JOURNAL OF COMPUTATIONAL PHYSICS **126,** 11–20 (1996)
ARTICLE NO. 0116

Traps and Snares in Eigenvalue Calculations with Application to Pseudospectral Computations of Ocean Tides in a Basin Bounded by Meridians

JOHN P. BOYD

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

Received September 7, 1995; revised January 12, 1996

We make several observations about eigenvalue problems using, as examples, Laplace's tidal equations and the differential equation satisfied by the associated Legendre functions. Whatever the discretization, only some of the eigenvalues of the *N*-dimensional matrix eigenvalue problem will be good approximations to those of the differential equation—*usually* the *N*/2 eigenvalues of smallest magnitude. For the tidal problem, however, the "good" eigenvalues are scattered, so our first point is: It is important to plot the "drift" of eigenvalues with changes in resolution. We suggest plotting the difference between a low resolution eigenvalue and the nearest high resolution eigenvalue, divided by the magnitude of the eigenvalue or the intermodal separation, whichever is smaller. Second, as a final safeguard, it is important to look at the Chebyshev coefficients of the mode: We show a numerically computed "anti-Kelvin" wave which has little eigenvalue drift, but is completely spurious as is obvious from its spectral series. Third, inverting the roles of parameters can drastically modify the spectrum; Legendre's equation may have either an infinite number of discrete modes or only a handful, depending on which parameter is the eigenvalue. Fourth, when the modes are singular but decay to zero at the endpoints (as is true of tides), a tanh-mapping can retrieve the usual exponential accuracy of spectral methods. Fifth, the pseudospectral method is more reliable than deriving a banded Galerkin matrix by means of recurrence relations; the pseudospectral code is simple to check, whereas it is easy to make an untestable mistake with the intricate algebra required for the Galerkin method. Sixth, we offer a brief cautionary tale about overlooked modes. All these cautions are applicable to *all* forms of spatial discretization including finite difference and finite element methods. However, we limit our illustrations to spectral schemes, where these difficulties are most easily resolved. With a bit of care, the pseudospectral method is a very robust and efficient method for solving differential eigenproblems, even with singular eigenmodes.

1. INTRODUCTION

Differential eigenvalue problems are so common that there would seem to be little new to say. The book by Boyd [1] describes a simple procedure [2–4] which, for sufficiently smooth solutions, is guaranteed to give an error which decreases exponentially fast with N, the number of degrees of freedom. The steps are the following. First, choose a set of spectral basis functions $\{\phi_j(x)\}$ that satisfy the boundary conditions: the sines and cosines of a Fourier series for periodic problems, spherical harmonics for problems in latitude and longtitude on a sphere, Chebyshev polynomials for a finite, nonperiodic interval, or rational Chebyshev functions for an infinite interval. Next, approximate the eigenmode $u(x)$ by the N-term truncation of the spectral series with (as yet) unknown coefficients $\{a_j\}$:

$$u(x) \approx u_N(x) \equiv \sum_{j=1}^{N} a_j \phi_j(x). \tag{1.1}$$

Third, substitute $u_N(x)$ into the differential equation and demand that the residual be zero at each of N "collocation" or "interpolation" points $\{x_j\}$, where there is a canonical choice of such points associated with each of the standard basis sets [1]. These collocation conditions translate a linear differential eigenproblem into an N-dimensional matrix eigenvalue problem, where the eigenmode is now a column vector containing the spectral coefficients a_j. The final step is to call library software to solve the matrix eigenvalue problem. If the eigenmodes are also needed, one may evaluate them by simply substituting the appropriate matrix eigenmode as the coefficients of $u_N(x)$ and then summing the series at each point of the graph. If $u(x)$ is nonsingular on the expansion interval, then the error decreases as $\exp(-qN)$ for some constant q. (For infinite interval problems, the error falls as $\exp(-qN^r)$ for some $r < 1$ (typically $\frac{1}{2}$ or $\frac{2}{3}$)).

It all sounds too good to be true. Our theme is that actually, the pseudospectral method *is* as good as advertised—but only if one is a little bit careful. It is a truism in auto safety that most accidents occur within 10 miles of home. Its arithmurgical equivalent is that no calculation is too simple to screw up.

All our cautions apply to finite difference, finite element, and finite volume computations of eigenvalues, too. We use only spectral methods, however, because (i) they are the most efficient and (ii) their exponential accuracy allows us to easily immunize our results against these diseases.

0021-9991/96 $18.00

2. TWO EXAMPLES: LEGENDRE'S EQUATION AND LAPLACE'S TIDAL EQUATIONS

The associated Legendre functions satisfy

$$\{D^2 + V_0(1-\mu^2) + E\}, u = 0, \quad \mu \in [-1, 1], \quad (2.1)$$

where μ is the cosine of colatitude and the differential operator D is

$$D \equiv (1-\mu^2)\frac{d}{d\mu} \quad (2.2)$$

with

$$u = P_n^s(\mu), \quad V_0 = n(n+1), \quad E = -s^2 \quad (2.3)$$

The periodicity of the sphere in longitude demands that s be an integer. With V_0 as the eigenparameter, we see that the Legendre equation is a typical Sturm–Liouville problem with an infinity of discrete eigenmodes.

Through the change of coordinate

$$\mu = \tanh(y/M)) \quad (\text{"Mercator coordinate"}), \quad (2.4)$$

where M is a constant, Legendre's equation can also be written, using subscript "y" to denote differentiation with respect to that coordinate and taking $M = 1$ for simplicity,

$$u_{yy} + \{V_0 \operatorname{sech}^2(y) + E\}u = 0, \quad y \in [-\infty, \infty]. \quad (2.5)$$

This is identical in form with the one-dimensional Schroedinger equation of quantum mechanics with E as the kinetic energy and the potential energy $V(y) = -V_0 \operatorname{sech}^2(y)$. However, E is the eigenvalue and V_0 is a fixed parameter which specifies the strength of the potential.

