Why Newton’s method is hard for travelling waves: Small denominators, KAM theory, Arnold’s linear Fourier problem, non-uniqueness, constraints and erratic failure

John P. Boyd
Department of Atmospheric, Oceanic and Space Science, 2455 Hayward Avenue, Ann Arbor, MI 48109, United States

Available online 12 December 2006

Abstract
Nonlinear travelling waves and standing waves can be computed by discretizing the appropriate partial differential equations and then solving the resulting system of nonlinear algebraic equations. Here, we show that the “small denominator” problem of Kolmogorov–Arnold–Moser (KAM) theory is equally awkward for numerical algorithms. Furthermore, Newton’s iteration combined with continuation in a parameter often exhibits “erratic failure” even in the absence of bifurcation. Wave resonances can interlock a countable infinity of branches in an extremely complex topology, as will be illustrated through the fifth-degree Korteweg–deVries equation. Continuation can easily jump, unsuspected, from one branch to another. Constraints, sometimes finite and sometimes infinite in number, are usually needed to specify a unique solution. This confluence of numerical difficulties can be overcome only by combining the latest numerical algorithms with a strong understanding of travelling wave physics.

© 2006 IMACS. Published by Elsevier B.V. All rights reserved.

Keywords: Nonlinear equations; Bifurcation; Small denominators; Fifth-order Korteweg–deVries equation

“...to do this business right is a thing of far greater difficulty than I was aware of.”
Sir Isaac Newton, in a letter to Edmund Halley.

1. Introduction

For a quarter of a century, I have been computing travelling waves as the solutions to nonlinear eigenvalue problems [4,3]. Even so, the study of the equatorially-trapped ocean Kelvin wave, which began in 1980, is still a work in progress [7].

The reason is that solving a system of nonlinear algebraic equations, as results from discretizing the partial differential equations for a set of travelling waves, is still an unsettled research frontier. There is a whole spate of recent books with recipes for nonlinear rootfinding [1,12,13,8,9]. One might be pardoned for thinking that nonlinear eigenvalue systems are a solved problem, reduced merely to a software download. This article will argue the opposite view.
2. A linear illustration of the small denominator problem

Arnold [2], pp. 112–115, shows through the “homological” equation that even linear equations may have complicated solutions.

In contrast, the differential equation

\[ v_{xx} = -f(x), \quad v(x + 2\pi) = v(x) \quad \forall x \]

is trivially solved by expanding both sides in a Fourier series:

\[ f(x) = \sum_{n=1}^{\infty} f_n \cos(nx), \quad v(x) = \sum_{n=1}^{\infty} \frac{f_n}{n^2} \cos(nx) \]

Suppose we modify the problem slightly by replacing the differential operator by a difference operator:

\[ u(x + h) - 2u(x) + u(x - h) = -f(x) \]

where \( h \geq 0 \) is the grid spacing. The solution is

\[ u(x; h) = \sum_{n=1}^{\infty} \frac{h^2 f_n}{\sin^2(nh)} \cos(nx) \]

In the limit \( h \to 0 \), the difference operator is equal to the second derivative operator. It would seem as though the change is trivial.

However, if we introduce

\[ H = \frac{h}{\pi}, \]

the \( n \)th term in the Fourier series for the difference solution \( u(x) \) is infinite whenever the following resonance condition is satisfied:

\[ H = \frac{m}{n}, m, n \text{ integers} \]

In other words, whenever \( H \) is a rational number, the infinite Fourier series for \( u(x) \) has at least one unbounded term!

Similar difficulties arise in the Poincaré–Lindstedt/Stokes expansions of celestial mechanics and travelling wave problems. Just as for Arnold’s example, these expansions are also Fourier series in which some of the terms have arbitrarily small denominators. Such “small denominator” problems are still big trouble more than a century after they were first identified.

Number theory provides more bad news. First, the geometry of the unboundedness of \( u(x; \pi H) \) is very complicated because every finite interval in \( H \) contains an infinite number of rationals interspersed with irrational numbers. Second, a theorem asserts that for any irrational number, one can always find a rational number close to it in the following sense.

**Theorem 1** (Approximation of an irrational by a rational). For any irrational number \( H \), there are rational approximations of arbitrary accuracy such that

\[ |H - \frac{m}{n}| < \frac{1}{n^2}, \quad m, n \text{ integers} \]

A proof and algorithm are in [2], p. 110.

