (i) have no singularities for real y except perhaps at ∞ and (ii) decay exponentially fast with $|y|$ as $|y| \to \infty$ along the real axis. The rational functions we defined as $TB_n(y)$ always give exponentially fast convergence for such functions.

The second class is composed of $u(y)$ which lack exponential decay with $|y|$, but which have asymptotic power series in $1/y$ for $|y| \gg 1$. As discussed in Section 5, there is always some set of orthogonal rational functions that will work. However, it may be necessary to augment the $TB_n(y)$, which are the images of $\cos(nt)$ under the map $y = L\cot(t)$, by additional rational functions, defined by (5.9) and denoted $SB_n(y)$, which are the images of $\sin(nt)$ under the same map. In other words, after transforming the problem from $y \in [-\infty, \infty]$ to $t \in [0, \pi]$, it may be necessary to use a general Fourier series in t instead of the cosine series that is sufficient for $u(y)$ that decay exponentially fast with $|y|$. Table III and Theorem 5 summarize the cases.

The third class of functions is one seemingly irrelevant: solutions on a *finite* interval with bounded singularities at the endpoints. The trick is to use a hyperbolic tangent mapping to stretch the domain to the whole y-axis and *then* apply the rational Chebyshev functions. In terms of the original coordinate, one is using $TB_n(\text{arctanh}[X])$ as a basis on $X \in [-1, 1]$.

Though it is a bit confusing to discuss so many different cases, this is a very welcome confusion because it shows that the $TB_n(y)$ series can be combined with other devices and mappings. In this article, we have tried to provide a good theoretical and empirical foundation and suggest the possibilities of orthogonal rational functions.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation under Grants OCE-8305648 and OCE-8509923.

REFERENCES

1. C. E. GROSCH AND S. A. ORSZAG, *J. Comput. Phys.* **25**, 273 (1977).
2. J. P. BOYD, *J. Comput. Phys.* **25**, 43 (1982).
3. J. P. BOYD, *SIAM J. Numer. Anal.* (1987), in press.
4. J. P. BOYD, *J. Comput. Phys.* **57**, 454 (1985).
5. D. GOTTLIEB AND S. A. ORSZAG, *Numerical Analysis of Spectral Methods: Theory and Applications* (Soc. Ind. and Appl. Math., Philadelphia, (1977).
6. P. M. MORSE AND H. FESHBACH, *Methods of Theoretical Physics, Part II* (McGraw-Hill, New York, 1953), p. 1210,
7. J. P. BOYD, *J. Comput. Phys.* **54**, 382 (1984).
8. M. ABRAMOWITZ AND I. STEGUN, *Handbook of Mathematical Functions*, (Dover, New York, 1965).
9. J. P. BOYD, *J. Comput. Phys.* **64**, 266 (1986).
10. A. B. CAIN, J. H. FERZIGER, AND W. C. REYNOLDS, *J. Comput. Phys.* **56**, 272 (1984).
11. G. STRANG AND G. J. FIX, *An Analysis of the Finite Element Method* (Prentice-Hall, Englewood Cliffs, NJ, 1973), p. 265.
12. F. STENGER, *SIAM Rev.* **23**, 165 (1981).
13. C. I. CHRISTOV, *SIAM J. Appl. Math.* **42**, 1337 (1982).
14. J. R. HIGGINS, *Completeness and Basis Properties of Sets of Special Functions* (Cambridge Univ. Press, Cambridge, 1977).

15. J. P. Boyd, *Monthly Weather Rev.* **106**, 1184 (1978).
16. J. R. Lund and B. V. Riley, *IMA J. Numer. Anal.* **4**, 83 (1984).
17. K. Bowers and J. Lund, *SIAM J. Numer. Anal.* (1985), in press.
18. J. P. Boyd, *Physica* **21D**, 227 (1986).
19. J. P. Boyd, *Monthly Weather Rev.* **106**, 1192 (1978).
20. A. E. Gill, in *Numerical Models of Ocean Circulation* (Nat. Acad. Sci., Washington, D.C., 1975).
21. J. P. Boyd and D. W. Moore, *Dyn. Atmos. Oceans* **10**, 51 (1986).
22. J. P. Boyd, *J. Comput. Phys.* (1987), in press.