The reversal of the roles of the parameters changes everything. There is only a finite number of discrete, spatially localized "bound states." If $V_0 = \nu(\nu+1)$ for some positive but otherwise arbitrary ν, the discrete modes, from page 1651 of [22], are

$$E_j = -(\nu - j)^2, \quad j = 0, 1, 2, ..., j_{\max}, \quad (2.6)$$

where $j_{\max}$ is the largest integer smaller than ν. The rest of the spectrum is a *continuous* spectrum: for every positive E, there is an eigenmode which is bounded (but oscillatory rather than decaying) as $|y| \Rightarrow \infty$.

Our second example is the set of three linear equations for atmospheric and oceanic tides due to Laplace [6–10],

$$\begin{aligned} \lambda u - \mu v - s\zeta &= 0 \\ \mu u - \lambda v + D\zeta &= 0 \\ su - Dv - \varepsilon\lambda(1-\mu^2)\zeta &= 0, \end{aligned} \quad (2.7)$$

where λ is the frequency divided by 2 Ω, where Ω is the angular frequency of the earth's rotation, μ is the cosine of colatitude, s the zonal wavenumber, $\varepsilon = 4\,\Omega^2\,a^2/gH$ is "Lamb's parameter," where a is the radius of the earth, g is the gravitational constant, and H is the mean depth of the ocean, ζ is the surface height displacement (for ocean tides), and u and v are the velocities multiplied by the sine of colatitude (and by i for the north–south current v) and nondimensionalized using $1/(2\Omega)$ as the time scale and a as the length scale. In different applications, any of the three parameters $(\lambda, s, \varepsilon)$ may be the eigenvalue while the other two are fixed.

In atmospheric tides ε is the eigenparameter. As reviewed in Chapman and Lindzen's monograph [8], the analogy with classical Sturm–Liouville theory led to some 30 years of confusion. It was assumed that ε, which is the square of the vertical wavenumber for atmospheric tides, always had to be positive. It was finally realized in the 1960s, independently by Kato and Lindzen, that because the tidal equations have so-called "apparent singularities" at the "critical" or "inertial" latitudes, classical Sturm–Liouville theorems need not apply. There are also an infinite number of modes with *negative* ε for the diurnal tide.

We shall return to the issue of "missing modes" in a later section. In the next section, we describe another difficulty that arises for tides in an ocean bounded by meridians.

To understand some of the peculiarities of the tides, it is helpful to also use the "equatorial beta-plane," which is a consistent asymptotic approximation to the spherical tidal equations in the limit $\varepsilon \Rightarrow \infty$ for fixed λ, s:

$$\begin{aligned} \lambda u - yv - s\zeta &= 0 \\ yu - \lambda v + \zeta_y &= 0 \\ su - v_y - \varepsilon\lambda\zeta &= 0. \end{aligned} \quad (2.8)$$

Though only an approximation, this has the great advantage that it can be solved exactly by Hermite functions.

3. PSEUDOSPECTRAL METHODS FOR PROBLEMS WITH WEAK ENDPOINT SINGULARITIES

Both Legendre's equation and the tidal equations are singular at the endpoints where the meridians (curves of constant longitude) converge to a single point. When the domain is the entire surface of a sphere, however, these coordinate singularities are only *apparent*. (The reason is that we are free to rotate the coordinate system and, therefore, the location of the coordinate singularities, to arbi-

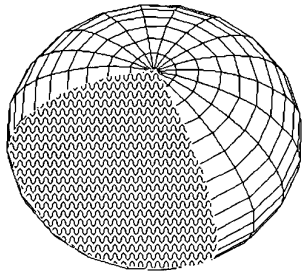

FIG. 1. Schematic of an ocean basin (wavy shading) bounded by meridians.

trary points on the sphere, but this cannot change the positions of any *real* singularities of the eigenmodes.) The implicit boundary conditions at the poles are that the solution somehow stay bounded, even though the coefficients of the differential equation are singular. It is unnecessary to *explicitly* impose boundary conditions at the endpoints $\mu = \pm 1$ because the pseudospectral method automatically picks out those modes—the Associated Legendre functions for (2.1)—which are ''nice'' at the poles. Boyd [5] shows that Fourier series, Chebyshev polynomials, and spherical harmonics (which are the exact eigenmodes!) all work fine for (2.1).

In the quantum mechanics application of Legendre's equation (mapped to the infinite interval, (2.2)) and also for tides in an ocean bounded by meridians, the situation is different. The zonal wavenumber s is no longer quantized, but it is the eigenvalue and may assume arbitrary values. For large $|y| \gg 1$, Legendre's equation (2.2) reduces to

$$u_{yy} + Eu = 0, \quad |y| \to \infty \tag{3.1}$$

which implies that the solutions are asymptotically

$$u \sim \exp(-\sqrt{-E}y) \quad \text{as } y \Rightarrow \infty. \tag{3.2}$$

This shows that exponential decay is possible only when $E < 0$. When we translate back to spherical coordinates, (3.2) is equivalent to

$$u \sim (1 - \mu)^{\sqrt{-E}/2} \quad \text{as } \mu \Rightarrow 1. \tag{3.3}$$

In other words, the solution has an algebraic branch point at both poles unless $(\frac{1}{2})\ (-E)^{1/2}$ is equal to an integer. The analysis for the tidal equation yields the same answer; all solutions of spherical problems will have branch points of the form of (3.3). (The associated Legendre functions of odd wavenumber include a factor of $(1-\ \mu^2)^{1/2}$ which is just what is needed so that the rest of the eigenfunction is analytic (in fact, a polynomial!).)

Figure 1 is a schematic that illustrates why the change from a global ocean or atmosphere (with λ or ε as an eigenparameter) to an ocean basin bounded by great circles of constant longitude (with zonal wavenumber s as an eigenparameter) alters both the physics and the mathematics. On an ocean-covered planet, the whole area around the coordinate singularity is water and there is nothing special about the pole. In contrast, at the poles, the coastlines of the meridian-bounded ocean basin meet in a sharp-angled corner. The pole is physically distinct because it is the only point in the hemisphere where the coastline makes an *abrupt*, *discontinuous* turn. The solutions of partial differential equations, even the Laplace equation, are usually singular in the corners. The most familiar physical manifestations of these corner singularities is in structural mechanics where corners are regions of high stress. Many homes have networks of cracks radiating from the corners of door openings. Commercial jets have rounded windows because the first passenger jet, the *Comet*, had to be withdrawn from service after repeated crashes initiated by cracks that began at the corners of its square windows and then opened up to break the aircraft's back.