In reality, the situation is much less dire. First, if \( f(x) \) is a periodic function which is analytic in a strip of width \( |\Im(x)| \leq \mu \) in the complex \( x \)-plane, then its Fourier coefficients must satisfy the bound

\[ |f_n| \leq M \exp(-\tilde{\mu} n) \]

for some positive constant \( M \) and some \( \tilde{\mu} \) which is smaller than \( \mu \) by an arbitrarily tiny amount. It follows that \( f(x) \) can be approximated by the \( N \)-term truncation of its Fourier series with an error which is exponentially small in \( N \).
Fig. 1. Fourier coefficients for the solution of Arnold’s difference equation when the forcing is the first 20 non-zero cosine coefficients of $f(x) = (1/2)((1 - p^2)/(1 + p^2) - 2p \cos(x)) - 1 = \sum_{n=1}^{\infty} p^n \cos(nx)$. The 20 curves shown are the values of $h^2 f_n / \sin^2(n\pi H)$ for $n = 1, 2, \ldots, 20$. The curves need no labels because in the left-half of the interval, the curves are vertically stacked from $n = 1$ at the top to $n = 20$ at the bottom.

On a computer, $f(x)$ is always approximated, very accurately, by a trigonometric polynomial and then the number of resonances is finite on any finite interval in $H$, restricted to $H = m/n$ with $n \leq N$.

Second, difference operators are normally applied only with small $h$. If the interval is restricted to $H \in [0, 1/N - \delta]$ for some $\delta > 0$, then there are no resonances. In most small denominator problems, one is usually not quite that lucky, but it is often true that restrictions on the parameter range reduce the number of resonances.

Fig. 1 is a numerical example. Although there are 10 resonances on the (arbitrarily chosen) interval, $H \in [0, 1/10]$, the left-half of the interval is resonance-free. On the right-half of the interval, each resonance produces a narrow spike. However, even though the graph was made using a rather dense grid (1001 points on $H \in [0, 1/10]$), none of the spikes is graphically unbounded. Indeed, the left five of the spikes never rise above 1/100 whereas the peak of $u(x; \pi H)$ is about 0.4.

The analytical solution (4) shows that for each of 10 discrete values on $H \in [0, 1/10]$, $u(x; \pi H)$ is infinite, a complication that will not go away. However, the graph shows that if one picked a random value on this same interval, the probability is very high that the resonances would be invisible!

Theoretical treatments of small denominators use the following lemma:

**Theorem 2.** Let $\sigma$ denote an arbitrary positive number. Then for almost every real $H$ there exists $K(H, \sigma)$ such that

$$\left| H - \frac{m}{n} \right| \geq \frac{K}{|n|^{2+\sigma}} \quad (9)$$

A number $H$ for which the inequality is true is said to be of “type ($K, \sigma$). The set of numbers for which this inequality fails is of measure zero ([2], p. 114).

It then follows that a solution to the difference equation exists for almost all $H$. The bland “almost all”, however, hides a geometry of many tiny intervals where the resonance makes $u(x; \pi H)$ huge, and utterly changes the solution. The sets in $H$ where solutions are bounded by some constant $Q$ are Cantor sets.

3. Newton’s iteration

“Solving systems of nonlinear equations is perhaps the most difficult problem in all of numerical computation. It is not unheard of for even 5 equations to be very difficult to solve; one can “hide” a solution in 5-dimensional space rather easily.”

John R. Rice [15], p. 355.
Nonlinear differential equations, after discretization by spectral or finite difference methods, yield a nonlinear system of algebraic equations of the form \( \vec{r}(\vec{u}) = \vec{0} \) where \( \vec{r} \) is a vector whose elements are the residuals of the discretized differential equation and \( \vec{u} \) is composed of the grid point values or spectral coefficients of \( u(x) \). The most popular strategy to solve such systems is Newton’s iteration, which in generalized form is

\[
\vec{u}^{(m+1)} = \vec{u}^{(m)} - \tau \delta, \quad \frac{\delta}{\|\delta\|} = \frac{\vec{r}(\vec{u}^{(m)})}{\|\vec{r}(\vec{u}^{(m)})\|} \tag{10}
\]

where the elements of the “Jacobian” matrix are \( J_{ij} \equiv \partial r_i/\partial u_j \) and \( \tau \) is a constant, the “relaxation parameter”; classically, \( \tau = 1 \).

Unfortunately, Newton’s iteration requires an “initialization” or “first guess” \( \vec{u}^{(0)} \), and fails to converge if the initialization is too far from a root. Three popular strategies for expanding the domain of convergence are:

1. line search (underrelaxation);
2. continuation;
3. minimization/steepest descent thinking.

