Journal of Computational Physics **179,** 216–237 (2002)
doi:10.1006/jcph.2002.7052

Deleted Residuals, the QR-Factored Newton Iteration, and Other Methods for Formally Overdetermined Determinate Discretizations of Nonlinear Eigenproblems for Solitary, Cnoidal, and Shock Waves

John P. Boyd

Department of Atmospheric, Oceanic and Space Science, University of Michigan, 2455 Hayward Avenue, Ann Arbor Michigan 48109-2143
E-mail: jpboyd@engin.umich.edu

Received June 20, 2000; revised February 22, 2002

Solitary waves, cnoidal waves, and shock waves can be computed by solving a nonlinear eigenvalue problem, which in discretized form is a system of nonlinear algebraic equations. Unfortunately, many such systems are *singular* because the solution is not unique until one or more additional constraints are imposed. For example, if the waves are translationally invariant and $u(X)$ is a solution, then so also is $u(X + \Phi)$ for arbitrary Φ. To obtain a unique solution, one must impose an additional condition to reduce the one-parameter family of solutions by constraining Φ. We describe five methods for coping with such singular systems: (i) reformulation of the problem, (ii) deleting residuals, (iii) Keller's bordered matrix scheme, (iv) QR-factored, overdetermined Newton iteration, and (v) pseudoinverse-Newton iteration. We illustrate these ideas using the cnoidal waves of the Korteweg–deVries equation, the traveling shocks of the Korteweg–deVries-Burgers equation, and the weakly nonlocal solitary waves of the nonlinear equatorial beta-plane equations. Finite difference, Fourier and rational Chebyshev pseudospectral methods, and spectrally upgraded finite differences are applied. Reformulation and deleting residuals are the cheapest strategies, but the QR-factored Newton iteration is needed for the shock waves, which lack the symmetry of the other two wave species.

Key Words: solitary wave; nonlinear algebraic equations; Newton's iteration; spectrally upgraded finite differences.

1. INTRODUCTION

If a nonlinear wave is assumed to travel at a phase speed c, then the time-dependent partial differential equation can be reduced to a nonlinear differential eigenproblem by the

0021-9991/02 $35.00

TABLE I
Comparison of Model Equations

Property	KdV	KdV–Burgers	Rossby
Wave type	Cnoidal	Shock	Weakly nonlocal soliton
Symmetry	Symmetric	Unsymmetric	ϕ is symmetric in x, y
No. of zero eigenvalues	2	1	∞
Cheapest method, fd	Deleted residual	QR	Deleted residual
Cheapest method, spectral	Reformulation	Deleted residual	—
Unsuccessful methods	Reformulation (fd only)	Reformulation, fd deleted Residual, and fd Keller's bordered	Reformulation

substitution $\partial/\partial t \to -c\partial/\partial X$, $\partial/\partial x \to \partial/\partial X$, where $X \equiv x - ct$ is a new spatial coordinate that moves with the wave. In this moving reference frame, the wave is steady. For example, the Korteweg–deVries equation of water wave theory becomes, with subscripts denoting differentiation with respect to the subscripted variable,

$$u_t + uu_x + u_{xxx} \to u_{XXX} + (u - c)u_X = 0 \quad \text{[KdV ODE]}. \tag{1}$$

The travelling wave solutions of the KdV ODE are solitary waves and cnoidal waves as found analytically by Korteweg and deVries more than a century ago. But the KdV equation is merely an example; a vast number of other time-dependent wave equations have similar traveling-wave solutions and may similarly be reduced to differential eigenproblems. In this article, we use a shock, a cnoidal wave, and a nonlocal solitary wave of three different model equations to illustrate this diversity (Table I).

The usual boundary conditions are either spatial periodicity or decay-to-a-constant. Unfortunately, these boundary conditions plus the differential equation are usually insufficient to specify a unique solution. For example, if the differential equation has coefficients which are independent of the spatial coordinates, then any solution can be translated in position by an arbitrary amount Φ, and $u(x - \Phi)$ still solves both the differential equation and the periodic or decay boundary conditions. However, there may be other or even an infinite number of additional degrees of freedom.

For example, the KdV equation has an additional dilational degree of freedom so that *two* additional equations must be appended to the usual differential equation/boundary conditions combination in order to specify a unique solution. An even more dramatic example is the Rossby wave in geophysical fluid dynamics. Small-amplitude perturbation theory shows that a Rossby wave is unique only after specification of an arbitrary zonally averaged mean current, whether in so-called quasigeostrophic models [24, 30] or the shallow-water wave equations for equatorially trapped waves [13] (discussed further below). Since the mean current $U(y)$ can be specified by an infinite series of orthogonal functions, it is equivalent to an *infinite* number of additional constraints.

In the discretized problem, these degrees of freedom are zero eigenvalues of the Jacobian matrix of the nonlinear system of algebraic equations.

2. NEWTON'S ITERATION

Finite-difference and pseudospectral methods convert a differential equation into a system of nonlinear algebraic equations, which is written as

$$\vec{r}(\vec{u}) = \vec{0}, \tag{2}$$

where the elements of $\vec{r}$ are the residuals of the differential equation plus the discretization of whatever additional constraints are needed to define a unique solution. The elements of $\vec{u}$ may be either Fourier or Chebyshev coefficients or the values of $u(x)$ at the points of the computational grid, as in finite-difference schemes. All of the algorithms for solving the discretized system are variations on Newton's iteration, also called the Newton–Ralphson method.

Let $\vec{u}^{(m)}$ denote the approximation to the solution of $\vec{r} = \vec{0}$ after the mth iteration. Expanding each component of $\vec{r}$ as a multivariable Taylor series about $\vec{u} = \vec{u}^{(m)}$ and discarding all but the linear and constant terms gives a linear matrix equation whose solution defines the next iterate, $\vec{u}^{(m)}$. Generalizing slightly to include the "underrelaxation factor" τ, the Newton iteration is

$$\vec{u}^{(m+1)} = \vec{u}^{(m)} - \tau\vec{\delta}, \tag{3}$$

where $\vec{\delta}$ is the solution of

$$\vec{\vec{J}}\,\vec{\delta} = \vec{r}\left(\vec{u}^{(m)}\right). \tag{4}$$

The elements of the Jacobian matrix $\vec{\vec{J}}$ are

$$J_{ij} \equiv \frac{\partial r_i}{\partial u_j}. \tag{5}$$

The standard Newton iteration sets $\tau = 1$. If the first guess $\vec{u}^{(0)}$ is too far from the root, Newton's iteration will diverge. Both theoretically and empirically, the domain of convergence can be greatly expanded by "underrelaxation," that is, by choosing $\tau < 1$.

We employed a simple line search to optimize the underrelaxation parameter τ. The residual $\vec{r}(\vec{u})$ was evaluated for

$$\tau_j \equiv 1/2^{(j-1)/2}, \quad j = 1, 2, \ldots j_{max} \tag{6}$$

for some j_{max}. The final Newton's correction at the mth iterate used whatever τ_j gave the smallest $\|\vec{r}\|$. Since the cost of factoring the Jacobian is an order of magnitude greater than the cost of a single evaluation of $\vec{r}(\vec{u})$, this line search strategy increases the cost per iteration only slightly. In the "end game" where $\vec{u}$ is very close to the minimum, the search will revert to the standard Newton's iteration.

If k additional constraints are required for uniqueness, then the Jacobian will be a *rectangular* $(n + k) \times n$ matrix. However, Newton's iteration is still well defined. The only complication is that both the linear algebra problem at each iteration and the nonlinear system itself cannot be solved in the classical sense of simultaneously making all components of the residual equal to zero. We can, however, minimize the norm of $\vec{r}$ by applying a QR factorization and thus compute the correction $\vec{\delta}$ and $\vec{u}$ itself in a least-squares sense. The total error still goes to zero rapidly as the number of grid points n is increased.

3. METHODS FOR SINGULAR SYSTEMS OF NONLINEAR EQUATIONS

In this article, we use five different strategies for coping with the difficulty that a straightforward discretization gives a singular Jacobian matrix for all three of the wave species discussed here (and many other types of waves as well). These are listed along with their costs for a dense matrix in Table II. With a finite-difference discretization, the Jacobian matrix is sparse, but the relative ordering of the methods by cost from cheapest (top) to most expensive (bottom) is not altered.