It follows that in a bounded basin, we shall *lose spectral accuracy* because of corner singularities. Rounding the corners, as in aircraft windows, is not feasible because the coastline would no longer follow lines of constant longitude: a one-dimensional eigenvalue problem would become two-dimensional. Fortunately, there is a good alternative: Apply the Mercator transformation and then use any spectral basis which gives exponential convergence for functions analytic on the real y-axis [11–13].

The rational Chebyshev functions $TBj(y)$ [14] are a good choice. These functions are so named because they are rational functions in y. However, the easiest way to implement them is to recognize that they are the images under a change of coordinate of the terms of an ordinary Fourier cosine series.

We expand

$$u(\mu) = \sum_{j=0}^{N-1} a_j \cos(jt), \tag{3.4}$$

where the trigonometric coordinate t is related to the cosine of colatitude μ via

$$\mu = \tanh(L \cot(t)) \Leftrightarrow t = \operatorname{arc\,cot}\left\{\frac{\operatorname{arc\,tanh}(\mu)}{L}\right\}, \tag{3.5}$$

where L is a user-choosable constant, the ''map parameter.'' (For simplicity, we take $L = 1$ in most calculations here, but this can be varied to improve the rate of convergence [14–16].) The collocation points are evenly spaced in t:

$$t_i = \pi \frac{(2i-1)}{2N},$$
$$i = 1, 2, ..., N \Leftrightarrow \mu_i = \tanh\left(L \cot\left\{\pi \frac{(2i-1)}{2N}\right\}\right). \tag{3.6}$$

The magic of the transformation is that the grid points are clustered with *exponential* density near the endpoints in the sense that the separation between nearest neighbors close to $\mu = \pm 1$ decreases exponentially fast with N:

$$\mu_i \approx 1 - 2\exp\left\{-\frac{4L}{\pi(2i-1)}N\right\}, \quad \frac{i}{N} \ll 1. \tag{3.7}$$

In contrast, the nearest neighbor separation for a Chebyshev expansion in μ is $O(1/N^2)$. Squeezing the grid points close to the singularities recovers spectral accuracy [11–13].

In practice, the Mercator transformation is not always necessary. One can show that the coefficients of a Chebyshev series in μ will asymptotically decay as

$$a_j \sim \text{const}/n^{s+1} \quad \text{if branch pt. of type } (1-\mu)^{s/2},$$
$$s \neq \text{even integer}. \tag{3.8}$$

(Spherical harmonics series have similar properties [1].) If s is *large*, the coefficients of a Chebyshev or Legendre series will decay as a high inverse power of j, and this gives high accuracy for moderate N. For tidal modes, there is also a tendency towards equatorial confinement which increases with ε, especially for the lowest modes.

Thus, O'Connor [6] obtained good results even without the mapping. His choice of $\varepsilon = 22$ confined the lowest few modes, which would otherwise have very slowly convergent series because s is small, so that he was able to obtain six places of accuracy for all eigenmodes with $N < 100$. For smaller ε, however, the low modes (small s) would have very slowly convergent spectral series without the Mercator transformation.

Figure 2 compares the spectral coefficients for the second symmetric mode for the fortnightly tide as computed using the Mercator coordinate with rational Chebyshev basis functions and also the coefficients computed using Chebyshev polynomials in the coordinate μ. (The Chebyshev series is not quite the same basis as employed by O'Connor, but should be equally sensitive to the polar branch points of the eigenmode.)

We see that if only moderate accuracy of three or four decimal places is needed, Chebyshev polynomials are quite effective. However, the Chebyshev coefficients plateau and the coefficients beyond a_{10} decrease very slowly with increasing degree. (The plateau is at $a_{10} \sim 10^{-3}$, instead of at $O(1)$, because this mode is equatorially trapped, so the singular factor $(1-\mu^2)^{s/2}$ is weakened by the small amplitude of the mode near the singular points, the north pole and south poles.) The Mercator coordinate method actually converges slower at first, but like the turtle in the Aesop proverb of the rabbit and the turtle, the Mercator/*TBj* scheme eventually overtakes the Chebyshev polynomial series. Boyd [12] gives other examples of a "crossover point," here at about a_{22}, where the asymptotically faster method finally overtakes the initially faster-but-asymptotically slower Chebyshev polynomial series.

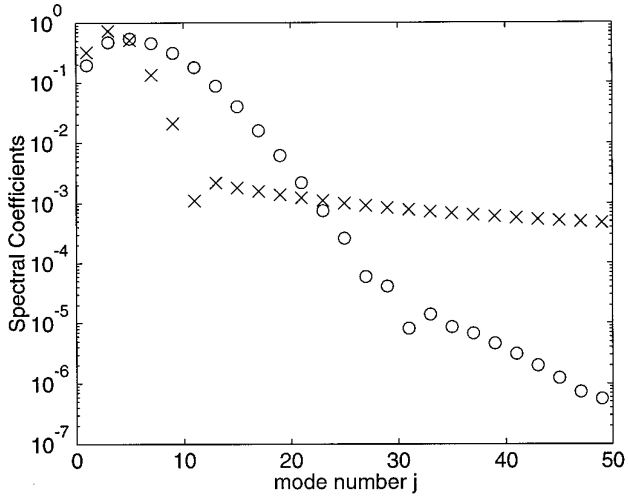

FIG. 2. Absolute values of the spectral coefficients for the zonal velocity for the fortnightly tide ($\lambda = \frac{1}{27}$) with Lamb's parameter $\varepsilon = 22$ for the second symmetric mode. The zonal wavenumber (eigenvalue) is $s = -0.500$. Circles: TB_j coefficients after the globe has been transformed to the Mercator coordinate with map parameter $L = 2$. X's: coefficients of Chebyshev polynomials. (Because the east–west current is symmetric with respect to the equator, all the odd degree coefficients are zero, so only the even coefficients $\{a_0, a_2, a_4, ...\}$ are shown.)