The “Armijo line search” is a simple, computationally cheap underrelaxation. Define \( \tau_j \equiv 1/2^{(j-1)/2} \) for \( j = 0, 1, 2, \ldots, j_{\text{max}} \) and accept whichever \( \tau_j \) minimizes \( \|\vec{r}(\vec{u}^{(m)} - \tau_j \delta)\| \).

Continuation was inspired by the fact that (i) most real-world problems depend on a parameter \( \lambda \) and (ii) it is usually easy to guess the solution for extreme values of \( \lambda \). One may then march from \( \lambda_0 \) to other values of \( \lambda \) in small steps, using the converged answer for \( \lambda = \lambda_j \) as the first guess for \( \lambda = \lambda_{j+1} \). If the steps in \( \lambda \) are sufficiently small, one can successfully trace an entire branch of solutions even when initially clueless about \( u(x; \lambda) \) for some parameter values. The prediction can be improved by fitting a polynomial through the last few solutions to extrapolate a first guess for \( \vec{u}(\lambda_{j+1}) \).

When continuation and line search both fail, a third strategy is to recall that a solution to the system of \( N \) equations in \( N \) unknowns is also a global minimum of the \( L_2 \) norm of the residual,

\[
\rho(\vec{u}) \equiv \frac{1}{2} \|\vec{r}(\vec{u})\|^2 = \frac{1}{2} \sum_{j=1}^{N} (r_j(\vec{u}))^2 \tag{11}
\]

Therefore, one can find roots by the “steepest descent” method: marching down the gradient of the residual norm function \( \rho \). This is a reliable but often excruciatingly slow algorithm. Modern software therefore blends the Newton and steepest descent iterations together in so-called “trust region” methods, which resemble steepest descent far from a root, but revert to Newton’s iteration very close to the solution.

Unfortunately, for wave problems, even these tools are sometimes inadequate.

4. The unexpected success of some quasi-Newton methods

Newton’s method is expensive because the Jacobian matrix must be recomputed and refactorized at every iteration. This has inspired a whole slew of “quasi-Newton” schemes that cheat on the recompute-and-factor requirement. In theory, quasi-Newton methods should have always a smaller domain of convergence and a slower rate of convergence than the “honest” Newton method.

However, Fig. 2 compares two quasi-Newton calculations of an equatorially-trapped Kelvin wave in a shear flow. The lower curve cheats by computing the Jacobian matrix just once and then recycling this matrix throughout all the remaining iterations. The upper curve is the result of recomputing the Jacobian matrix at the fifth iteration, which immediately causes the iteration to diverge!

Tuminaro et al. [16] observed the same phenomenon in the computation of nonlinear steady-state solutions to the three-dimensional Navier–Stokes equations. Their Newton correction was computed through an inner iteration, which was a standard scheme for solving a linear matrix problem. The larger the number of inner iterations, the closer their scheme to the standard Newton’s method. They found, however, that the number of Newton or quasi-Newton “outer” iterations was dramatically reduced by poisoning the accuracy of the matrix solve by using only a few inner iterations.
5. An illuminating example: the fifth-order Korteweg–deVries equation

Haupt and Boyd [10] studied the one-space-dimensional wave equation.

\[-\nu u^{(xxxx)} + u^{(xxx)} + (u - c)u^{(x)} = 0 \quad \text{[FKdV Eq.]} \tag{12}\]

where \(X \equiv x - ct\), subject to the periodic boundary condition that \(u(X) = u(X + 2\pi)\), \(c\) the phase speed of the travelling wave and \(\nu\) is a parameter independent of \(X\).

The periodic travelling waves of period \(2\pi\) consist of an infinite number of branches of solutions which can be labelled by the shortest period of the branch. The branch with period \(P(k) = 2\pi/k\) (where \(k\) must be an integer) has a limit-point (pitchfork) bifurcation from the trivial solution, \(u(X) = 0\), at infinitesimal amplitude and the phase speed \(c_{\text{limit}}(k; \nu = 0) = -k^2 - \nu k^4\) and exists only for \(c > c_{\text{limit}}(k, \nu)\).