The "reformulation" and "deleted residuals" strategy both modify the discretized problem so as to obtain a nonsingular system with the same number of equations as unknowns; the standard Newton iteration then applies. The other three algorithms employ linear algebra methods (bordered matrix, QR factorization, and pseudoinverse, respectively) to solve the matrix problem at each Newton step when the Jacobian matrix is either singular or rectangular.

3.1. Reformulation

If $u(x)$ is symmetric with respect to its maximum, translational invariance can be suppressed by designing the numerical methods to impose symmetry with respect to $x = 0$. This forces the peak to be at the origin; the symmetry-assuming discretization gives a nonsingular system of algebraic equations.

In the Appendix, we show how to construct high-order finite differences that impose symmetry. For spectral methods, the analogous technique is to restrict the basis set to symmetric functions only—for a Fourier series, this means using a basis of cosine functions only. For both finite-difference and pseudospectral algorithms, the grid is restricted to $x \geq 0$ only.

For the unsymmetric shocks of the Korteweg–deVries–Burgers equation, alas, this sort of reformulation is useless.

3.2. Deleted Residuals

The idea is to simply replace one of the usual finite-difference or pseudospectral residual conditions by an explicit constraint, such as $u(0) = 1$. For the Korteweg–deVries cnoidal wave and equatorial Rossby solitary waves, this succeeds, but it fails for the unsymmetric shocks of the Korteweg–deVries–Burgers equation.

3.3. Keller's Bordered Matrix Algorithm

Keller has developed a bordered matrix algorithm for singular rectangular matrices [20]. For expository simplicity, we explain his ideas for the case of a single zero eigenvalue of the

TABLE II
Five Strategies for Singular Nonlinear Algebraic Systems

Method	Cost for dense matrix	Matrix type
Reformulation	$(2N^3/3)$	Square/nonsingular
Deleted residuals	$2(N^3/3)$	Square/nonsingular
Keller bordered matrix	$2(N^3/3)$ (+)	Square/singular
QR overdetermined	$4(N^3/3)$	Rectangular/nonsingular
SVD	$8(N^3/3)$	Square/singular

Jacobian matrix, but as he shows in [20], the generalization to an arbitrary number of zero eigenvalues is so straightforward that he expressed surprise that it had not been previously published.

The Keller/Newton method is identical to the standard iteration except for how $\vec{\vec{J}}\vec{\delta} = \vec{r}$ is solved. First, under the assumption of a single zero eigenvalue, or equivalently, that the Jacobian matrix has a nullspace of dimension one, partition the system as

$$\begin{vmatrix} \vec{\vec{A}} & \vec{B} \\ \vec{C}^T & D \end{vmatrix} \begin{vmatrix} \vec{x} \\ z \end{vmatrix} = \begin{vmatrix} \vec{g} \\ h \end{vmatrix}, \tag{7}$$

where $\vec{\vec{A}}$ is a square matrix and $\vec{B}$ and $\vec{C}$ are column vectors, all of dimension $(N-1)$. (The Nth row and column are the "border" of the $(N-1)\times(N-1)$ matrix $\vec{\vec{A}}$, so algorithms based on such a partition are "bordered matrix" methods.)

Next, compute the LU factorization of the upper left block $\vec{\vec{A}}$:

$$\vec{\vec{L}}\vec{\vec{U}} = \vec{\vec{A}}. \tag{8}$$

Then, do one triangular backsolve to compute a vector $\vec{U}_1$:

$$\vec{\vec{L}}\vec{U}_1 = \vec{B}. \tag{9}$$

Let $\vec{\phi}$ denote the Jacobian eigenvector of zero eigenvalue. This is found by performing a triangular backsolve as

$$\vec{\phi} \equiv \begin{vmatrix} \vec{\phi}_{upper} \\ -1 \end{vmatrix}, \quad \vec{\vec{U}}\vec{\phi}_{upper} = \vec{U}_1. \tag{10}$$

The correction is then, with ζ as an arbitrary constant,

$$\vec{\delta} = \begin{vmatrix} \vec{\delta}_{upper} \\ 1 \end{vmatrix} + \zeta\vec{\phi}, \quad \vec{\vec{L}}\vec{\vec{U}}\vec{\delta}_{upper} = \vec{g} - \vec{B}, \tag{11}$$

where $\vec{g}$ is the upper $(N-1)$ part of $\vec{r}$ and $\vec{B}$ is the upper $(N-1)$ elements of the Nth column of the Jacobian matrix. The constant ζ cannot be determined from the singular matrix problem alone. However, we can exploit this extra degree of freedom to impose at each iteration the uniqueness-needed constraint, such as $u(0) = 1$.

The cost of Newton's iteration with bordered-matrix computation is only slightly more expensive than the standard Newton iteration, the difference being $O(N^2)$ for a dense matrix versus an overall cost of $O([2/3]N^3)$. This is true even when there are m zero eigenvalues instead of one, and $\vec{B}$ and $\vec{C}$ must be generalized to $m\times(N-1)$ matrices, and similarly for the other components.

The flaw in this bordered matrix algorithm is that it implictly assumes that the upper left block $\vec{\vec{A}}$ is nonsingular even though the full matrix $\vec{\vec{J}}$ has a zero eigenvalue. *Generically*, this assumption is true, and Keller's bordered matrix strategy has been widely applied with success.

Unfortunately, discretizations of differential equations are not necessarily "generic" or "random." For the finite-difference discretization of the Korteweg–deVries–Burgers equation, alas, the bordered matrix method fails because the subblock $\vec{\vec{A}}$ is singular, too. We

reordered rows and columns so as to vary the elements that were deleted, but this did not help. The pseudoinverse, which does not make any *a priori* assumptions, succeeded in solving the square-but-singular Korteweg–deVries–Burgers matrix problem.

3.4. QR-Overdetermined Newton Iteration

The QR–Newton scheme applies when the Jacobian matrix is rectangular. Unfortunately, the QR factorization of the Jacobian matrix is twice as expensive as the LU factorization of a square matrix when the matrices are dense, and at least four times more costly for banded matrices. Still, the algorithm never failed in our tests.

The strategy of retaining all the finite-difference residuals and all the constraints and minimizing the residual of the formally overdetermined system is the safe, logical strategy. There is, in contrast, no general theory for how the computation will be corrupted by junking residuals at selected points, and indeed "deleted residuals" fails for the shock waves discussed later.

A system with more equations than unknowns gives a rectangular Jacobian matrix. It is still possible to solve the matrix problem in a least-squares sense by QR factorization. The orthogonal matrix $\vec{\vec{Q}}$ is computed only implicitly as the product of a series of rank-one Householder transformations. (Explicit computation of $\vec{\vec{Q}}$ more than doubles the cost but does reduce round-off error in the backsolve step; it proved unnecessary here.) The other factor $\vec{\vec{R}}$ is triangular so that the solution to

$$\vec{\vec{J}}\vec{\delta} = \vec{r} \tag{12}$$

may be obtained cheaply from

$$\vec{\vec{R}}^T\{\vec{\vec{R}}\vec{\delta}\} = \vec{\vec{J}}^T\vec{r}, \tag{13}$$

where the backsolves cost $O(N^2)$ operations each because $\vec{\vec{R}}$ is triangular. The round-off error of the "seminormal equation" (13) is often reduced, in applications to linear overdetermined systems, by one step of iterative refinement. This is unnecessary here because the Newton iteration effectively does the same job of "refinement."

An obvious alternative strategy is to multiply the original overdetermined problem by the transpose of the Jacobian to obtain the so-called "normal" equations

$$\vec{\vec{J}}^T\vec{\vec{J}}\vec{\delta} = \vec{\vec{J}}^T\vec{r} \quad \text{[normal equations]}, \tag{14}$$

where superscript "T" denotes the matrix transpose. The "squared Jacobian" matrix

$$\vec{\vec{\mathcal{J}}} \equiv \vec{\vec{J}}^T\vec{\vec{J}} \tag{15}$$

is a *square*, $N \times N$ symmetric matrix. In exact arithmetic, this alternative is equivalent to the QR factorization; $\vec{\vec{R}}$ is just the transpose of a Cholesky factor of $\mathcal{J}$; that is,

$$\vec{\vec{R}}^T\vec{\vec{R}} = \vec{\vec{J}}^T\vec{\vec{J}}. \tag{16}$$

It follows that, ignoring round-off error, the QR and normal equation strategies always give the *same answer*. However, the QR factorization employs only orthogonal matrices

and therefore is much more resistant to disastrous accumulation of round-off error than the normal equations. Both compute corrections $\vec{\delta}$ which minimize the L_2 norm of the residual of the matrix problem (12), $\|\vec{\vec{J}}\vec{\delta} = \vec{r}\|$.