The spectral method is easy to implement. The chain rule shows that

$$\frac{d}{d\mu} = -\frac{\sin^2(t)}{L\,\text{sech}^2\{L\cot(t)\}}\frac{d}{dt} \Leftrightarrow D = -\frac{\sin^2(t)}{L}\frac{d}{dt} \tag{3.9}$$

which can be iterated to obtain higher derivatives. The operator D takes an especially simple form because $(1-\mu^2) = \text{sech}^2(L\cot(t))$, which is just what is needed to cancel the same factor in the denominator of the "metric factor" on the left of (3.9). Thus, the Legendre equation becomes, in terms of the trigonometric coordinate t,

$$\frac{\sin^4(t)}{L^2}\frac{d^2u}{dt^2} + 2\frac{\sin^3(t)\cos(t)}{L^2}\frac{du}{dt} + \{V_0\,\text{sech}^2[L\cot(t)] + E\}u = 0. \tag{3.10}$$

The $N \times N$ matrix eigenvalue problem equivalent to (3.10) is (with E as eigenvalue)

TABLE I

Matlab Code to Solve Legendre's Equation with Mercator Coordinate Transformation

```
N=60   % Number of collocation points,          M=1;   % Map parameter
for j=1:N % Collocation points t;
t(j)=pi*(2*j-1)/(2*N);   sechsq(j)=sech(L*   cos(t(j))/sin(t(j))   )^2;
end
phi2=zeros(n,n);        % grid point values of basis function
phiDD2=zeros(n,n);    % grid point values of operator D (squared!)
for i=1:N    % Outer loop over collocation point (row) index
ss=sin(t(i));    cc=cos(t(i));
          for j=1:N   % Inner loop over basis function (column) index
        jj=j-1;
         phi2(i,j)=cos(jj*t(i));
           pt=-jj*sin(jj*t(i));   ptt=-jj*jj*phi2(k,j);
             phiDD2(k,j)= ss*ss*ss*(ss*ptt +2*cc*pt)/(L*L);
          end
end
for i=1:N,         for j=1:N
AA(i,j)= phiDD2(i,j)+nu*sechsq(i)*phi2(i,j) ;
BB(i,j)=-phi2(i,j);
end                end
[VV,EE]=eig(A,B);
% The eigenvalues E are the diagonal elements of EE; the spectral
% coefficients of the eigenvectors are the columns of  VV.
```

$$\mathbf{Aa} = E\mathbf{B}, \tag{3.11}$$

where the elements of the column vector **a** are the coefficients of the Fourier series (3.4) and the elements of the square matrices are

$$\begin{aligned} A_{ij} &= \frac{\sin^4(t_i)}{L^2}\{-(j-1)^2\cos((j-1)t_i)\} \\ &\quad + 2\frac{\sin^3(t_i)\cos(t_i)}{L^2}\{-(j-1)\sin((j-1)t_i)\} \\ &\quad + V_0\,\mathrm{sech}^2[L\cot(t_i)]\cos((j-1)t_i), \\ B_{ij} &= -\cos((j-1)t_i). \end{aligned} \tag{3.12}$$

Table I is a complete Matlab code for solving the Legendre eigenvalue problem. The pseudospectral scheme for the tidal equations is identical except that because the latter is a system of three equations in three unknowns, the matrices (of total dimension $3N$) must be evaluated as nine $N \times N$ blocks. The code omits the lines needed to graph the eigenfunctions, compare results for different N, and so on, and it may be hard to follow in detail if one is unfamiliar with Matlab. However, the major theme of the table is its *brevity*: just 20 lines. Even with the change of coordinate, the pseudospectral method is not tricky or complicated!

4. DISTINGUISHING "GOOD" EIGENVALUES FROM "BAD" EIGENVALUES

One special curse of eigenvalue calculations is that an N-dimensional matrix usually has N eigenvalues. When the matrix is the discretization of a differential equation (by any method), all the matrix eigenvalues will *not* be good approximations to those of the differential equation. The reason is that the eigenfunctions of the differential equation oscillate more and more rapidly as the mode number j increases. For a classical Sturm–Liouville problem, the error is smallest for the smallest eigenvalue and then increases rapidly until the matrix eigenvalues become useless as approximations to those of the differential equation as illustrated in Fig. 3. For ''nice'' eigenvalue problems, one typically obtains about $N/2$ matrix eigenvalues that are within a few percentages of those of the differential equation, which motivated the ''eigenvalue Rule-of-Thumb'' on page 100 of [1].

However, the $N/2$ estimate may be wildly in error for problems such as the quantum mechanics application of Legendre's equation, where there are only a handful of localized bound states—three for the case illustrated below—regardless of the numerical resolution. Since the elements of the continuous spectrum are oscillatory for large $|y|$ and thus have an infinite number of crests and troughs on $y \in [-\infty, \infty]$, it is obvious that a truncated sum of rational Chebyshev functions, or any other standard basis set, cannot furnish an accurate representation. We have

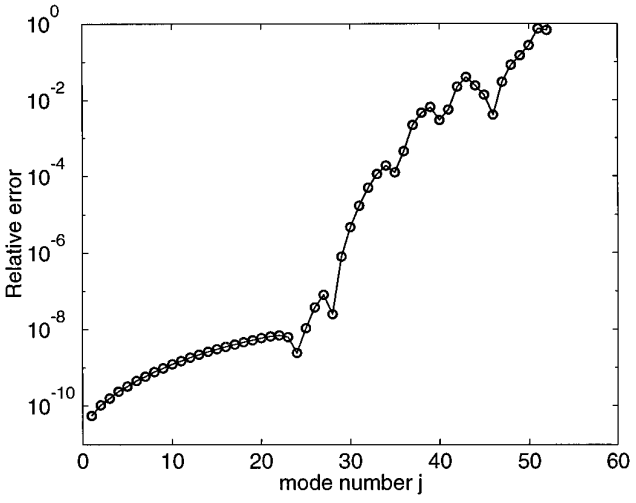

FIG. 3. Relative errors, as a function of matrix mode number j, for Legendre's differential equation for zonal wavenumber one, as computed using the tanh/cot map with $L = 1$, $N = 60$ basis functions. The exact eigenvalues are $\lambda_j = -j\,(j+1)$, $j = 1, 2, \ldots$. [$\lambda_j = V_0$].

to use other strategies to compute the continuous spectrum [17].