For small amplitude,

\[u(X; 0, k) \approx a \cos(kX) + O(a^2) \tag{13}\]

Stokes’ perturbation expansions can be calculated as power series in \(a\) by expanding \(c = \sum_{k=0}^{\infty} c_k a^k\) and similarly for \(u\), and matching powers of \(a\), which yields at \(i\)th order:

\[a^i : -\nu u_{i,xxxx} + u_{i,xxx} - c_{0i} u_{i,x} = F_i(X) \tag{14}\]

This is a linear equation, just like Arnold’s example, and similarly \(u_i\) and \(F_i\) can be expanded in Fourier series as

\[u_i(X) = \sum_{k=1}^{N} u_{ik} \cos(kX), \quad F_i(X) = \sum_{k=1}^{N} f_{ik} \sin(kX) \tag{14}\]

Details are discussed in [10], but what matters is that it is trivial to compute the Fourier coefficients of \(u(X)\) at each order from those of the inhomogeneous function \(F_i(X)\) just as for Arnold’s example:

\[u_{ik} = \frac{f_{ik}}{vk^5 + k^3 + c_0k} \tag{15}\]

When the FKdV equation is discretized by a Fourier–Galerkin method, the Jacobian matrix is a dense matrix. The perturbation theory is, in numerical jargon, a quasi-Newton method: The Jacobian matrix is approximated by a diagonal matrix which is the limit of the Jacobian matrix as the amplitude \(a \to 0\). All the terms of the diagonal matrix come
from the \textit{linear} terms in the differential equation. Thus, at each order in the Stokes’ series, we must solve a \textit{linear}, \textit{forced} differential equation by means of a Fourier series—exactly as in Arnold’s exemplary problem.

Unfortunately, this also means that the Stokes’ series is also bedeviled by small divisors. The first solution branch, with $u(X) \sim a \cos(X)$ for $a \ll 1$, has a zero divisor whenever

$$
\nu = \nu_{\text{res}}(m) = \frac{1}{1 + m^2} 
$$

Physically, the zero divisor arises whenever the infinitesimal amplitude phase speeds of the $k = 1$ and $k = m$ branches coincide. Wave theorists call such a coincidence of phase speeds a “resonance”.

Haupt and Boyd [10] showed that the Stokes’ series can be salvaged by a more cunning expansion in which the resonances are built-in at lower order than in the classical expansion. Similarly, Kolmogorov, Arnold and Moser showed that the “small divisor” problem could be overcome, to obtain periodic solutions in celestial mechanics and dynamical systems theory, by using a quasi-Newton iteration that was, in comparison to Poincaré–Lindstedt series, more Newton and less quasi. The branch of mathematics that evolved from their ideas is “KAM Theory”.

6. Resonance and mode-switching

When the amplitude of the wavenumber one branch is infinitesimal, an \textit{infinite} number of other branches bifurcate from the primary branch at each of the resonances, $\nu_{\text{res}} = -1/(1 + m^2)$, for each integer $m > 1$ as shown schematically in Fig. 3. Because of the coincidence of linear phase speeds at the $m$th resonance, one can add a small but otherwise arbitrary amount of $\cos(mX)$ and still satisfy the differential equation. An exact crossing of branches is a special circumstance; the bifurcation is not stable under perturbation. When the amplitude of $\cos(X)$ is small but finite, the bifurcation breaks as shown in the lower panel of the figure. Because the branches no longer quite intersect, each branch switches identities at the “busted bifurcation” so that a branch which is predominantly wavenumber one to the left of the resonance is dominated by wavenumber $m$ to the right of the resonance.

Fig. 3. Schematic of bifurcation (upper panel) and near-bifurcation (lower panel) from the wavenumber one branch of waves in the FKdV equation.
At the wavenumber three resonance, as illustrated in Fig. 4, three modes become resonantly coupled. The mode which is dominated by wavenumber one for $\nu > -1/10$, labeled “$U$” in the lower right panel and illustrated in the lower right panel, becomes a wave dominated by wavenumber three as $\nu$ tends to large negative values as was confirmed by continuing the mode to much more negative $\nu$ than illustrated. The mode which is dominated by wavenumber one for $\nu < -1/10$, labelled “$L_{\text{top}}$” in the figure (upper right), joins with a wavenumber three branch in a limit point for $\nu$ slightly larger than $-1/10$.

Fig. 4 shows the near-bifurcation at the wavenumber three resonance. Two of the three solution branches connect in a limit point so that there are only two solution curves. At finite amplitude, these two solution curves never intersect. An equally complicated structure arises at each of the infinite number of bifurcations of the wavenumber one solution branch. At infinitesimal amplitude, there are also other primary solution branches in which the wave is $a \cos(kX)$: each of these wavenumber $k$ branches in turn has a countable infinity of bifurcations. Each of the double infinity of branches requires two integers $(k, m)$ to indicate the wavenumber $k$ of the primary wave and the wavenumber $m$ of the resonant wave. The topology of the solution branches at finite amplitude where the bifurcations are broken is thus extremely complicated.