The normal equations make the following definition precise:

DEFINITION 3.1 (Formally Overdetermined Determinate System). A system of algebraic nonlinear equations $\vec{r}(\vec{u})$ is said to be "formally overdetermined but determinate" if

1. The number of equations (length of $\vec{r}$) is greater than the number of unknowns, the elements of $\vec{u}$.
2. $\mathcal{J} \equiv \vec{\vec{J}}^T \vec{\vec{J}}$ is a *nonsingular* matrix, that is, all its singular values are greater than zero.

The QR–Newton iteration converges to a minimum $\rho_{min}(N)$ of the norm of the residual vector $\vec{r}(\vec{u})$. The right-hand side of the normal equations, $\vec{\vec{J}}\vec{r}$, is the gradient of $(1/2)\rho^2$ where $\rho \equiv \|\vec{r}\|$; this gradient is the zero vector at a minimum of ρ. From this and using the invertibility of $\vec{\vec{J}}^T \vec{\vec{J}}$ embodied in the definition of formally overdetermined determinate, one can show that, sufficiently close to the minimum that the linearization about the mimimum is legitimate, the iteration reduces to $\vec{\delta}^{(m+1)} \approx (1 - \tau)\vec{\delta}^{(m)}$. This shows that the iteration is an attractor and the correction $\vec{\delta}$ is converging to zero.

Also, costs of both the QR and normal equation methods are the same—and much more expensive than the linear algebra expenses of the deleted residuals approach. For a *dense* matrix, both the QR factorization and the symmetry-exploiting creation of the squared Jacobian matrix followed by a Cholesky factorization cost $O([4/3]N^3)$ floating-point operations for the factorization and $O(4N^2)$ operations per backsolve, exactly double the corresponding costs with an LU factorization. For banded matrices, the ratio of QR/LU factorization costs is much higher because, as is obvious from the normal equations, the QR method effectively *doubles the bandwidth*, and the factorization cost rises quadratically with bandwidth. The QR–Newton iteration was *six* times more expensive in execution time for the KdV problem and *fifteen* times more costly for the equatorial wave problem with only an infinitesimal improvement in accuracy over the "deleted residuals" strategy.

3.5. Pseudoinverse/SVD

All matrices can be given a singular value decomposition (SVD) in the form

$$\vec{\vec{J}} = \vec{\vec{U}}\vec{\vec{S}}\vec{\vec{V}}^T, \tag{17}$$

where $\vec{\vec{S}}$ is a diagonal matrix with σ_j, the "singular values," on the main diagonal and where the columns of $\vec{\vec{V}}$ are orthogonal to each other, as are the columns of $\vec{\vec{U}}$. The pseudoinverse, or "Moore–Penrose pseudoinverse," is defined as

$$\vec{\vec{J}}^+ = \vec{\vec{V}}\vec{\vec{S}}^+\vec{\vec{U}}^T, \tag{18}$$

where

$$\sigma^+ \equiv \begin{cases} 1/\sigma & \text{if } \sigma 0 \\ 0 & \text{if } \sigma = 0. \end{cases} \tag{19}$$

If all the singular values σ_j are nonzero, the pseudoinverse is the usual inverse. However, when $\vec{\vec{J}}$ is singular, then one or more singular values is zero. The pseudoinverse nevertheless

generates a particular solution:

$$\vec{\vec{J}}\vec{\delta} = \vec{f} \to \vec{\delta}_p = \vec{\vec{J}}^{+}\vec{f}. \tag{20}$$

The general solution is obtained by adding copies of each of the columns of $\vec{\vec{V}}$ associated with zero singular values to the particular solution. The arbitrary constants multiplying these $\vec{\vec{V}}$ columns can be chosen to enforce the constraints, such as the condition $u(0) = 1$ employed with the Korteweg–deVries–Burgers shocks below.

The pseudoinverse/SVD method is powerful because it always succeeds. We obtained rapid convergence from a square, singular Jacobian matrix for the Korteweg–deVries–Burgers shocks when Keller's bordered matrix algorithm failed.

The drawback of the pseudoinverse method is that it is very expensive. The cost of a full SVD with explicit computation of all factors is $O(13N^3)$ for a dense matrix of dimension N, or in other words about 20 times the cost of LU factorization. Fortunately, the algorithm can be modified to avoid explicit computation of the factors to reduce the cost to only $O([8/3]N^3)$, but this is still four times the cost of an LU factorization.

The pseudoinverse is useful only when the number of zero eigenvalues is *not known*. The other four strategies implicitly assume that the waves are sufficiently well understood that one knows precisely how many residuals to delete or how many extra rows to add to obtain a rectangular matrix of full rank (for the QR factorization), or how many rows and columns are the "border" in Keller's scheme. As costly as it is, the pseudoinverse can succeed when the physics is murky because it does not require any *a priori* assumptions.

4. KORTEWEG–DEVRIES CNOIDAL WAVES

4.1. KdV Free Parameters

KdV steadily propagating solutions which decay exponentially for large $|X|$ on $X \in [-\infty, \infty]$ are called "solitary waves" or "solitons"; the spatially periodic solutions are dubbed "cnoidal waves." Large-amplitude cnoidal waves have a single tall narrow peak on each spatial period which is well approximated by the sech^2 shape of the solitary wave; solitons are really just the limit of cnoidal waves for large amplitude with fixed spatial period.

THEOREM 4.1 (KdV ODE Symmetries and Eigenfunctions). *If $u(x; c, P)$ is a solution to the KdV ODE* (1) *for a given phase speed c and a given spatial period P (which could be infinite), then a solution is also given by*

$$\text{(i)}\quad v(X; c, P) \equiv u(X + \phi; c, P), \quad \forall\phi \quad [\textit{translational invariance}], \tag{21}$$

$$\text{(ii)}\quad w(X; c, P) \equiv -\lambda + (1 + \lambda)u\left([1 + \lambda]^{1/2}X; c, P\right), \quad \forall\lambda \quad [\textit{dilational invariance}]. \tag{22}$$

Equivalently, if $u(X; c, P)$ is a spatially periodic solution to the KdV ODE, the differential equation linearized about u,

$$\Delta_{XXX} + (u - c)\Delta_X + u_X\Delta = 0, \tag{23}$$

has two eigensolutions (of zero eigenvalue):

1. $\Delta_1(X) \equiv u_X$.
2. $\Delta_2(X) \equiv -1 + u(x) + (X/2)u_X$.

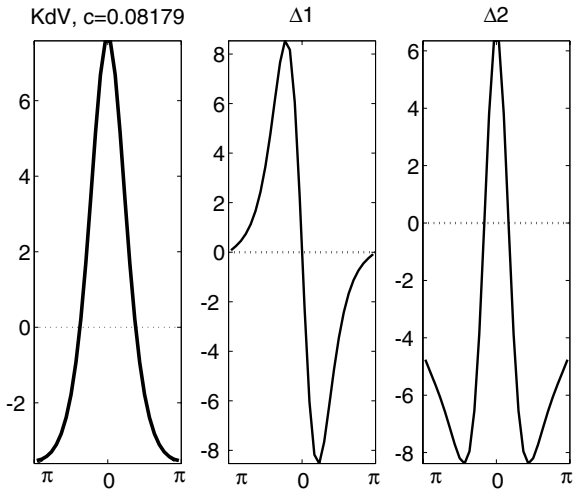

FIG. 1. (Left) The zero-mean KdV cnoidal wave for $c = 0.08$; the constant in the Fourier series is zero and the coefficient of $\cos(X)$ is 5. The soliton approximation, $u \approx -3.728 + 11.431\,\mathrm{sech}^2\,(0.976\,X)$, is plotted as a dashed curve in the left panel, but this is graphically indistinguishable from the periodic solution. This confirms the claim that the solitary wave is the large-amplitude limit of the spatially periodic cnoidal waves. (Middle) The translational eigenfunction, $\Delta_1 = du/dX$. (Right) The dilational eigenfunction, $\Delta_2 = -1 + u + (X/2)\,du/dX$.

The first half of the theorem is proved by substituting the new solutions into the ODE to verify that they are indeed solutions. The eigenfunctions follow by specializing to new solutions, which differ by arbitrarily small translations or dilations from the solutions which generate them.

A more comprehensive discussion of parity and other symmetries and their numerical consequences is given in Chap. 8 of [12] (see also [1, 18, 28]).

