ELSEVIER

Comput. Methods Appl. Mech. Engrg. 175 (1999) 281–309

Computer methods in applied mechanics and engineering

www.elsevier.com/locate/cma

Compatibility conditions for time-dependent partial differential equations and the rate of convergence of Chebyshev and Fourier spectral methods

John P. Boyd*, Natasha Flyer[1]

Department of Atmospheric, Oceanic and Space Science and Laboratory for Scientific Computation, University of Michigan, 2455 Hayward Avenue, Ann Arbor, MI 48109, USA

Received 1 December 1997; revised 23 June 1998

Abstract

Compatibility conditions for partial differential equations (PDEs) are an infinite set of relations between the initial conditions, the PDE, and the boundary conditions which are necessary and sufficient for the solution to be C^∞, that is, infinitely differentiable, everywhere on the computational domain including the boundaries. Since the performance of Chebyshev spectral and spectral element methods is dramatically reduced when the solution is not C^∞, one would expect that the compatibility conditions would be a major theme in the spectral literature. Instead, it has been completely ignored. Therefore, we pursue three goals here. First, we present a proof of the compatibility conditions in a simplified form that does not require functional analysis. Second, we analyze the connection between the compatibility conditions and the rate of convergence of Chebyshev methods. Lastly, we describe strategies for slightly adjusting initial conditions so that the compatibility conditions are satisfied.

Mathematics, after all, is more than an artform.
Seki Takakazu, mathematician of the *wasan* school (d. 1708)

1. Introduction

1.1. Corner singularities and all that

The solution to an elliptic partial differential equation is weakly singular in the corners of the domain unless the forcing and boundary data are special.These 'corner' singularities are well known in science and engineering because they have observable consequences. In solid structures, corners are points of very high stress. Many homes have cracks in the drywall that radiate from the corners of doorjambs or windows; the corners of airliner windows are rounded to eliminate corner singularities since cracks in the fuselage of a transonic aircraft are undesirable. The corners of a fluid-filled region are filled with small, nested vortices called 'Moffatt eddies' which are observed both in the laboratory and in numerical models.

It is less well known that hyperbolic and parabolic partial differential equations are equally prone to singularities in the corners of the space–time domain—that is to say, at the spatial boundaries at the initial time, $t = 0$ (Fig. 1). Unless the initial conditions, boundary data and forcing satisfy a countable infinity of so-called 'compatibility' conditions, the kth spatial derivative of $u(x, t)$ will be unbounded at the spatial boundary for

* Corresponding author. E-mail: jpboyd@engin.umich.edu

[1] Present address: National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307, USA.

0045-7825/99/$ – see front matter
PII: S0045-7825(98)00358-2

some finite order k. For parabolic equations, the singularities quickly diffuse away so that the solution will become smooth and infinitely differentiable even if it was not C^∞ at the boundary at $t = 0$. For hyperbolic equations in the absence of damping, however, the weak singularities propagate away from the boundary and persist forever.

Either way, corner singularities are a serious difficulty for high order numerical methods because spectral and h–p finite element schemes implicitly assume that the solution is *smooth*, differentiable at all points including the boundary at least through derivatives of the same order as the algorithm. The usual 'infinite order' of Chebyshev and Fourier spectral schemes is only a fantasy when the solution has space–time corner singularities.

In contrast to the corner pathologies of the purely spatial domain of an elliptic equation, space–time corner singularities are usually *unphysical*. Numerical weather prediction is an obvious example. By convention, data is collected from radiosondes twice a day, simultaneously over the entire globe, to provide the initial conditions for weather forecasting. By international agreement, the twice-daily initializations are slaved to Greenwich mean time. If the initial data fed to the computer models fails to satisfy all the infinite number of compatibility conditions, then the model flow is weakly singular everywhere on the earth's surface at the initial time. Since the models are dissipative, these singularities are diffused away, and the forecast is singular *only* at the initial time, at noon and midnight in Greenwich, England. Can one really believe that the singularity structure of the atmosphere is slaved to the historical accident that a Renaissance monarch happened to own a manor on a particular river, which he donated for an astronomical observatory?

The truth is that the great engine of the atmosphere does not begin to turn at the start of a computer forecast, but rather has had five billion years for viscosity to adjust the flow to satisfy the infinite set of compatibility conditions. An incompatible initial condition is not merely a numerical difficulty but also a corruption of the physics.

Numerical forecasters have known for decades that the classical initial value problem was unsatisfactory. On a global scale, the winds and pressure are correlated so as to be what meteorologists call 'balanced'. However, radiosondes and satellites are prey both to measurement errors and errors in interpolating the measurements to the computational grid. These errors are random and unbalanced. If raw data is used as the initial conditions for a forecast, the computed winds and pressure will oscillate with large-amplitude, spurious gravity waves, a kind of computational St. Vitus Dance [9,28]. Numerical forecasts are therefore begun with a preprocessing step in which the data is balanced through an 'initialization' procedure, which is quite effective in suppressing these spurious oscillations. (However, present-day initialization algorithms do *not* adjust the measurements so that the initial state satisfies the compatibility conditions at the earth's surface.)

It may well be that many engineering and physics simulations would benefit from a compatibility-condition-enforcing 'initialization'. In the concluding sections, we will discuss a number of strategies for mitigating the ill effects of space–time corner singularities.