7. Branch-jumping

“one problem with continuation methods is that, while in theory paths of roots should never cross, in practice they often come close enough to permit path jumping, unless the path is followed with impractically tight tolerances”.


Near a “busted bifurcation”, it is very easy for a numerical continuation method to fail to “turn the corner” and jump to a different branch. Fig. 5 shows that the signs of branch-jumping are rather subtle. There is a sharp change in the slope of the mesh lines at the jump, and then the slope flattens again; the third Fourier coefficient, which was rising,
Fig. 5. Modejumpsing near the wavenumber three resonance. Six steps with a spacing in $\nu$ of $1/100$ jumps from the mode labelled “U” to that labeled $L_{\text{top}}$ in the previous figure when passing from $\nu = -0.105$ to $-0.115$. Left: Mesh plot of $u(X; \nu)$. Right: The third Fourier coefficient $a_3$ is solid while the second Fourier coefficient is dashed. The amplitude of the lowest Fourier coefficient, that of $\cos(X)$, is fixed at 1.

abruptly dives to a large negative value. But that is all; the spoor of branch-jumping is as unostentatious as the crushed grass left by a tiger on the prowl.

8. Erratic failure of Newton/continuation

Table 1 shows the results of applying continuation with small steps $\delta\nu$ in the coefficient of the fifth derivative of the FKdV equation. The first guess for Newton’s iteration was obtained by extrapolating a polynomial $P_{\text{extrap}}$ fitted through previously computed solutions. The details of the table are not important; what matters is that sheer complexity and variety of the pattern of failure and success.

Fig. 6 is a graph that shows a similar erratic pattern of a failure for a different wave equation. The gaps could be filled by simply using a much smaller step size in the continuation parameter $\epsilon$, but why does the continuation wander in and out of convergence apparently without pattern?

9. Non-uniqueness and constraints

A further complication is that travelling wave equations usually do not have a unique solution unless additional constraints are imposed. For example, the KdV equation, $u_{XXX} + (u - c) u_X = 0$ has two Lie symmetries: (i) if $u(X)$ is a solution, then so is $u(X + \phi)$ for any constant $\phi$ (translational invariance) and (ii) if $u(X)$ is a solution, then so is $v(X) = -\epsilon + (1 + \epsilon) u(\sqrt{1 + \epsilon} x)$ for arbitrary $\epsilon$ (dilational symmetry). These Lie group symmetries imply that the Jacobian matrix has eigenfunctions of zero eigenvalue which are the discretizations of $u_X$ and $-1 + u(X) + (X/2)u_X$. It follows that the KdV equation does not have a unique solution unless the differential equation is augmented with additional constraints.

Sometimes these constraints can imposed implicitly. For example, as explained in Appendix C of [5], the translational invariance can be destroyed, and the peak fixed at $X = 0$, by using a Fourier basis of cosine functions only, implicitly
imposing $u(X) = u(-X)$, and the dilational invariance can be suppressed by omitting the constant from the Fourier series so as to compute travelling waves of zero mean when averaged over one spatial period. For many wave equations, this sort of “reformulation” or alternatively, deleting one condition on the residual of the differential equation, is sufficient. For many other wave equations, however, the constraints must be imposed explicitly.

Worse still, theory is not always sufficiently advanced to provide a priori knowledge of the number of constraints. A remedy is to generalize Newton’s method slightly to formally-overdetermined systems, using the QR algorithm to factor the resulting rectangular Jacobian matrix [6].

In many wave problems in fluid mechanics, one may specify an arbitrary mean flow $U(y)$, that is, an arbitrary velocity averaged over the direction of propagation, and varying in a direction perpendicular to the propagation direction. Because a smooth but otherwise arbitrary $U(y)$ can be represented analytically only as an infinite series, specifying the mean flow is equivalent to imposing an infinite number of constraints (i.e., the spectral coefficients of the series for $U(y)$).

Table 2 from [6] shows that whether the waves are cnoidal waves, shocks, solitons or weakly non-local solitary waves, there is no simple pattern to the number of constraints, or to which methods for imposing the constraints is successful.