Figure 1 illustrates the KdV cnoidal wave for a typical case. The middle and right panels show the translational and dilational eigenfunctions of the linearized KdV equation (23); the discretization of this linearized ODE is the Jacobian matrix (without constraints and therefore singular!).

4.2. Implicit Constraints in Finite-Difference and Spectral Methods

The assumption that the KdV solution is symmetric with respect to $X = 0$ implicitly removes the translational invariance: If the point of symmetry is fixed at the origin, the freedom to shift the cnoidal wave left or right is destroyed.

Unfortunately, one additional constraint must be *explicitly* imposed. When the KdV equation is solved by a pseudospectral method using a Fourier series, as in [6, 10], the easiest constraint is

$$\int_0^P u\,dx = 0, \tag{24}$$

where P is the spatial period. The symmetry condition is imposed by using only a cosine basis; the integral constraint is imposed by omitting the constant from the Fourier series so that the basis functions are $\{\cos([2\pi/P]X), \cos(2[2\pi/P]X), \ldots\}$.

Thus, reformulation of the problem is completely successful in eliminating the singularity for the Fourier pseudospectral method. Futhermore, the size of the discrete basis is halved (from sines, cosines, and constant to just cosines). Since the pseudospectral Jacobian matrix is dense, this reduces costs by a factor of eight!

In finite-difference methods, reformulation is only partially successful. As explained in the Appendix, it is easy to built symmetry into differences, even high-order ones, and thus suppress the translational mode. However, it is necessary to *explicitly* impose an additional constraint.

The integral constraint can be imposed by adding one dense row of the Jacobian matrix with elements that are all one. This is equivalent to approximating the integral by the rectangle rule; for a *periodic* integrand, the rectangle rule has an error that decreases *exponentially* fast with the number of grid points.

The cost of the LU factorization of a banded matrix with one full row is only slightly larger than the cost of a strictly banded matrix of the same size and bandwidth. In the "deleted residuals" scheme, the integral constraint is cheap to impose.

Unfortunately, when the Jacobian matrix is multiplied by its transpose, as required by the QR–Newton algorithm, the resulting square matrix, $\mathcal{J}$, is *dense*. Thus, the integral constraint completely wrecks the sparsity of the QR-factored Newton method. We experimented with approximating $\mathcal{J}$ by a banded matrix, but the banded approximation to $\mathcal{J}$ was singular.

Fortunately, a unique solution can also be obtained by specifying the maximum amplitude of the wave, instead of the integral:

$$u(X = 0) = u_{max}. \tag{25}$$

Sixth-order polynomial interpolation, consistent with the sixth-order approximation of third derivatives, gives

$$u_{max} - \left\{ \frac{1225}{1024} u_1 - \frac{245}{1024} u_2 + \frac{49}{1024} u_3 - \frac{5}{1024} u_4 \right\} = 0. \tag{26}$$

Imposing this as r_1, the first element of the vector $\vec{r}$, gives a banded Jacobian matrix. The other elements of $\vec{r}$ are defined by

$$r_{j+1} = \tilde{r}_j \equiv \{u_{XXX} + (u - c)u_X\}|_{X=X_j}, \tag{27}$$

where the derivatives are evaluated by finite-difference approximations given in the Appendix.

4.3. Numerical Results: KdV Equation

With a finite-difference discretization, the deleted residual strategy is much cheaper than the QR-overdetermined Newton iteration. Nevertheless, since this QR–Newton has not been applied before, its performance is interesting.

Figure 2 shows a typical QR–Newton iteration for the KdV equation. For a given N, there is an inescapable finite-difference error, but in a conventional problem, the error in solving the *discretized algebraic system* can be reduced to machine precision. Here, however, the residual of the nonlinear algebraic equations reaches a finite minimum after a handful of iterations.

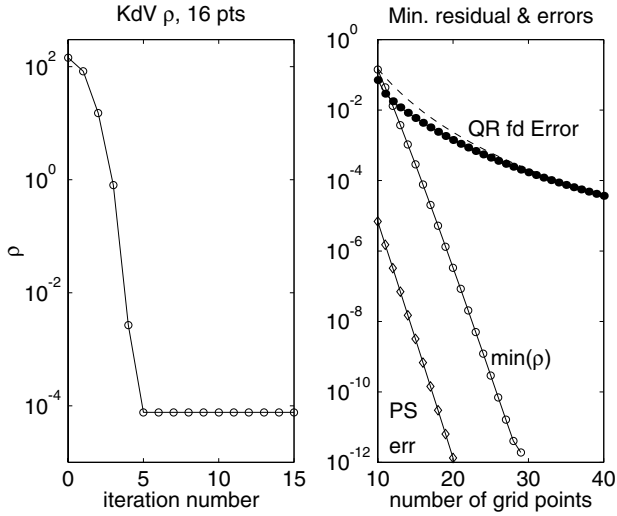

FIG. 2. (Left) The norm of the residual, ρ, versus iteration number for a typical KdV case. For a standard Newton's iteration where the number of equations equals the number of unknowns, ρ would decrease to machine precision. Here, however, it flattens at a finite minimum where $\rho_{min} = 7.7E-5$. (Right) Solid-line-with-diamonds: Error in the pseudospectral method versus the number of collocation points or degrees of freedom. Open circles: Minimum of the L_2 norm of the finite-difference/QR residual, ρ, versus the number of grid points, N. Solid circles (upper curve): The L_2 norm of the difference between the exact solution (Appendix D of [10]) and the finite-difference-with-constraint numerical solution. (Thin dashed curve) A constant multiplied by N^6; the error asymptotes to this, consistent with the expected sixth-order accuracy of the nine-point centered-difference approximation. Note that the error of the pseudospectral method with 10 grid points (or cosines) is smaller than the error of the sixth-order finite-difference method with 40 grid points! KdV cnoidal wave of phase speed $c = 0.08179$ and $P = 2\pi$ with the constraint $u(0) = 7.07031$. (These correspond to a solution with zero mean and lowest coefficient $a = 5$.)

Nevertheless, if N is sufficiently large, the least-squares solution to the overdetermined algebraic system, as found by the QR factorization, is a good approximation to the solution of the differential equation. The two upper curves of Fig. 2 confirm the expected sixth-order convergence of the finite-difference approximation to the exact solution. Although the formally overdetermined $(N+1) \times N$ linear algebra problem is satisfied only approximately at each iteration, the QR–Newton algorithm nevertheless gives very accurate results that improve with N at the expected rate.

The bottom curve (open circles) shows that for a given number of grid points N, the QR–Newton iteration converges to a minimum residual norm, $\rho_{min}(N)$. Remarkably, the straight-line appearance of the graph on this log-linear plot shows that ρ_{min} is decreasing *exponentially* with N. The error $u(X) - u_{approx}(X; N)$, in contrast, decreases only at a sixth-order rate and is orders-of-magnitude larger except for rather small N.

As the amplitude u_{max} increases, one needs more and more grid points to achieve a given accuracy, but the qualitative behavior of both error curves and $\rho_{min}(N)$ do not change.

The error for the much faster deleted residual strategy is not graphed separately because it is *visually indistinguishable* from the error $\|u - u_{approx}\|$ for the QR–Newton iteration. The deleted residual strategy is preferable because in Matlab on our Dell workstation, it was typically *six* times faster and required less than one-tenth as many operations as the QR-factored Newton scheme.

5. EQUATORIAL ROSSBY WEAKLY NONLOCAL SOLITARY WAVES

The theory of equatorial Rossby solitary waves has been developed and reviewed in a long series of articles and books [3, 8, 10, 25]. For small amplitude, the analytic perturbation theory of [3, 5] is accurate in predicting the "core" of the solitary waves.

Higher latitudinal mode solitary waves are "weakly nonlocal": the core of the coherent structure does not decay to zero at large distances from the maximum of the wave, but instead decays to a small-amplitude, quasi-sinusoidal oscillation. The existence of nonlocal equatorial waves was first predicted in [7]. The author's reviews and monograph [9, 11, 10] describe the general theory of nonlocal solitary waves and their spatially periodic generalizations; Figure 3 illustrates the mode-three Rossby nonlocal soliton for a typical case.