First, though, we shall discuss a rather broad theorem, embracing both parabolic and hyperbolic equations in an arbitrary number of spatial dimensions, which will illustrate the ubiquity of compatibility conditions. To provide concrete examples, we follow this with a discussion of three specific linear partial differential equations in one space dimension. To illustrate the numerical consequences of initial conditions that flunk the compatibility conditions, we shall use only Chebyshev polynomial series. However, singularities are just as bad for Legendre series, spectral p-type finite elements, and other high order numerical methods.

The method of separation of variables has acquired a bad reputation because its eigenseries converge slowly. However, for many problems, we show that Chebyshev series are no faster because of singularities in the space–time corners.

To a mathematician, an initial value problem lives only in the present and onward, as if the universe were created in a thunderclap at $t = 0$. In reality, most physical systems have a history which wraps all physical initial conditions —all initial conditions that generate C^∞ solutions—in an infinite set of boundary constraints.

2. An exemplary theorem: compatibility conditions for the multidimensional wave and diffusion equations

Since every time-and-space-dependent partial differential equation probably has an infinite set of compatibility conditions, different for each PDE, it is impossible to cover all cases. The following theorem, however, is

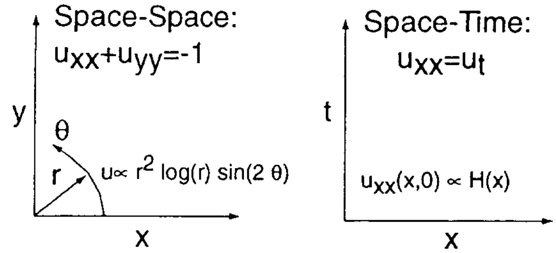

Fig. 1. Left: Elliptic partial differential equations, even if linear and constant coefficient with constant forcing, are singular at the corners of spatial domains. With homogeneous Dirichlet conditions, the Poisson equation on the square is singular at the corner like $r^2 \log(r) \sin(2\theta)$ in terms of a polar coordinate system centered on the corner as shown. Right: A parabolic or hyperbolic equation, even if linear and constant coefficient, is singular at the spatial boundary at $t = 0$; the second derivative of the solution to the diffusion equation or wave equation is usually discontinuous at the space–time corner, proportional to the step function $H(x)$.

representative. We will restrict it to homogeneous equations with homogeneous boundary conditions. However, note that a linear PDE with inhomogeneous boundary conditions can always be transformed into a problem with homogeneous boundary conditions by writing $u = w + I$ where $I(x, t)$ is any function that satisfies the inhomogeneous boundary conditions and w is a new unknown that satisfies the same PDE with a different inhomogeneous term and homogeneous boundary conditions. Similarly, an inhomogeneous PDE such as $u_t = Lu - f$ can be rendered homogeneous by writing $u = w + E$ where E solves the elliptic problem $LE = f$ where time appears as a parameter.

THEOREM 1 (Compatibility: diffusion and wave equations). The domain is $[0, T] \times \Omega$ *where* Ω *is a d-dimensional spatial domain with a boundary* $\partial\Omega$ *which is a* C^∞ *manifold of dimension* $(d-1)$. *Let L be an elliptic operator of the form*

$$L = \sum_{i=1}^{d} \sum_{j=1}^{d} A_{ij}(x) \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} + \sum_{j=1}^{d} B_j(x) \frac{\partial}{\partial x_j} + \sum_{j=1}^{d} C_j(x) \tag{1}$$

The generalized linear diffusion problem is

$$u_t = Lu\,, \qquad u(x, t=0) = u_0(x) \in \Omega\,, \tag{2}$$

The generalized wave equation is

$$u_{tt} = Lu\,, \qquad u(x, t=0) = u_0(x)\,, \qquad u_t(x, t=0) = v_0(x) \in \Omega\,, \tag{3}$$

Both are subject to homogeneous Dirichlet or Neumann boundary conditions

$$u = 0 \text{ [Dirichlet] } \quad or \quad \nabla u \cdot \vec{n} = 0 \text{ [Neumann] } \quad \in \partial\Omega\, \forall\, t \tag{4}$$

where $u_0(x)$ *is assumed to be* C^∞ *on all of the boundary.*

Then, the necessary and sufficient compatibility conditions for $u(x, t)$ *to be* C^{2k} *is that*

(i) Dirichlet boundary conditions:

$$L^k u_0 = 0 \quad \text{and} \quad L^k v_0 = 0 \quad \forall\, x \in \partial\Omega\,, \quad k = 0, 1, \ldots \tag{5}$$

(ii) Neumann boundary conditions:

$$L^k\{\nabla u_0 \cdot \vec{n}\} = 0 \quad \text{and} \quad L^k\{\nabla v_0 \cdot \vec{n}\} = 0 \quad on\ [0, T] \times \partial\Omega \tag{6}$$

where in each case the second condition applies only to the wave equation, which has the extra initial condition.

In words, the compatibility conditions for the wave and diffusion equations are IDENTICAL in the sense that the same constraints must be imposed on both initial conditions for the wave equation as on the single initial condition for the diffusion equation.

PROOF. For simplicity, we shall give the argument only for homogeneous Dirichlet boundary conditions and only prove that these conditions are necessary. The complete proof, for more general boundary conditions and inhomogeneous equations and including proof of sufficiency, is given by [26,27,13,35] [parabolic PDEs] and [25,32,33,35] [hyperbolic].