How do we tell the "good" eigenvalues from the "bad," i.e., inaccurate eigenvalues? The obvious answer is to compare the eigenvalues for different N; only those whose difference or "resolution-dependent drift" is small can be believed.

But which pairs of eigenvalues should be compared? For many classical problems, the eigenvalues for different N are in direct correspondence after sorting the modes by eigenvalue, smallest first. (Sorting is usually necessary because library eigenvalue routines typically return the eigenvalues and eigenvectors in a very unpredictable order.) One can then use what we shall dub the "ordinal drift,"

$$\delta_{j,\text{ordinal}} \equiv |\lambda_j^{(N_1)} - \lambda_j^{(N_2)}|, \tag{4.1}$$

where $\lambda_j^{(N)}$ is the jth eigenvalue (after the eigenvalues have been sorted) as computed using N spectral coefficients.

For some problems, however, there are spurious modes of small eigenvalues which are inserted *between* the true eigenvalues. The "ordinal drift" may falsely imply that many numerical eigenvalues are inaccurate when they in fact match closely the eigenvalues of larger N and different mode numbers j. In such cases, it is therefore much more useful to assess modes by the "nearest neighbor" drift:

$$\delta_{j,\text{nearest}} \equiv \min_{k\in[1,N_2]} |\lambda_j^{(N_1)} - \lambda_k^{(N_2)}|. \tag{4.2}$$

The nearest neighbor drift is the difference between the jth eigenvalue at low resolution and whatever high resolution eigenvalue is closest to it.

The second issue is: what should we compare the drift to? The obvious choice is the magnitude of the jth eigenvalue itself. However, for the Legendre equation, $\lambda_j = -j\,(j+1)$. When j is large, an error which is $O(j)$ is unacceptable because the "intermodal separation,"

$$\sigma_j \equiv \begin{cases} |\lambda_1 - \lambda_2|, & j = 1, \\ \frac{1}{2}\{|\lambda_j - \lambda_{j-1}| + |\lambda_{j+1} - \lambda_j|\}, & j > 1, \end{cases} \tag{4.3}$$

is $O(2j)$ so that the numerical eigenvalue is no better than a random guess in the correct range. However, comparing the resolution-dependent drift to the eigenvalue, which is $O(j^2)$ and, therefore, large, compared to the intermodal separation for large j, may give the appearance that the eigenvalue is "good."

We therefore recommend assessing the goodness of eigenvalues by making a log/linear plot of one of the two ratios,

$$r_{j,\text{ordinal}} \equiv \frac{\min(|\lambda_j|, \sigma_j)}{\delta_{j,\text{ordinal}}} \tag{4.4}$$

$$r_{j,\text{nearest}} \equiv \frac{\min(|\lambda_j|, \sigma_j)}{\delta_{j,\text{nearest}}} \tag{4.5}$$

which are both *large* when the eigenvalue approximation is *accurate*. Figure 4 shows these ratios for Legendre's

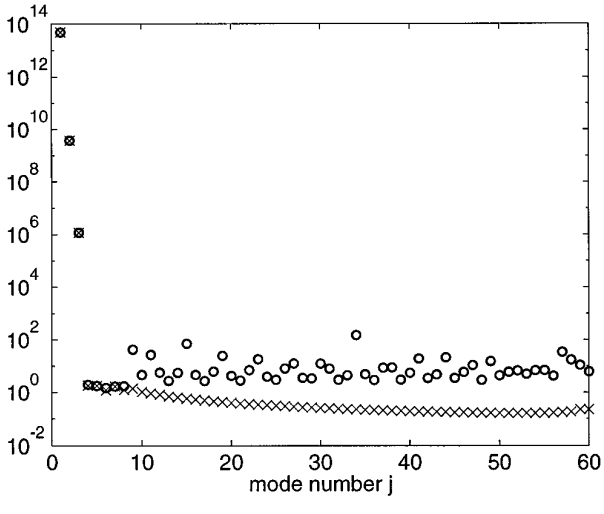

FIG. 4. The drift ratios $r_{j,\text{nearest}}$ (circles) and $r_{j,\text{ordinal}}$ (x's) are plotted on a logarithmic scale versus mode number j. The modes are ordered by absolute value of eigenvalue with $j = 1$ the smallest. The ratios compare $N_1 = 60$ and $N_2 = 90$, where N is the number of collocation points (terms in the spectral series). The map parameter $L = 1$. The differential equation, which is Legendre's with $V_0 = 8.75$, has only three discrete eigenvalues ("bound states"). The drift ratios are defined so that the "good" eigenvalues appear at the top of the graph.