The equatorial waves solve the nonlinear shallow-water wave equations on the so-called equatorial beta-plane.

$$(u-c)u_X + vu_y - yv + \phi_X = 0, \tag{28}$$

$$(u-c)v_X + vv_y + yu + \phi_y = 0, \tag{29}$$

$$-c\phi_X + \{u\phi\}_X + \{v\phi\}_y = 0, \tag{30}$$

where c is the phase speed, ϕ is the total depth of the fluid, u and v are the nondimensional

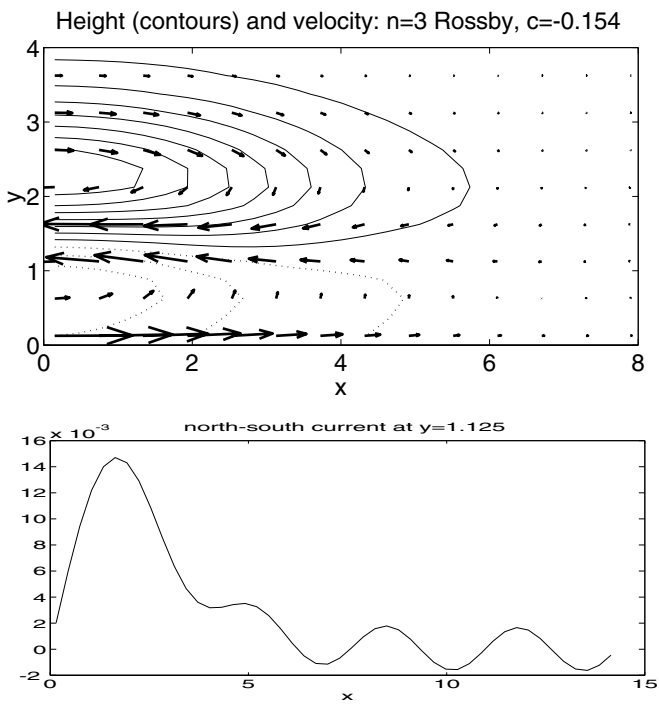

FIG. 3. (Top) Contours of the height anomaly ϕ' ($\phi > 0$ is solid; $\phi' < 0$ is dashed) with superimposed horizontal velocity vectors. (The anomaly $\phi' \equiv \phi - 1$ where ϕ is the total fluid depth.) Only the upper right quadrant of the X–y plane is shown because of the fourfold symmetry of all fields. (Bottom) A cross section of the north–south current at $y = 1.125$. The "far field" oscillations which make the wave "weakly nonlocal" are invisible on the top graph, which explains the adjective "weakly." The cross section shows clearly that the flow is not decaying to zero, but rather to small-amplitude oscillations, which is the "nonlocal" part.

east–west and north–south currents, $X = x - ct$ is longitude in a frame of reference that is moving with the wave, and y is nondimensional latitude. The parity symmetries are

$$u(X, y) = u(-X, y) = u(X, -y), \tag{31}$$

$$v(X, y) = -v(-X, y) = -v(X, -y), \tag{32}$$

$$\phi(X, y) = \phi(-X, y) = \phi(X, -y). \tag{33}$$

Thus, the same symmetry-exploiting derivative matrices as used for the KdV equation can be applied in both longitude and latitude here. The only important difference is that the boundary conditions for large $|y|$ are

$$u, v \to 0, \quad \phi \to 1, \quad |y| \to \infty. \tag{34}$$

These are imposed by omitting weights in the bottom four rows of the y-differentiation matrices, implicitly assuming that the grid point values of all unknowns at latitudes beyond the largest latitudinal grid point are zero.

The complication is that in fluid mechanics in general and equatorial flows in particular, solitary and cnoidal waves are not unique without the specification of a mean [i.e., zonally averaged] east–west current $U(y)$. This implies that the number of extra constraints is equal to the number of latitudinal grid points n_y. Denoting the number of east–west grid points by n_x, the rectangular Jacobian matrix is of size $(3n_x n_y + n_y) \times (3n_x n_y)$.

Unfortunately, a rigorous proof of this assertion is lacking. However, perturbation theory shows that for weak mean currents, the zonally averaged flow $U(y)$ is arbitrary—for any mean flow, so long as it is small, one can find a solitary wave. [4, 17, 22, 23]. Without constraints on the mean flow, our numerical computations show that the Jacobian matrix is always singular.

Our detailed results for equatorial solitary waves are published in [13]. There we explain why it is preferable, to produce simpler agreement with theory, to employ the modified constraints

$$\int_{P/4}^{P/2} u(X, y_j)\, dx = 0, \quad j = 1, 2, \ldots, n_y, \tag{35}$$

which were discretized using the rectangle rule.

As in the KdV case, the most efficient finite-difference strategy is deleted residuals. It is not possible to eliminate the arbitrary mean current by any reformulation of the problem. However, as for the KdV cnoidal wave, the imposition of symmetry on the finite-difference method both enormously reduced cost (by a factor of four in two dimensions) and also suppressed the translational degree of freedom.

The QR–Newton iteration was also very similar to the KdV problem: the residual of the discretized algebraic system of equations ρ had a finite minimum for each set of (n_x, n_y), as illustrated in Fig. 4. Also like the KdV problem, the minimum residual seems to decrease faster with the number of grid points than does the finite-difference error.

However, there are a couple of puzzling differences. One is that the exponential decrease of ρ_{min} for the KdV cnoidal wave has been replaced by a (roughly!) tenth-order decrease. Another difference is that on a log–log plot (not shown), the KdV ρ_{min} is concave downward whereas the residual minimum for nonlocal equatorial waves is concave upward.

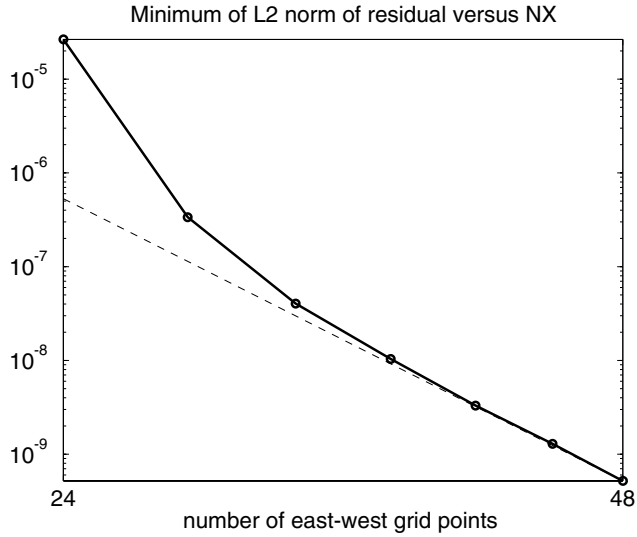

FIG. 4. Nonlocal equatorial waves. The heavy solid curve with circles is the L_2 norm of the minimun residual for a given number of east–west grid points; the number of latitudinal points is half that of the number of the east–west points. The scale is logarithmic for both the residual and the number of grid points. Although the solution used eight-order differences, the thin dashed line is proportional to n_x^{-10}, which gives a better fit. The phase speed $c = -0.15411$; the computational domain was $X \in [0, 10] \otimes Y \in [0, 5]$.

Because of the complexity and cost of the problem and also because the finite-difference/deleted residuals method was so successful, we did not attempt to apply the Keller bordered matrix algorithm, nor the pseudoinverse, nor did we do any pseudospectral calculations for this problem. We conjecture, because of the similar results for the methods that we did apply to the KdV case, that all these untested algorithms would have been performed well, if not always rapidly, for nonlocal equatorial Rossby waves.

6. SHOCKS OF THE KORTEWEG–DEVRIES–BURGERS (KDV–BURGERS) EQUATION

The Korteweg–deVries–Burgers (KdV–Burgers) equation for steadily traveling shocks, first proposed by Grad and Hu [16] for plasmas and by Johnson [19] for water waves, is

$$u_{xx} - \nu u_x + (u/2 - c)u = 0, \tag{36}$$

where c is the phase speed. Bona and Schonbek [2] prove existence and uniqueness of the shocks. As $x \to \pm\infty$ where the shocks asymptote to constants $u_{\pm\infty}$, the differential equation collapses to $(u/2 - c)u \sim 0$, which implies that the constants $u_{\pm\infty}$ must be either 0 or 2c. It is easy to show that if $v(x; \nu)$ is a shock for unit phase speed, then a shock for general c is given by $u = cv(\sqrt{c}x; \nu\sqrt{c})$. Thus, there is no loss in generality in specializing to $c = 1$ and solving

$$u_{xx} - \nu u_x + (u/2 - 1)u = 0, \quad u \to 2 \text{ as } x \to -\infty, \quad u \to 0 \text{ as } x \to \infty, \tag{37}$$

for $\nu > 0$.