If the boundary condition is independent of time as specified, then the time derivatives of the solution to all orders must be zero on the boundary:

$$\frac{\partial^j u}{\partial t^j}(t=0) = 0 \quad \in [0, T] \times \partial\Omega . \tag{7}$$

Since the left-hand side of the diffusion equation, which is just u_t, is zero on the boundary, it follows that the right-hand side must vanish on $\partial\Omega$, too:

$$Lu = 0 \quad \in [0, T] \times \partial\Omega . \tag{8}$$

It follows that if

$$Lu_0 \neq 0 \quad \in \partial\Omega , \tag{9}$$

then there will be a contradiction: the boundary condition together with the partial differential equation itself demands that $Lu = 0$ at the boundary, but the application of L to the initial condition u_0 does not vanish. The resolution of the contradiction is that if the initial condition does not satisfy the lowest order compatibility condition $Lu_0 = 0$, then Lu has a jump discontinuity at the boundary at $t = 0$; the magnitude of the discontinuity is simply the value of Lu_0 on $\partial\Omega$.

To obtain the higher order compatibility conditions, assume that the conditions hold for all $j < k$, that is

$$L^j u = 0 \quad \in [0, T] \times \partial\Omega \quad j = 1, 2, \ldots, k-1 . \tag{10}$$

Apply the operator L to the diffusion equation $(k-1)$ times, which gives, after interchanging the time and space derivatives,

$$(L^{k-1} u)_t = L^k u \tag{11}$$

However, $L^{k-1} u = 0$ on $\partial\Omega$ for all t, so its time derivative is zero. It follows that

$$L^k u = 0 \quad \in [0, T] \times \partial\Omega \tag{12}$$

Once again, if L^k applied to the initial condition does not vanish on the boundaries, then we have a contradiction which forces a jump discontinuity in $L^k u$ at the boundary. Thus, for the solution to have $2k$ continuous derivatives including the boundary, it is necessary that the initial condition satisfy the compatibility conditions (5): $L^k u_0 = 0$ on the boundary $\partial\Omega$.

For the wave equation, the proof is identical except that we differentiate each side of Eq. (3) with respect to t to get

$$Lu_t = u_{ttt} \tag{13}$$

Making the substitution $v = u_t$, the rest of the proof proceeds as above.

3. The one-dimensional diffusion equation

3.1. Compatibility conditions with forcing and time-dependent boundary conditions

The ideas of the general but abstract theorem of the preceding section can be fleshed out by looking at a set of simple examples. In this section, we analyze the one-dimensional diffusion equation. In contrast to the theorem (and the remaining subsections), in this section we shall allow both a general, time-dependent inhomogeneous term and also general, time-dependent boundary conditions:

$$u_t = u_{xx} + f(x, t) , \qquad u(x, 0) = Q(x) \quad x \in [0, \pi] \tag{14}$$

subject to the time-dependent Dirichlet boundary conditions

$$u(0,t)=\alpha(t)\,,\qquad u(\pi,t)=\beta(t) \tag{15}$$

The initial-value problem is well-posed for any smooth $Q(x)$. However, if $u(x,t)$ is to be an analytic function for all x on the domain and all t, then the lowest order compatibility condition is that the initial condition $Q(x)$ must satisfy the boundary conditions:

$$Q(0)=\alpha(0)\,,\qquad Q(\pi)=\beta(0) \tag{16}$$

If this condition is not satisfied, then u is forced to jump discontinuously from $Q(0)$ to the boundary condition $\alpha(0)$ at $x=0$, and similarly at the other boundary. For finite t, this discontinuity is diffused into a layer of finite thickness δ where δ increases steadily with time. For sufficiently small time, however, any numerical method with a finite number N of degrees of freedom will be unable to resolve the layer where u changes from $Q(0)$ to $\alpha(0)$.

The need to avoid such jump discontinuities at the boundaries (by choosing the initial condition to have the same boundary values as the boundary conditions) is so obvious that it is usually imposed, even by very nuts-and-bolts engineers who have never heard the phrase 'compatibility condition'. However, the single pair of equations (16) is insufficient to ensure a solution which is analytic at $t=0$.

If we evaluate the diffusion equation at the left boundary and then apply the boundary condition, we obtain:

$$u_{xx}(0,t)=\alpha_t(t)-f(0,t) \tag{17}$$

(and similarly at the right boundary). Unless the initial condition is chosen so that

$$Q_{xx}(0)=\alpha_t(0)-f(0,0)\,,\qquad Q_{xx}(\pi)=\beta_t(0)-f(\pi,0) \tag{18}$$

the *second derivative* of $u(x,0)$ will have a jump discontinuity at the spatial boundaries.

Similarly, if the diffusion equation is differentiated twice with respect to x and evaluated at the left boundary, we find

$$u_{4x}(0,t)=u_{txx}(0,t)-f_{xx}(0,t) \tag{19}$$

However, Eq. (17) already provides a value for u_{xx} at the left boundary; time differentiation and then evaluation at $t=0$ shows that the fourth spatial derivative of $u(x,t)$ will be discontinuous at the boundaries at $t=0$ unless

$$\begin{aligned}Q_{4x}(0,t)&=\alpha_{tt}(0)-f_t(0,0)-f_{xx}(0,t)\\ Q_{4x}(\pi,t)&=\beta_{tt}(0)-f_t(\pi,0)-f_{xx}(\pi,t)\end{aligned} \tag{20}$$

This argument can be extended indefinitely: to be analytic in the space-time corners, the initial condition must satisfy a countable infinity of compatibility conditions.