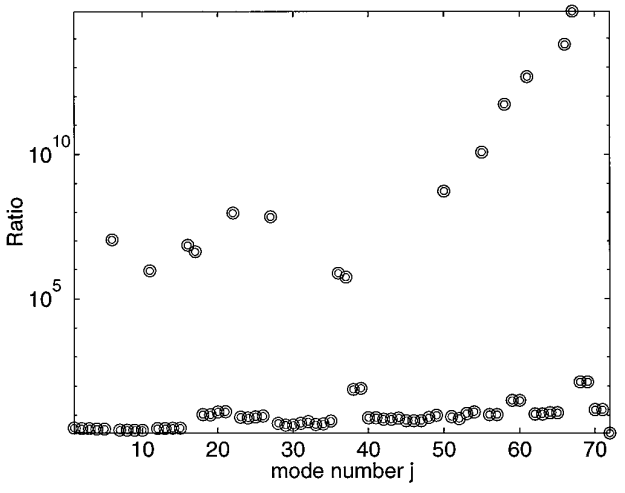

FIG. 5. The ratio $r_{j,\text{nearest}}$ versus mode number j, where the eigenvalues are ordered by increasing absolute value with $j = 1$ as the smallest and where the two resolutions are $N = 36$ and $N = 50$. Tides in a bounded basin for frequency ratio $\lambda = \frac{1}{27}$ with Lamb's parameter $\varepsilon = 22$. N of the matrix eigenvalues are infinite and have been purged before plotting. The size of the matrix problem is $3N \times 3N$.

differential equations for the quantum case, where the potential is strong enough to create precisely three bound states. The ''good'' eigenvalues, which are highly accurate approximations to the three bound states, appear at the top of the graph. The ''bad'' eigenvalues, such that these ratios are $O(1)$, appear at the bottom.

It is annoying to have to define two ''drift ratios'' instead of one. However, Fig. 4 shows that for this problem, the ''ordinal'' ratio is much closer to one and much less scattered than $r_{j,\text{nearest}}$. The reason is that it may happen by dumb luck that one of the high resolution eigenvalues is close to a low resolution eigenvalue. The result is that $r_{j,\text{nearest}}$ may be as large as $O(100)$, even though the approximate eigenvalue is a random number.

For some eigenvalue problems, one must simply live with the greater scatter of $r_{j,\text{nearest}}$. Figure 5 illustrates this ratio for the numerical eigenmodes of Laplace's tidal equation for the fortnightly tide ($\lambda = \frac{1}{27}$). Remarkably, the ''good'' eigenvalues form two clusters separated by a number of garbage modes.

In the equatorial beta-plane approximation, we can analytically compute the eigenvalues where n is the mode number in the usual geophysical convention:

$$s = \lambda\sqrt{\varepsilon}, \quad n = -1 \quad \text{[Kelvin wave]} \tag{4.6a}$$

$$s = -\frac{1}{2}\left\{\frac{1}{\lambda} \pm \sqrt{(1/\lambda^2) + 4\lambda^2\varepsilon - 4(2n+1)\sqrt{\varepsilon}}\right\}, \quad n = 0, 1, 2, \ldots. \tag{4.6b}$$

Although O'Connor's application needs only the modes whose height and zonal currents are symmetric with respect to the equation (n odd), the formula applies to the antisymmetric modes, too.

Table II compares the beta-plane formulas with numerical calculations on the globe. Although the approximation is accurate only for the three or four eigenvalues of smallest magnitude, it correctly reproduces the bimodal distribution of s. In particular, the analytical formula resolves the mystery of why the eigenvalues $s = -0.500$ and -25.99 (modes 11 and 67 in Fig. 4) are both computed very accurately: both correspond to the $n = 0$ mode.

A more accurate approximation to modes with large s may be obtained by noting that when $s^2 \gg \varepsilon$, it is a good first approximation to set $\varepsilon = 0$. In this case, Eq. (7.18) of Longuet-Higgins [7] shows that the tidal equations, after reduction, become the equation satisfied by the associated Legendre functions. Applying (2.6) with $E = -s^2$ and $V_0 = -s/\lambda$, we obtain

$$s = n + 1/2 - \frac{1}{2\lambda} \pm \frac{1}{2}\sqrt{\frac{1}{\lambda^2} - 4\left\{\frac{n+1/2}{\lambda} - 1/4\right\}}, \quad n = 0, 1, 2, \ldots, n_{\max}, \tag{4.7}$$

where $n_{\max}$ is the largest integer for which s is real and negative. Table III shows that the ''zero-epsilon'' approximation is quite accurate for the six modes of largest $|s|$ (right half of the table) and quite awful for the modes corresponding to the plus sign in (4.7). Although this approximation fails for small s and also does not predict any discrete modes with complex eigenvalues, it is more realistic than the beta-plane, in that it predicts only a finite number of discrete modes. The beta-plane model, like its quantum equivalent (the harmonic oscillator, also solved by Hermite functions), predicts an infinite number of discrete modes. In reality, there seems to be only a finite number of discrete modes while the rest of the spectrum is a continuous spectrum.

A plot of the reciprocals of the error ratios is a very useful tool for distinguishing accurate from inaccurate eigenvalues. However, analytical approximations, such as the beta-plane and ''zero-epsilon'' approximations for ocean tides, may be valuable also in understanding the numerical results. We must also emphasize that it is important to look at the spectral coefficients of the modes, and not just at the eigenvalues.

The equatorial beta-plane version of the tidal equations (2.8) provides a cautionary tale. The numerical solution gave a resolution-independent eigenvalue equal to the *negative* of the Kelvin wave whose exact eigenvalue is given by (4.6a). Inspection of the eigenvalues alone is *not sufficient* to identify the spuriousness of this mode because its

TABLE II

Comparison of Numerically Computed Eigenvalues s for Tidal Modes on the Sphere with the Predictions of Equatorial Beta-Plane for the Fortnightly Tide

Mode no.	β-plane	Spherical	Difference	β-plane	Spherical	Difference
−1	0.173	0.164	−0.0098	—	—	—
0	−0.530	−0.500	0.0301	−26.83	−25.99	0.84
1	−0.897	−0.820	0.0777	−26.47	−23.87	2.60
2	−1.275	−1.236	0.0390	−26.10	−21.62	4.48
3	−1.665	−1.900	−0.2350	−25.72	−19.22	6.51
4	−2.068	−3.085	−1.017	−25.34	−16.54	8.80

numerical eigenvalues vary as slowly with N as those of the true Kelvin wave.

Nevertheless, this mode is nonsense. It is known that if the beta-plane, which is unbounded in latitude, is modified by a coastal boundary along a circle of latitude, then this bounded ocean does support an ''anti-Kelvin'' wave which is boundary-trapped to the coast and propagates in the direction opposite that of the equatorially trapped Kelvin wave. Since the grid points span only a finite interval, this anti-Kelvin wave emerges from the infinite interval numerics.