The KdV–Burgers shock is translationally invariant: if $u(x)$ is such a solution, so is $u(x+\Phi)$ for arbitrary constant Φ. Because the shock is not symmetric with respect to $x=0$, unlike our cnoidal and solitary wave examples, it is not possible to remove the translational eigenmode by imposing a constraint of symmetry: reformulation *fails*.

6.1. Numerical Methods for KdV–Burgers Shocks

We computed the shocks using four different numerical methods: fourth-order finite differences, a Fourier pseudospectral method, and a rational Chebyshev pseudospectral algorithm. For the finite-difference method, the domain was the truncated interval $x \in [-x_B, x_B]$ with $u(\pm x_B) = u_{\pm\infty}$; the domain size x_B was chosen to be sufficiently large that u at $x = \pm x_B$ differs negligibly from its asymptotic limits.

The Fourier approximation consisted of (i) a linear polynomial plus (ii) an $(N+1)$-term trigonometric polynomial:

$$u = 2(P/2 - x)/P + \sum_{j=0}^{N/2} a_j \cos(j[2\pi/P]x) + \sum_{j=1}^{N/2} a_{N/2+j} \sin(j[2\pi/P]x). \tag{38}$$

In the deleted residual method, the residual was forced to be zero at a set of discrete points, but at only N points instead of the usual $(N+1)$:

$$x_j = -P/2 + (j-1)/(PN), \quad j = 1, 2, \ldots, N. \tag{39}$$

The constraint $u(0) = 1$ is imposed as the $(N+1)$st element of the residual vector

$$\sum_{j=0}^{N/2} a_j = 1. \tag{40}$$

This discretization gives a *nonsingular* system of $(N+1)$ algebraic equations in $(N+1)$ unknowns a_j.

The optimum choice of spatial period increases with N; a larger domain is necessary for larger N so that the errors $(u(\pm P/2) - u_{\pm\infty})$ at the edges of the truncated domain do not exceed the errors in truncating the Fourier series after N. After some experimentation, we set

$$P = \sqrt{50}\sqrt{N}. \tag{41}$$

The rational Chebyshev method [12] is equivalent to making the change-of-coordinate

$$x = L\cot(t) \quad \leftrightarrow \quad t = \operatorname{arccot}(x/L) \tag{42}$$

and then applying a Fourier cosine pseudospectral method. The solution is approximated by

$$u(x) \approx \sum_{j=0}^{N} a_j \cos(jt[x]). \tag{43}$$

N algebraic conditions come from collocation at the points

$$t_j = \pi \frac{2j-1}{2N}, \quad x_j = L\cot(t_j), \quad j = 1, 2, \ldots, N. \tag{44}$$

In the deleted residuals method, the $(N+1)$st condition is $u(0) = 1$, which in spectral terms is

$$\sum_{j=0}^{N} a_j = 1. \tag{45}$$

6.2. Spectral Upgrade

> Every low order BVP-solver contains a spectral solver within it. . . . Preconditioned iterations provide a way to promote a low order method to spectral accuracy. (J. P. Boyd [12], p. 301)

The notion of upgrading a low-order finite-difference or finite-element method to spectral accuracy is stated most clearly in Sect. 13.6 of the book [12], but the idea is implicit in the articles of Orszag [29] and Morchoisne [26], who independently devised the idea of solving pseudospectral discretization matrices by finite difference–preconditioned iteration, and of Deville and Mund [15], who showed that finite-element preconditioning worked even better. Historically, these authors began with pseudospectral discretizations and asked: How can we efficiently solve these dense matrix problems?

From a practical standpoint, this is looking at the telescope the wrong way around. The most interesting question is: Given a fast-but-inaccurate finite-difference computer code, how can it be improved, given that a spectral algorithm can achieve many more orders of accuracy on the same grid? The answer is to replace the finite-difference evaluation of the residual in the end game by a subroutine that evaluates the residual to pseudospectral accuracy. In our case, we used the grid-point values of $u(x)$ on the finite-difference grid, computed the Fourier interpolant, and evaluated the residual at each grid point by differentiation and evaluation of the Fourier series. As noted in [12], all the spectral stuff can be isolated in a single subroutine; the Jacobian matrix is the finite-difference matrix, and the main program knows only that a "black box" is taking the grid-point values of u as input and returning the residual of the differential equation.

As shown by Orszag, Morchoisne, and Deville and Mund, the finite-difference or finite-element Jacobians are sufficiently good approximations to the spectral Jacobian that the iteration—a quasi-Newton iteration because the true Jacobian (spectral) has been replaced by an approximation (finite difference)—still converges geometrically fast until the *spectral* residual is negligibly small. Thus, Newton's iteration, with merely the replacement of one residual evaluation routine by another, yields full spectral accuracy.

Spectral upgrades have not been previously applied when the underlying square matrix is singular, nor a QR–Newton iteration that solves a *rectangular* matrix at each step. We shall see below, however, that the finite-difference-to-spectral upgrade is successfull for Korteweg–deVries–Burgers shocks.

6.3. Algorithmic Failures

For the finite-difference method and shocks, deleted residuals failed even though we experimented with several variations. First, the residual at the central grid point, $x = 0$, was replaced by the constraint $u(0) = 1$. Second, the grid was staggered so that the grid points

nearest the origin were at $x = \pm h/2$; the residual at one was replaced by the (interpolated) condition $u(0) = 1$. This failed, too; the discrete solution satisfied the constraint but had a discontinuous slope at the origin. Third, we deleted two residuals and imposed both the condition $u(0) = 0$ and a condition of slope continuity at $x = 0$, both expressed through interpolation formulas as accurate as those for the derivatives in the differential equation. This did not work either.

With the pseudospectral method, one could use a grid of N points instead of $(N + 1)$ without complication since there is no penalty for evaluating the spectral series at an arbitrary point. In contrast, the finite-difference methods with $(N + 1)$ degrees of freedom could be easily posed only on a grid of $(N + 1)$ points, too, since the unknowns are defined as the values of $u(x)$ on a grid of points. In the spectral case, it was easy to squeeze the N points together to avoid leaving the gap that was inevitable with the finite-difference method.

We applied Keller's method, but this, too, failed. When a row and column is removed from a generic matrix with a single zero eigenvalue, the smaller matrix is nonsingular, as assumed by Keller's bordered matrix scheme. (By "generic," we mean a matrix which is randomly chosen from the set of all matrices of a given size with a single zero eigenvalue.) However, the finite-difference discretization of the KdV–Burgers equation is *not generic*, and its upper left $(N - 1) \times (N - 1)$ block is *singular, too*.

This was so surprising that we checked the finite-difference discretization by successfully applying the pseudoinverse/SVD method. The discretization was fine, but the bordered matrix algorithm failed.

Since the pseudoinverse method is quite expensive, the cheapest *successful* finite-difference method was to solve a formally overdetermined system by appending $u(0) = 1$ to an N-point discretization to obtain a system of $(N + 1)$ equations in only N unknowns, and then applying the QR–Newton iteration.

In contrast, the deleted residual strategy was successful for both spectral methods.

6.4. Errors

Figure 5 compares the errors for the first three methods. As expected, the spectral methods are much more efficient, both in terms of fewer grid points (left panel) and computer time (right panel). The rational Chebyshev method seems to be a little more efficient than the Fourier algorithm, but both converge at an exponential rate with N.

Nevertheless, if one is content with a modest accuracy of a few percent, the QR finite-difference method is competitive with the spectral methods. It is remarkable that for this problem, the QR-overdetermined Newton iteration is the cheapest successful way with a finite-difference discretization.

In contrast, the much less expensive deleted residual strategy was successful for the KdV and equatorial Rossby wave problems even with a finite-difference discretization. Why the difference? We can only speculate, but the lack of symmetry in the shock waves is probably the culprit.