The inhomogeneous forcing and boundary conditions add complication to the formulas but do not alter the essentials—the reader will note the great similarity between the argument in this subsection and the proof in the preceding section. We shall therefore simplify the formulas by setting $f=\alpha=\beta=0$ in the rest of this section. For this special case of homogeneous equation and boundary conditions, the compatibility conditions are easy to summarize: *all even derivatives* of the initial condition must be *zero* at *both* endpoints.

3.2. Formal solutions

To illustrate the practical consequences of these compatibility equations, we shall analytically solve the diffusion equation (with homogeneous boundary conditions and forcing):

$$u_t=u_{xx}\,,\qquad u(x,0)=Q(x) \tag{21}$$

Although our goal is to solve the problem on the finite interval $x\in[0,\pi]$ subject to homogeneous Dirichlet boundary conditions, it turns out the best way to do this is a two-step procedure which requires (i) solving the problem on an *infinite* interval followed by (ii) constructing a solution which is *spatially periodic.*

Infinite interval solution, denoted by $P(x,t)$ [40]:

$$P(x,t) = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} Q(y) \exp\left(-\frac{(y-x)^2}{4t}\right) dy \tag{22}$$

By the 'method of images', we can combine an infinite number of copies of $P(x,t)$ to obtain a solution which is periodic with 2π. First, split the infinite interval solution into its parts which are symmetric and antisymmetric with respect to $x = \pi/2$:

$$\begin{aligned} P_S(x,t) &= (1/2)\{P(x,t) + P(-x+\pi, t)\}, \\ P_A(x,t) &= (1/2)\{P(x,t) - P(-x+\pi, t)\} \end{aligned} \tag{23}$$

Note that P_S and P_A are each solutions of the diffusion equation, and could have been obtained by splitting the initial condition into its symmetric and antisymmetric parts before applying the infinite interval formula. Then, a spatially periodic solution to the diffusion equation is

$$u_P(x,t) = \sum_{m=-\infty}^{\infty} (-1)^m P_S(x - m\pi, t) + \sum_{m=-\infty}^{\infty} P_A(x - m\pi, t) \tag{24}$$

In modern language, the first sum in u_P is an 'alternating imbricate series' with $P_S(x,t)$ as the 'pattern function' while similarly the second sum is a regular 'imbricate' series with P_A as the pattern function [5,6]. The solution is periodic with period 2π since if we increase the argument x by 2π in either sum, we merely shift from one copy of the pattern function to another identical copy with the same sign.

The third step is to prove a theorem (Appendix A) that both series are individually antisymmetric with respect to both $x = 0$ and $x = \pi$, that is,

$$u_P(x,t) = -u_P(-x,t), \qquad u_P(x,t) = -u_P(-x+2\pi, t) \tag{25}$$

Antisymmetry with respect to a point implies that the antisymmetric function is zero at that point. Thus, the periodic solution u_P satisfies homogeneous Dirichlet boundary conditions at both $x = 0$ and $x = \pi$ and therefore is the desired solution to the Dirichlet problem on $x \in [0, \pi]$.

3.3. Example

The simplest initial condition that: (i) satisfies the zeroth compatibility condition of vanishing at both endpoints; and (ii) fails to satisfy the first-order condition that its second derivative is zero at $x = 0, \pi$ is the parabola:

$$Q(x) = \begin{cases} x(\pi - x), & x \in [0, \pi] \\ 0, & \text{elsewhere} \end{cases} \tag{26}$$

The infinite interval solution is

$$\begin{aligned} P(x,t) = {} & \left(-t + \frac{x\pi}{2} - \frac{x^2}{2}\right) \operatorname{erf}\left(\frac{x}{2\sqrt{(t)}}\right) - \left(-t + \frac{x\pi}{2} - \frac{x^2}{2}\right) \operatorname{erf}\left(\frac{x-\pi}{2\sqrt{(t)}}\right) \\ & + (\pi - x)\sqrt{\frac{t}{\pi}} \exp\left(-\frac{x^2}{4t}\right) + x\sqrt{\frac{t}{\pi}} \exp\left(-\frac{x^2}{4t}\right). \end{aligned} \tag{27}$$

This pattern function is symmetric with respect to $x = \pi/2$, so using it as the pattern function for an 'alternating' imbricate series generates a solution which both is spatially periodic with period 2π and also satisfies homogeneous Dirichlet boundary conditions at $x = 0, \pi$:

$$u(x,t) = \sum_{m=-\infty}^{\infty} (-1)^m P(x - m\pi, t) \tag{28}$$

At $t = 0$, the pattern function is a piecewise parabola, different from zero only on the interval $x \in [0, \pi]$. At the ends of this interval, the second derivative jumps discontinously from -2 to zero. For $t > 0$, the diffusion smooths the pattern function so that it (and its alternating imbricate sum) are analytic functions for all real x. However, for sufficiently small t, the discontinuity in the second derivative is smoothed over a very thin layer.