Table 1

<table>
<thead>
<tr>
<th>$\nu_{\text{Start}}$</th>
<th>$\delta_0$</th>
<th>Degree of $P_{\text{extrap}}(v)$</th>
<th>Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-0.07$</td>
<td>$-0.01$</td>
<td>1–3</td>
<td>NaN at $v = -1/10$</td>
</tr>
<tr>
<td>$-0.07$</td>
<td>$-0.005$</td>
<td>1, 3</td>
<td>Converges to $U$ mode everywhere</td>
</tr>
<tr>
<td>$-0.07$</td>
<td>$-0.005$</td>
<td>2</td>
<td>Diverges at $v = -0.105$</td>
</tr>
<tr>
<td>$-0.07$</td>
<td>$-0.005$</td>
<td>4</td>
<td>Diverges at $v = -0.140$</td>
</tr>
<tr>
<td>$-0.07$</td>
<td>$-0.0025$</td>
<td>1–4</td>
<td>Converges to $U$ mode everywhere</td>
</tr>
<tr>
<td>$-0.065$</td>
<td>$-0.01$</td>
<td>1</td>
<td>Mode-jumps on $v \in [-0.105, -0.115]$</td>
</tr>
<tr>
<td>$-0.065$</td>
<td>$-0.01$</td>
<td>2 or 3</td>
<td>Diverges at $v = -0.105$</td>
</tr>
<tr>
<td>$-0.065$</td>
<td>$-0.005$</td>
<td>4</td>
<td>Diverges at $v = -0.115$</td>
</tr>
<tr>
<td>$-0.065$</td>
<td>$-0.005$</td>
<td>1</td>
<td>Converges to “$U$” mode everywhere</td>
</tr>
<tr>
<td>$-0.065$</td>
<td>$-0.005$</td>
<td>2</td>
<td>Diverges at $v = -0.105$</td>
</tr>
<tr>
<td>$-0.065$</td>
<td>$-0.005$</td>
<td>3</td>
<td>Converges to “$U$” mode everywhere</td>
</tr>
<tr>
<td>$-0.065$</td>
<td>$-0.005$</td>
<td>4</td>
<td>Diverges at $v = -0.12$</td>
</tr>
<tr>
<td>$-0.065$</td>
<td>$-0.005$</td>
<td>5</td>
<td>Diverges at $v = -0.115$</td>
</tr>
<tr>
<td>$-0.065$</td>
<td>$-0.0025$</td>
<td>1–4</td>
<td>Converges to “$U$” mode everywhere</td>
</tr>
</tbody>
</table>
Table 2

<table>
<thead>
<tr>
<th>Equation/Name</th>
<th>KdV</th>
<th>KdV–Burgers</th>
<th>Rossby</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wave type</td>
<td>Cnoidal</td>
<td>Shock</td>
<td>Non-local soliton</td>
</tr>
<tr>
<td>No. of zero eigenvalues</td>
<td>2</td>
<td>1</td>
<td>∞</td>
</tr>
<tr>
<td>Cheapest method, fd</td>
<td>Deleted residual</td>
<td>QR</td>
<td>Deleted residual</td>
</tr>
<tr>
<td>Cheapest, spectral</td>
<td>Reformulation</td>
<td>Deleted residual</td>
<td>–</td>
</tr>
<tr>
<td>Failures</td>
<td>Reformulation (fd only)</td>
<td>Reformulation, fd deleted residual, fd Keller-bordered</td>
<td>Reformulation</td>
</tr>
</tbody>
</table>

10. Summary

“We make an extreme, but wholly defensible, statement: There are no good, general methods for solving systems of more than one nonlinear equation. Furthermore, it is not hard to see why (very likely) there never will be any good, general methods.”

W.H. Press et al. [14], p. 379.

To this we add: Amen. The “small denominators” problem is described in the pure mathematics world of existence proofs and perturbation theory as a technical complication. In the computer science and numerical analysis literature, bifurcations are treated as a common annoyance, treated well by modern continuation/underrelaxation/trust region strategies. The reality is that when the linear dispersion relation bends back upon itself so that the frequency is a double-valued function of wavenumber, the geometry of travelling waves is broken up into an infinite number of bifurcations. Small denominators and Newton iteration failures are the difficult, barely-treatable symptoms of this topological illness.

Acknowledgements

This work was supported by the National Science Foundation through grants OCE99863683 and OCE0451951. I also thank Thiab Taha and the other organizers of the IMACS 2005 conference on “Nonlinear Evolution Equations and Wave Phenomena” and the two reviewers.

References