Figure 6 shows that the spectral upgrade is very effective. The left panel illustrates that for a given number of points, the spectral upgrade is much more accurate by orders of magnitude than the pure finite-difference program, and yet the two differ only in the single subroutine that evaluates the residual of the differential equation (plus the constraint $u(0) = 1$). The accuracy is not quite as good as the Fourier pseudospectral method for the same N; the residual of the spectral upgrade does not quite go to zero because there is a tiny component of the error which is not properly preconditioned by the finite-difference

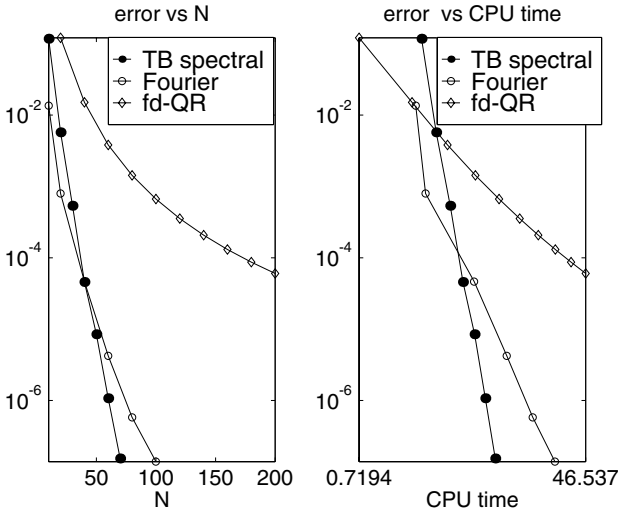

FIG. 5. Comparison of numerical errors for computing shocks of the KdV–Burgers equation, plotted versus the number of grid points N (left) or CPU time (right, with a logarithmic time scale). Black disks: Rational Chebyshev method with map parameter $L = 2\sqrt{N}$. Fourier pseudospectral method with the spatial period $P = \sqrt{50N}$. Diamonds (upper curve in each panel): Fourth-order finite differences with QR-overdetermined Newton iteration with $x_B = 25$. All calculations performed using Matlab 5.2 on Apple G4 at 450 Mhz.

Jacobian. (In the usual, nonsingular case, Orszag shows that the eigenvalues of the product of the inverse of the finite-difference matrix times the pseudospectral Jacobian is a matrix whose eigenvalues all lie in the narrow range between 1 and 2.4, but this is not true here.) Nevertheless, the right panel shows that for achieving a given accuracy, the spectral upgrade is far faster than either of its "parents," the Fourier pseudospectral method and the fourth-order finite-difference/QR scheme.

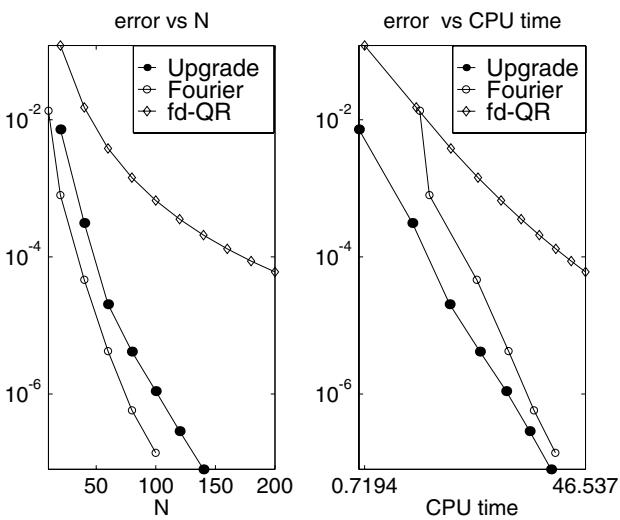

FIG. 6. Same as previous figure except that the TB (rational Chebyshev) pseudospectral method was replaced by the Fourier pseudospectral upgrade of the finite-difference QR algorithm. The upgrade used the same N-dependent domain as the Fourier pseudospectral method.

7. CONCLUSIONS

When a naive discretization of a nonlinear eigenvalue problem is singular, additional constraints must be appended to obtain a unique solution. The cheapest fix is to replace residual conditions from the differential equation by the constraints. This simple-minded deleted residuals method worked for all combinations of discretizations and test problems except for finite-difference discretizations of shocks of the Korteweg–deVries–Burgers equation. For that exceptional case, the best algorithm is to apply the QR factorization to a rectangular Jacobian.

The most efficient discretization is to apply a "spectral upgrade" to a finite-difference program. The code is identical to a pure finite-difference method except for the addition of a single subroutine to evaluate the residual spectrally.

One novelty of this article is that spectral upgrading has never previously been applied with rectangular matrices or QR solution of the finite-difference preconditioner. (Actually the finite-difference QR–Newton method, with or without upgrade, has never been previously applied to singular nonlinear problems.) Fourth-order differences-with-Fourier upgrade is the fastest algorithm for the KdV–Burgers shocks.

At least three issues remain for future research. First, it would be desirable to develop better tools for predicting the number of constraints required for a given wave problem. Our equatorial Rossby constraints, one per latitudinal grid point, are justified only by perturbation theory and empirical success.

Second, there is no theory explaining either the success or the failure of the deleted residuals strategy. For the pseudospectral algorithms, it easy to squeeze the grid so as to leave no obvious gap. This perhaps explains why deleting residuals always works with these algorithms though it failed for shocks with finite differences where there was a gap at the origin, left by the residual that was replaced with the constraint $u(x=0)=1$. Speculation and plausibility arguments are no substitute for theory, however.

Third, the preconditioned QR–Newton iteration with a Fourier upgrade did not converge to zero residual, only to a very small residual. Although the algorithm was successful anyway in the sense that the finite-difference-with-Fourier-upgrade for the shocks gave accuracies that decreased exponentially with N and were only slightly less than those of the (much slower) pure Fourier pseudospectral algorithm, it would be nice to understand why this minor difficulty arises. Until this is understood, the spectral upgrade method for singular wave equations must come with a mild warning sticker: fast algorithm that *probably* will work.

Even with these unresolved issues, it is clear that there are good methods for solving wave problems that must be constrained, beyond the differential equations and boundary conditions, to specify a unique solution. Pseudospectral methods and the QR–Newton iteration are the most robust; finite-difference-with-spectral-upgrade and the deleted residuals strategy are fastest.

APPENDIX: PARITY AND FINITE DIFFERENCES

To approximate derivatives for both the KdV and equatorial wave equations, we use centered nine-point differences, which are sixth-order accurate for the third derivative of the KdV equation, and eighth-order accurate for the first derivative KdV term and for all derivatives of the equatorial wave equations, which contain only first derivatives. With h as

the grid spacing, the approximations are

$$\frac{df^j}{dx^j}(x) \approx w_4 f(x+4h) + w_3 f(x+3h) + w_2 f(x+2h) + w_1 f(x+h) - w_1 f(x-h) - w_2 f(x-2h) - w_3 f(x-3h) - w_4 f(x-4h). \quad \text{(A.1)}$$

The first derivative weights are

$$w_1 = \frac{4}{5}\frac{1}{h}, \quad w_2 = -\frac{1}{5}\frac{1}{h}, \quad w_3 = \frac{4}{105}\frac{1}{h}, \quad w_4 = -\frac{1}{280}\frac{1}{h}. \quad \text{(A.2)}$$

The third derivative weights are

$$w_4 = \frac{7}{240}\frac{1}{h^3}, \quad w_3 = -\frac{3}{10}\frac{1}{h^3}, \quad w_2 = \frac{169}{120}\frac{1}{h^3}, \quad w_1 = -\frac{61}{30}\frac{1}{h^3}. \quad \text{(A.3)}$$

(Neither of the problems studied here has even-order derivatives, but this is just a coincidence and not intrinsic to wave equations that require additional constraints to specify unique solutions.)

The boundary conditions in X are spatial periodicity; that is,

$$u(X) = u(X+P), \quad \forall X \quad \text{[periodic boundary condition]}. \quad \text{(A.4)}$$

It is conceptually useful to visualize the interval for a periodic problem as a *ring*; this mental model makes it obvious that we can apply centered differences everywhere on the grid. Near $X = P/2$, where the stencil requires values of u at $X > P/2$, we can invoke the periodicity to replace the missing values of $u(P/2 + jh)$ by $u(-P/2 + jh)$. The resulting differentiation matrices are so-called "cyclic" matrices. These are banded matrices except for small triangular blocks in the upper right and lower left corners [27, 31].