Fig. 2 shows how this solution evolves. The graph of u itself changes almost invisibly from the initial time

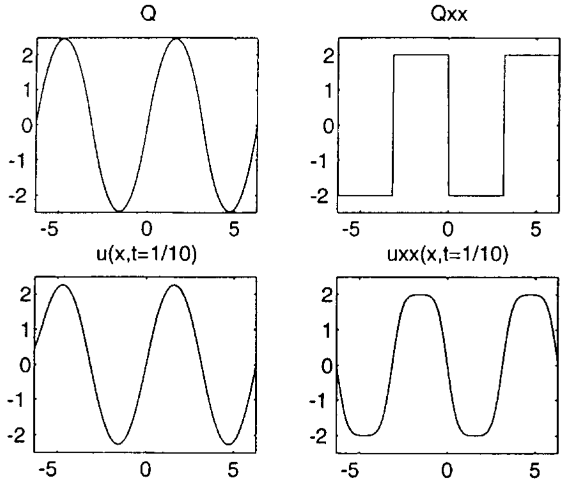

Fig. 2. Upper panels: Initial condition (left) and its second derivative (right). Bottom panels: the solution and its second derivative at $t = 1/10$.

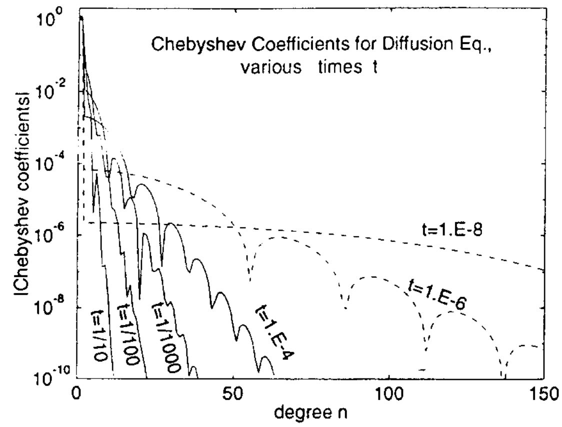

Fig. 3. Chebyshev coefficients for $u(x, t)$ at various times t where u is the solution to the diffusion equation from an initial condition which is a quadratic polynomial.

(upper left) to the later time (lower left). However, the jump discontinuities of the second derivative (upper right) are visibly smoothed by diffusion to give the graph of u_{xx} in the lower right. Although the primary interest in a parabolic initial condition is when the physical domain is restricted to $x \in [0, \pi]$, the figure shows the periodic solution on a larger domain in x. This extension to a larger interval makes it much easier to see the discontinuous second derivative of the initial condition and its subsequent smoothing.

3.4. The rate of Chebyshev series convergence

In a Chebyshev spectral method, $u(x, t)$ is approximated by a truncated series of Chebyshev polynomials with coefficients a_n. (The basis functions are actually $T_j\{[2/\pi](x - \pi/2)\}$.) Fig. 3 shows that the rate of convergence of the coefficients is very fast even for moderate times ($t = 1/10$, leftmost curve) but becomes slower and slower for smaller times.

The convergence-slowing boundary layer diffuses away quickly. For very small t, the rate of convergence is awful, but the slowly decreasing coefficients are very small ($O(10^{-5})$ for $t = 10^{-8}$). In many practical situations, a Chebyshev approximation with a moderate number of degrees of freedom would be quite satisfactory.

Still, it *is* alarming that an initial condition as smooth as a parabola for a differential equation as simple as the unforced, one-dimensional diffusion equation can reduce a spectral method from infinite order convergence to wretchedly slow convergence.

4. The one-dimensional wave equation

4.1. Formal solutions

The one-dimensional wave equation is

$$u_{tt} = u_{xx} \tag{29}$$

subject to the initial conditions

$$u(x, 0) = R(x)\,, \qquad u_t(x, 0) = 0 \tag{30}$$

For simplicity, we have set the inhomogeneous term in the differential equation and also the second initial condition equal to zero and will similarly apply only homogeneous boundary conditions. These restrictions are irrelevant to the underlying ideas, and there is an infinite set of compatibility conditions even when these restrictions are all relaxed.

Although wave propagation is very different from diffusion, constructing a solution is very similar. D'Alembert solved the wave equation on a domain without walls:

Infinite-interval solution:

$$P(x,t) = \frac{1}{2}\{R(x+t) + R(x-t)\} \tag{31}$$

If $R(x)$ is a function that decays as fast as $O(1/|x|^2)$, then this solution may be 'imbricated' to construct a spatially periodic solution. For simplicity, we assume that $R(x)$ is symmetric with respect to $x = \pi/2$; if it is not, then the antisymmetric part should be summed in a non-alternating series as done above for the diffusion equation. For such a symmetric initial condition, an alternating imbricate series generates:

Spatially-periodic solution:

$$u_P(x,t) \equiv \sum_{m=-\infty}^{\infty} (-1)^m \frac{1}{2}\{R(x+t-m\pi) + R(x-t-m\pi)\} \tag{32}$$

The third step is to invoke the imbricate theorem of Appendix A to show that this series is zero at both $x = 0$ and $x = \pi$ and thus satisfies homogeneous Dirichlet boundary conditions

$$u_P(0,t) = u_P(\pi,t) = 0 \tag{33}$$

To satisfy homogeneous Neumann conditions, the procedure is similar except that the part of $R(x)$ which is symmetric with respect to $x = \pi/2$ is converted into a regular imbricate series

$$u_{\text{Neumann}}(x,t) \equiv \sum_{m=-\infty}^{\infty} \frac{1}{2}\{R(x+t-m\pi) + R(x-t-m\pi)\} \tag{34}$$

(Similarly, the antisymmetric part of R is fashioned into an alternating imbricate series and added to the sum above.) One can prove by the methods of Appendix A that

$$\frac{\partial u_{\text{Neumann}}}{\partial x}(0,t) = \frac{\partial u_{\text{Neumann}}}{\partial x}(\pi,t) = 0 \tag{35}$$

4.2. Compatibility conditions

Homogeneous Dirichlet boundary conditions:

$$R_{2kx}(0,0) = R_{2kx}(\pi,0) = 0\,, \quad k = 0,1,2,\ldots \tag{36}$$

where R_{2kx} denotes the x-derivative of order $2k$. Similarly, if $u_t(x,0) = S(x)$, all its even derivatives must also vanish at both $x = 0$ and $x = \pi$. (We set $S(x) = 0$ for simplicity elsewhere in this section).