Figure 6 compares the spectral coefficients and zonal velocity for the true and false Kelvin waves. The coefficients of the physical mode converge exponentially as shown by a near-linear decrease on a logarithmic plot. In contrast, the coefficients of the anti-Kelvin wave *diverge* exponentially fast! As a result, the zonal velocity oscillates wildly with latitude, dominated by the highest few basis functions that are retained in the truncation.

5. GALERKIN-BY-RECURRENCE: THE DANGER OF BEING TOO CLEVER

The usual spectral basis functions satisfy a vast number of recurrence relations. When the discretization is Galerkin's method—multiplying the residual by each of the N lowest basis functions in turn and integrating to obtain N integral constraints to determine the coefficients—the required integrations can often be bypassed by manipulations of the recurrence relations instead. In the presilicon era, Galerkin's method was very appealing because for simple problems, the Galerkin matrix is *banded* with mostly zero matrix elements. In tidal theory [8–10], for example, Hough derived a tridiagonal matrix by using the spherical harmonics recurrences. He then converted the matrix into a continued fraction to create a rapidly convergent iteration for the eigenvalues: a numerical method that was not too onerous even for paper-and-pencil calculations. Fox and Parker's [18] monograph is mostly a catalogue of similar recurrence computations using Chebyshev polynomials.

However, we live in an age when even a personal computer has a megaflop execution speed, and our priorities are different. Banded matrices are still useful when the eigenproblem is to be solved *symbolically* [19], but the QZ algorithm causes ''fill-in'' so that it is just as expensive for a sparse matrix as a dense matrix. A banded Galerkin discretization matrix saves us nothing over a dense pseudospectral method if the linear algebra is to be handed off to the QZ algorithm. Unfortunately, the use of recurrence relations can cause a lot of trouble.

The reason is that to test a program, one usually simpli-

TABLE III

Same as Table II except that the $\varepsilon = 0$ Approximation (4.7) Is Compared with the Numerical Calculations of s

Mode no.	$\varepsilon = 0$: Eq. 4.7	Spherical	Difference	$\varepsilon = 0$: Eq. 4.7	Spherical	Difference
−1	—	0.164	—	—	—	—
0	0	−0.500	−0.50	−26	−25.99	0.01
1	−0.084	−0.820	−0.74	−23.92	−23.87	0.05
2	−0.276	−1.236	−0.96	−21.72	−21.62	0.10
3	−0.619	−1.900	−1.28	−19.38	−19.22	0.16
4	−1.199	−3.085	−1.81	−16.81	−16.54	0.27
5	−2.169	$-7.21 \pm i\,1.00$	Complex	−13.83	−13.34	0.49
6	−4.35	$-6.21 \pm i\,5.32$	Complex	−9.65	Complex	—

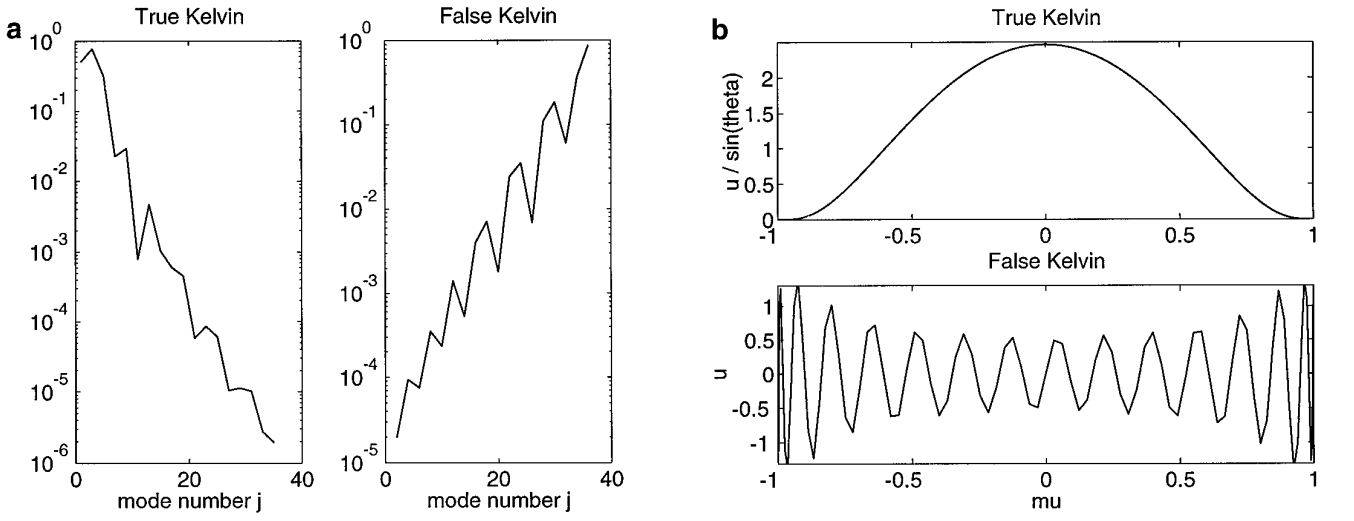

FIG. 6. a. Absolute values of the spectral coefficients for the Kelvin and anti-Kelvin modes. b. Graphs of $u/\sin(\theta)$ for the Kelvin wave (where θ is colatitude) and u itself for the spurious anti-Kelvin mode.

fies the differential equation to one which has a known, exact solution. When the tidal equations are thus simplified, the exact solution is a single spherical harmonic (for each unknown) and the Galerkin matrix is *diagonal*. Thus, the test can tell us nothing about whether we have correctly applied and coded the recurrence relations for the *off-diagonal* elements.

In contrast, the elements of the pseudospectral matrix are nontrivial even when the differential equation is simplified to constant coefficients. A pseudospectral program can pass a simplified test and yet fail for the real problem, only when the coefficients of the differential equation are coded incorrectly.

O'Connor ran afoul of this problem [6]. A small slip in the messy algebra of the recurrence relations required a corrigendum because the tabulated eigenvalues were corrupted.