Because the KdV cnoidal wave is symmetric with respect to $X = 0$, we choose an evenly spaced grid that spans only half the spatial period, $P = 2\pi$:

$$X_j = \pi\frac{2j-1}{2N}, \quad j = 1, 2, \ldots, N. \quad \text{(A.5)}$$

To apply centered differences near the ends of the grid, we use symmetry and periodicity to supply missing grid point values. Abbreviating $u(X_j)$ by u_j, symmetry with respect to $X = 0$ gives

$$u_{-j} = u_j, \quad j = 1, 2, 3, 4. \quad \text{(A.6)}$$

Periodicity on an evenly spaced grid implies that $u_{N+j} = u_{-N+j}$ for all j. Symmetry with respect to the origin demands $u_{-N+j} = u_{N-j}$. The two constraints together supply the needed values near $X = P/2$:

$$u_{N+j} = u_{N-j}, \quad j = 1, 2, 3, 4. \quad \text{(A.7)}$$

In the array below, for example, the second weight in each element is derived by these symmetry constraints. For a nine-point grid, for either the first or third derivatives with

appropriate w_j, the differentiation matrix is

$$\begin{vmatrix}
0-w_1 & w_1-w_2 & w_2-w_3 & w_3-w_4 & w_4 & 0 & 0 & 0 & 0 \\
-w_1-w_2 & 0-w_3 & w_1-w_4 & w_2 & w_3 & w_4 & 0 & 0 & 0 \\
-w_2-w_3 & -w_1-w_4 & 0 & w_1 & w_2 & w_3 & w_4 & 0 & 0 \\
-w_3-w_4 & -w_2 & -w_1 & 0 & w_1 & w_2 & w_3 & w_4 & 0 \\
-w_4 & -w_3 & -w_2 & -w_1 & 0 & w_1 & w_2 & w_3 & w_4 \\
0 & -w_4 & -w_3 & -w_2 & -w_1 & 0 & w_1 & w_2 & w_3+w_4 \\
0 & 0 & -w_4 & -w_3 & -w_2 & -w_1 & 0 & w_1+w_4 & w_2+w_3 \\
0 & 0 & 0 & -w_4 & -w_3 & -w_2 & -w_1+w_4 & w_3 & w_1+w_2 \\
0 & 0 & 0 & 0 & -w_4 & -w_3+w_4 & -w_2+w_3 & -w_1+w_2 & 0+w_1
\end{vmatrix} . \tag{A.8}$$

The matrices for larger numbers of points are similar except that additional rows, centered on the diagonal and without boundary modifications, must be inserted in the middle.

In the equatorial problem, the north–south current v is *antisymmetric* with respect to $X = 0$; that is,

$$v(-X, y) = -v(X, y), \quad \forall X, y. \tag{A.9}$$

The same differentiation matrix applies except that signs of the second weight must be changed in each element with two weights.

In latitude, the equatorial waves also have definite parity with respect to the equator. For large y, however, the proper boundary conditions are that all wave fields decay exponentially to zero. This condition was imposed by simply omitting the second weight of each pair in the four bottom rows of the latitudinal differentiation matrix, which is otherwise identical to the east–west (X) differentiation matrix.

As explained in the main text, parity automatically imposes the "translational" constraint, thereby eliminating one row from the rectangular Jacobian matrix. It also has another helpful effect: the differentiation matrices become ordinary banded matrices rather than cyclic matrices. This reduces costs, albeit only modestly; cyclic matrices, too, can be factored very efficiently [27, 31].

ACKNOWLEDGMENTS

Supported by NSF Grant OCE9986368. I thank the Associate Editor and the three reviewers for helpful comments.

REFERENCES

1. G. W. Bluman and J. D. Cole, *Similarity Methods for Differential Equations* (Springer-Verlag, New York, 1974).
2. J. L. Bona and M. E. Schonbek, Travelling-wave solutions to the Korteweg–deVries–Burgers equation, *Proc. R. Soc. Edinburgh* **101A**, 207 (1985).
3. J. P. Boyd, Equatorial solitary waves, Part I: Rossby solitons, *J. Phys. Oceanogr.* **10**, 1699 (1980).
4. J. P. Boyd, Equatorial solitary waves, Part IV: Kelvin solitons in a shear flow, *Dyn. Atmos. Oceans*, **8**, 173 (1984).
5. J. P. Boyd, Equatorial solitary waves, Part 3: Modons, *J. Phys. Oceanogr.* **15**, 46 (1985).

6. J. P. Boyd, Solitons from sine waves: analytical and numerical methods for non-integrable solitary and cnoidal waves, *Physica D*, **21**, 227 (1986).
7. J. P. Boyd, Non-local equatorial solitary waves, in *Mesoscale/Synoptic Coherent Structures in Geophysical Turbulence: Proc. 20th Liege Coll. on Hydrodynamics*, edited by J. C. J. Nihoul and B. M. Jamart (Elsevier, Amsterdam, 1989), p. 103.
8. J. P. Boyd, Nonlinear equatorial waves, in *Nonlinear Topics of Ocean Physics: Fermi Summer School, Course LIX*, edited by A. R. Osborne (North-Holland, Amsterdam, 1991), p. 51.
9. J. P. Boyd, Weakly nonlocal solitary waves, in *Nonlinear Topics of Ocean Physics: Fermi Summer School, Course LIX*, edited by A. R. Osborne (North-Holland, Amsterdam, 1991), p. 527.
10. J. P. Boyd, *Weakly Nonlocal Solitary Waves and Beyond-All-Orders Asymptotics: Generalized Solitons and Hyperasymptotic Perturbation Theory*, Mathematics and Its Applications (Kluwer, Amsterdam, 1998), Vol. 442.
11. J. P. Boyd, The devil's invention: Asymptotics, superasymptotics and hyperasymptotics, *Acta Appli.* **56**, 1 (1999).
12. J. P. Boyd, *Chebyshev and Fourier Spectral Methods* (Dover, Mineola, NY, 2001), 2nd ed.
13. J. P. Boyd, Equatorial solitary waves. Part 5. Initial value experiments, co-existing branches and tilted-pair instability, *J. Phys. Oceangr.*, in press.
14. J. E. Dennis, Jr. and R. B. Schnabel, *Numerical Methods for Nonlinear Equations and Unconstrained Optimization* (Prentice-Hall, Englewood Cliffs, NJ, 1983).
15. M. Deville and E. Mund, Chebyshev pseudospectral solution of second-order elliptic equations with finite element pre-conditioning, *J. Comput. Phys.* **60**, 517 (1985).
16. H. Grad and P. N. Hu, Unified shock profile in a plasma, *Phys. Fluids* **10**, 2596 (1967).
17. R. J. Greatbatch, Kelvin wave fronts, Rossby solitary waves and nonlinear spinup of the equatorial oceans, *J. Geophys. Res.* **90**, 9097 (1985).
18. P. E. Hydon, *Symmetry Methods for Differential Equations: A Beginner's Guide*, Cambridge Texts in Applied Mathematics (Cambridge Univ. Press, Cambridge, UK, 2000), Vol. 22.
19. R. S. Johnson, A nonlinear equation incorporating damping and dispersion, *J. Fluid Mech.* **42**, 49 (1970).
20. H. B. Keller, The bordering algorithm and path following near singular points of higher nullity, *SIAM J. Sci. Stat. Comput.* **4**, 573 (1983).
21. H. B. Keller, *Numerical Methods for Two-Point Boundary-Value Problems* (Dover, New York, 1992).
22. M.-T. Li, Equatorial solitary atmospheric waves in meridional shear flow, *Trop. Ocean Atmos. Newsl.* **1**, 15 (1984).
23. B. Long and P. Chang, Propagation of an equatorial Kelvin wave in a varying thermocline, *J. Phys. Oceanogr.* **20**, 1826 (1990).
24. R. Mallier, Stuart vortices on a beta-plane *Dyn. Atmos. Oceans* **22**, 213 (1995).
25. H. G. Marshall and J. P. Boyd, Solitons in a continuously stratified equatorial ocean, *J. Phys. Oceanogr.* **17**, 1016 (1987).
26. Y. Morchoisne, Resolution of Navier-Stokes equations by a space-time pseudospectral method, *Rech. Aerosp.* **5**, 293 (1979).
27. I. M. Navon, PENT: A periodic pentadiagonal systems solver, *Commun. Appl. Numer. Methods* **3**, 63 (1987).
28. P. J. Olver, *Applications of Lie Groups to Differential Equations* (Springer-Verlag, New York, 1986).
29. S. A. Orszag, Spectral methods for problems in complex geometries, *J. Comput. Phys.* **37**, 70 (1980).
30. L. G. Redekopp, On the theory of solitary Rossby waves, *J. Fluid Mech.* **82**, 725 (1977).
31. C. Temperton, Algorithms for the solutions of cyclic tridiagonal systems, *J. Comput. Phys.* **19**, 317 (1975).