Homogeneous Neumann boundary conditions:

$$R_{2k+1,x}(0,0) = R_{2k+1,x}(\pi,0) = 0\,, \quad k = 0,1,2,\ldots \tag{37}$$

(For nonzero initial condition on the first time derivative, all odd derivatives of the initial condition $S(x)$ must similarly vanish at both endpoints for Neumann boundary conditions.)

These can be easily generalized. For example, if the boundary conditions are altered to $u(0,t) = \alpha(t)$, $u(\pi,t) = \beta(t)$ and the differential equation is generalized to $u_{tt} = u_{xx} - f(x,t)$, then the first-order compatibility conditions are

$$u_{xx}(0,0) = \alpha_{tt}(0) + f(0,0)\,, \qquad u_{xx}(\pi,t) = \beta_{tt}(0) + f(\pi,0) \tag{38}$$

We shall restrict ourselves to $f = \alpha = \beta = 0$ in the rest of the section.

4.3. Examples

The simplest initial condition that satisfies homogeneous boundary conditions but flunks the compatibility conditions on the second derivative is

$$R(x) \equiv \begin{cases} x(\pi - x), & x \in [0, \pi] \\ 0, & \text{otherwise} \end{cases} \tag{39}$$

This is identical with the initial condition we employed for the diffusion equation.

Fig. 4 shows that the time evolution is very different when the initial parabola divides into two and propagates instead of diffuses. The most important difference is that the discontinuity in the second derivative (right pair of graphs) is not healed by time, but merely propagates to the left or to the right. For all times, the second derivative has a jump of magnitude 2.

4.4. Chebyshev convergence for the wave equation

The exemplary solution illustrated in Fig. 4 has four jumps of magnitude two. For simplicity, we shall discuss effects of second derivative discontinuities by using a simpler function which has two jumps of unit magnitude at $x = \pm(1 - \epsilon)$ where $\epsilon > 0$:

$$J(x) \equiv \begin{cases} \dfrac{\{x^2 - (1-\epsilon)^2\}^2}{8(1-\epsilon)^2}, & |x| < 1 - \epsilon \\ 0, & \text{otherwise} \end{cases} \tag{40}$$

Since the spectral coefficients of any function are asymptotically determined by its worst singularities, it follows that the coefficients for our $u(x, t)$ or any function with jumps in the second derivative can be approximated by a weighted sum of those of $J(x; \epsilon)$ for appropriate values of ϵ. If

$$J(x; \epsilon) = \sum_{j=0}^{\infty} a_{2j} T_{2j}(x) \tag{41}$$

then the Chebyshev coefficients are given exactly by

$$a_{2n} = \frac{\chi_0 + \chi_1 n + \chi_2 n^2}{8\pi \cos^2(\tau)(n^5 - 5n^3 + 4n)} \tag{42}$$

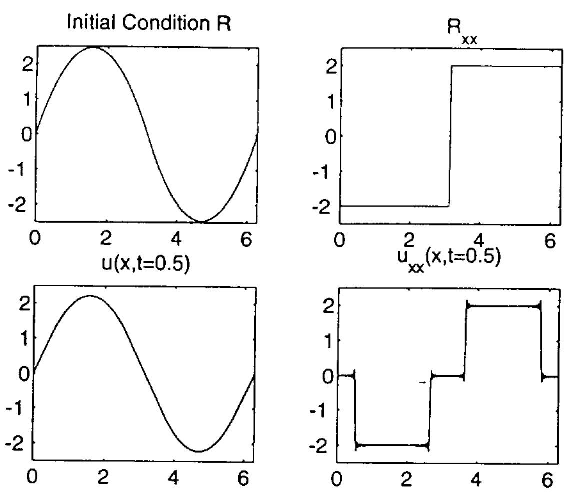

Fig. 4. Upper panels: Initial condition (left) and its second derivative (right). Bottom panels: the solution and its second derivative at $t = 0.5$ (right).

where the auxiliary parameter τ is

$$\begin{aligned} \tau &\equiv \arccos(1-\epsilon) \\ &\approx \sqrt{2}\sqrt{\epsilon} \qquad \epsilon \ll 1 \end{aligned} \tag{43}$$

and

$$\chi_0 = -\sin(2n\tau)(2+\cos(4\tau)), \qquad \chi_1 = (3/2)\cos(2n\tau)\sin(4\tau)$$
$$\chi_2 = (1/2)\sin(2n\tau)(1-\cos(4\tau)) \tag{44}$$

In the asymptotic limit,

$$a_{2n} \sim \frac{\sin^2(\tau)\sin(2n\tau)}{2\pi n^3}, \quad n \to \infty \tag{45}$$

Thus, the Chebyshev series of a function with a second derivative discontinuity has 'third-order' convergence, that is, the coefficient of T_n is decreasing as $O(n^3)$. If $u(x, t)$ or $J(x)$ were analytic in x everywhere on $x \in [0, \pi]$, then the Chebyshev or Legendre coefficients will decrease as $O(\exp(-qn))$ for some constant q; this is the usual 'geometric' or 'spectral' convergence. The proportionality constant in a_{2n} for $J(x)$ becomes smaller and smaller as the jumps move closer and closer to the endpoints (ϵ, $\tau(\epsilon) \to 0$), but the spectral convergence is destroyed for all $\epsilon < 1$.