The recurrence relations are seductive both because of their historical importance and their intrinsic aesthetic beauty. Sometimes, though, it is better to resist the temptation to be clever. KISS—keep it simple, stupid—is a good philosophy for a number-cruncher.

6. MISSING A MODE: LOCAL VERSUS GLOBAL METHODS FOR SOLVING THE MATRIX EIGENVALUE PROBLEM

The QZ algorithm, employed here to solve the N-dimensional generalized matrix eigenvalue problem, has been refined into a very reliable piece of software that will take an arbitrary pair of square matrices and return all N eigenvalues, even if some are infinite, or the problem is otherwise deficient. (This algorithm is sometimes also called the QR algorithm, after the simpler scheme for standard eigenvalue problems from which the QZ method was derived.) It is available in most software libraries and is a built-in command in Matlab. Its main drawback is cost. Because it uses iteration, rather than a finite set of steps, a precise estimate of cost is impossible, but experience has shown that the QZ cost is $O(10\ N^3)$ floating point operations.

Alternatively, one can use a local iteration, that is, one that computes a single mode at a time. Orszag [2] shows that one can compute individual modes at a cost proportional to N^2, which is a vast improvement over the N^3 expense of the QZ method. Indeed, when one is tracing out a parameter space and repeatedly computing the same mode for slightly different parameters, such "local" iterations may be very useful.

The danger is that it is easy to miss modes with a one-mode-at-a-time method. Boyd used an iterative finite difference scheme [20], together with an artificial viscosity, to compute the eigenvalues of

$$u_{xx} + \left\{\frac{1}{x} - \lambda\right\} u = 0, \quad u(-6) = u(6) = 0. \qquad (6.1)$$

A pseudospectral QZ calculation, using a mapping to shift the integral of integration to detour around the singularity in the lower half-plane, confirmed the earlier calculations—except that the third and seventh eigenmodes had been missed by local iteration [21].

Admittedly, this is an unusually nasty eigenproblem because the eigenmodes are logarithmically singular on the interior of the interval. Still, in stability calculations, the eigenvalue could lie almost anywhere in the complex plane,

making it easy to miss a mode. If the omitted instability has the largest growth rate—one shudders.

For multi-dimensional eigenproblems, sheer cost may force one to use a cheap local iteration. Whenever possible, however, the QZ algorithm is safer.

7. SUMMARY

The pseudospectral discretization, combined with the QZ algorithm, is a very powerful way to solve linear differential eigenproblems if one takes a few elementary precautions. The exponential rate of convergence with N implies that, at least for one-dimensional problems, it is usually very easy to distinguish "good" eigenvalues from "bad" eigenvalues. One must remember, however, that there may be few, many, or an infinite number of discrete eigenmodes, depending both on the differential equation and also on which parameter of the differential equation is the eigenvalue. One must be careful to both plot the "drift" in eigenvalues with resolution and also look at the spectral coefficients of the eigenmodes, as shown by the "anti-Kelvin" wave above. The pseudospectral algorithm is easier to program and check than hand-derivation of the elements of the corresponding recurrence relation. We recommend solving the matrix eigenvalue problem, which results from the pseudospectral discretization, by the QZ algorithm. This is available in most software libraries. Although expensive in comparison to the alternatives ($O(10\,N^3)$ operations), the QZ (also called "QR") is a reliable "black box" that will find all eigenvalues of the matrix problem, even if some of these are infinite or the matrix is singular. The global, find-all-the-modes character of the QZ algorithm is important because one-mode-at-a-time iterative methods may easily miss modes.

ACKNOWLEDGMENTS

This work was supported by the NSF through grant OCE9119459. I thank William P. O'Connor for suggesting this problem, for useful discussions and for his careful review of the manuscript.

REFERENCES

1. J. P. Boyd, *Chebyshev and Fourier Spectral Methods* (Springer-Verlag, Heidelberg, 1989).
2. S. A. Orszag, *J. Fluid Mech.* **50,** 689 (1971).
3. J. P. Boyd, *Mon. Weather Rev.* **106,** 1192 (1978).
4. J. P. Boyd, *J. Math. Phys.,* **19,** 1445 (1978).
5. J. P. Boyd, *Mon. Weather Rev.* **106,** 1184 (1978).
6. W. P. O'Connor, *Proc. R. Soc. London A* **449,** 51 (1995); Corrigendum, to appear.
7. M. S. Longuet-Higgins, *Phil. Trans. R. Soc. London A* **262,** 511 (1968).
8. S. Chapman and R. S. Lindzen, *Atmospheric Tides, Thermal and Gravitational* (Gordon & Breach, New York, 1970).
9. S. S. Hough, *Phil. Trans. R. Soc. London A* **189,** 201 (1896).
10. S. S. Hough, *Phil. Trans. R. Soc. London A* **191,** 139 (1898).
11. J. P. Boyd, *J. Comput. Phys.* **64,** 266 (1986).
12. J. P. Boyd, *Appl. Math. Comput.* **29,** 49 (1989).
13. J. Lund and K. L. Bowers, *Sinc Methods for Quadrature and Differential Equations* (SIAM, Philadelphia, 1992).
14. J. P. Boyd, *J. Comput. Phys.* **69,** 112 (1987).
15. J. P. Boyd, *J. Comput. Phys.* **45,** 43 (1982).
16. J. P. Boyd, *J. Comp. Phys.* **70,** 63 (1987).
17. J. P. Boyd, *Comput. Phys.* **4,** 83 (1991).
18. L. Fox and I. B. Parker, *Chebyshev Polynomials in Numerical Analysis* (Oxford Univ. Press, Oxford, 1968).
19. J. P. Boyd, *J. Symbolic Comput.* **16,** 377 (1993).
20. J. P. Boyd, *J. Math. Phys.* **22,** 1575 (1981).
21. J. P. Boyd, *J. Comput. Phys.* **57,** 454 (1985).
22. P. M. Morse and H. Feshbach, *Methods of Theoretical Physics*, Vol. 2 (Wiley, New York, 1953).