Spectral elements, also called 'p-type finite elements', lack infinite order convergence. However, a calculation with sixth degree polynomials, which would have an error $O(h^6)$ where h is the size of an element, would be reduced to only third-order convergence for the wave equation. Only three-point differences or piecewise linear finite elements, which have only second-order accuracy anyway, are indifferent to the first-order compatibility conditions.

5. The axisymmetric wave equation

5.1. Dirichlet boundary conditions

In order to show compatibility conditions in complicated geometries, we will consider the wave equation in radial coordinates with Dirichlet boundary conditions. The initial condition will assume radial symmetry, making the solution independent of angle. The problem is given by

$$u_{tt} = u_{rr} + \frac{1}{r} u_r \quad t \geq 0 \quad r \in [0, 1] \tag{46}$$

$$u(1, t) = 0 \tag{47}$$

$$u(0, t) = \text{bounded} \tag{48}$$

$$u(r, 0) = u_0(r) \tag{49}$$

$$u_t(r, 0) = v_0(r) \tag{50}$$

The compatibility conditions are

$$L^k u_0|_{r=1} = 0 \tag{51}$$

$$L^k v_0|_{r=1} = 0 \tag{52}$$

for $k \geq 1$ where

$$L = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}. \tag{53}$$

The simplest solution that flunks the first-order compatibility condition is

$$u_0(r) = 1 - r^2, \qquad v_0(r) = 0. \tag{54}$$

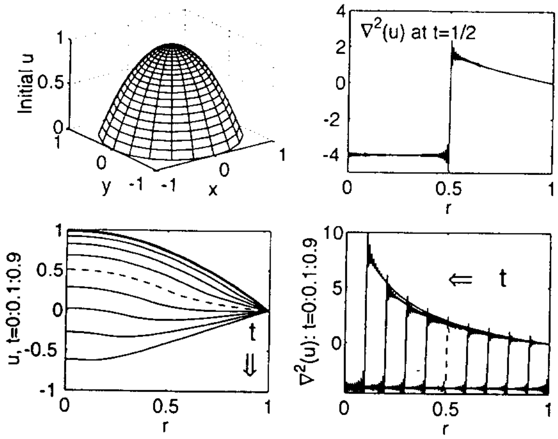

Fig. 5. Upper left: initial condition for exemplary solution of the axisymmetric wave equation. Upper right: $\nabla^2 u$ at $t = 1/2$. Lower left: sequence of $u(r,t)$ at intervals of 1/10 from $t = 0$ (top curve) to $t = 9/10$ (bottom). Lower right: same as lower left except for the $\nabla^2 u$.

Instead of vanishing at $r = 1$, the Laplacian of this initial condition is

$$\frac{\partial^2 u_0}{\partial r^2} + \frac{1}{r}\frac{\partial u_0}{\partial r} = -4 \tag{55}$$

for all r including the endpoint.

The time evolution is almost exactly the same as for the wave equation in one space dimension. At $t = 0$, there is a discontinuous jump in the Laplacian of the initial condition at the boundary from $Lu = -4$ to $Lu = 0$. The wave equation then propagates the discontinuity into the interior of the domain as illustrated by the two right panels of Fig. 5. The fuzzy black at the top and bottom of each jump is because of Gibbs' phenomenon, which creates oscillations near the jump in the truncated Bessel series used to make the graphs. As the jump converges on the origin, the amplitude rises since the same energy is concentrated on a smaller circle (lower right graph); at $t = 1$, the jump has converged to a point, and becomes infinite. Aside from this amplification-by-focusing, the evolution of the solution to the axisymmetric wave equation differs but little from that of $u_{tt} = u_{xx}$.

Homogeneous Neumann boundary conditions are very similar except that the compatibility conditions are that L^k acting on the radial derivatives of the two initial conditions, rather than the initial conditions themselves, must vanish at $r = 1$.

The solution to this problem as given by the method of separation of variables is described in Appendix B. The terms decrease asymptotically as $O(1/n^3)$, which seems to support Gottlieb and Orszag's claim ([15, pp. 33–35]), that expanding a function of radius as a Bessel series is a bad idea. However, because of the propagating second derivative discontinuity, the Chebyshev and Legendre series converge *just as slowly.*

A revisionist analysis of the method of separation of variables

In the analytic scheme for solving linear partial differential equations known as the method of 'separation of variables', the diffusion equation $u_t = Lu$ is solved by expanding in terms of the eigenfunctions of the elliptic operator L:

$$u = \sum a_n(t)\phi_n \tag{56}$$

$$L\phi_n = \lambda_n \phi_n \tag{57}$$

(For present purposes, it is not necessary to specify that the precise form of the elliptic operator L except that it is linear and all its eigenvalues are negative.) Substituting the eigenseries into the partial differential equation and applying the eigenrelation (57) shows the equations for each spectral coefficient become mutually independent, allowing the exact solution

$$a_n(t) = a_n(0)\exp(\lambda_n t) \tag{58}$$

where all the eigenvalues are negative. For the generalized wave equation of Section 2, $u_{tt} = Lu$, the theory is almost identical except that the spectral coefficients are $a_n(t) = a_n(0)\cos(\sqrt{-\lambda_n}t) + a_{t,n}(0)\sin(\sqrt{-\lambda_n}t)/\sqrt{-\lambda_n}$. If the wave or diffusion solution is required to satisfy homogeneous boundary conditions, then these same conditions are imposed on the eigenfunctions so that each term in the eigenseries individually satisfies the boundary conditions.

The only rub is that the eigenseries usually converges slowly. For example, the eigenseries for the wave equation, evolving from an initial parabola, is

$$u(x,t) = \frac{4}{\pi}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^3}\{\sin((2n+1)(x-t)) + \sin((2n+1)(x+t))\} \tag{59}$$

The Fourier sine coefficients decrease only as fast as $O(1/n^3)$. This has suggested to many that to obtain high accuracy, one should junk separation-of-variables as an eighteenth-century relic, and apply Chebyshev series or high order finite elements instead.

However, we have already seen that the Chebyshev series for this same solution also has only third order convergence because of the jump discontinuity in the second derivative. In some sense, the separation-of-variables series converges as fast as possible for a solution which has shipwrecked on Compatibility Condition Reef. The eigenfunctions themselves *individually* satisfy the *entire, infinite set* of compatibility conditions. This is easily proven through an inductive argument by applying the elliptic operator L to both sides of the eigenrelation (57) and invoking the boundary conditions that ϕ_n must satisfy.

This implies that any *finite* sum of eigenfunctions must satisfy the compatibility conditions, too. However, there is a subtlety. If the eigenseries is used to approximate a solution which flunks the kth compatibility conditions while passing those of lower order, then $L^k(u)$ for the *infinite* sum of the series will jump discontinuously at the boundaries. The series for $L^k u$ will converge very slowly with coefficients decreasing as $O(1/n)$ while that for u will have coefficients decreasing as $O(1/[n\lambda_n^k])$. Since $\lambda_n \sim O(n^2)$[typically], violation of the kth compatibility conditions usually implies that the eigencoefficients of u are decreasing as $O(1/n^{2k+1})$. For the one-dimensional wave equation, $L \equiv d^2/dx^2$; the parabolic initial condition flunks the condition, imposed by Dirichlet boundary conditions, that $Lu(0) = u_{xx}(0) = 0$ and so the Fourier coefficients decrease as $O(1/n^3)$ as expected.

Thus, the property that the eigenfunctions are individually compatible to all orders cannot rescue an initial condition which is incompatible; the eigenseries enforces compatibility at the boundaries by introducing discontinuities at the boundaries. The series approximation to $L^k u$ will exhibit Gibbs' phenomenon.

On the other hand, if the eigenseries converges with 'spectral' accuracy, that is, faster than $O(n^{-k})$ for arbitrarily large k, then this implies spectral convergence for all the derivatives of the series, too. Gibbs' phenomenon is not possible either for u or for any of its derivatives. The series is converging to a function which is truly and continuously compatible with all the boundary conditions. We summarize as follows.

THEOREM 2 (Compatibility and eigenseries). For a partial differential equation involving the elliptic operator L, an initial condition $u(x,0)$ is compatible to all orders if any of the following conditions is true:

(i) The eigenfunction series of the initial conditions converges exponentially fast. This means that if the initial condition is expanded as a sum of the eigenfunctions of L, the coefficients $a_n(0)$ of this series can be bounded by

$$|a_n(0)| \le C\exp(-qn) \tag{60}$$

for some constants C, q and n (and normalized eigenfunctions)

(ii) $u(x,0)$ is proportional to a single eigenfunction of L or

(iii) $u(x,0)$ is a sum of a finite number of eigenfunctions.

(The second and third conditions are special cases of the first condition.)

The initial condition is NOT compatible if the series converges algebraically with n, that is, if the tightest bound that can be imposed is

$$|a_n(0)| \le \frac{C}{n^k} \tag{61}$$

for some constants C and k where k is finite. (The sum of an algebraically-converging series has discontinuous derivatives of some finite order.)

On the other hand, if the initial condition flunks the compatibility conditions for a homogeneous partial differential equation with homogeneous boundary conditions, then the eigenseries converges no more slowly that a Chebyshev or Legendre method—and the coefficients, being the solutions of *uncoupled* ordinary differential equations in time, are far easier to compute.

However, eigenseries are less appealing if the boundary conditions are *inhomogeneous* because the eigenfunctions themselves satisfy homogeneous boundary conditions. It follows that the eigenseries will converge slowly and exhibit the Gibbs' phenomenon even if the initial condition satisfies all compatibility conditions.

There is a remedy: Homogenize the boundary conditions. Let $I(x, t)$ denote a function that interpolates the boundary conditions but is otherwise arbitrary. The new unknown

$$w \equiv u - I \tag{62}$$

will solve

$$w_t = Lw - \{I_t - LI\} \tag{63}$$

subject to homogeneous boundary conditions. This 'homogenization of the boundary conditions' is fairly straightforward even in multiple space dimensions (through transfinite interpolation) and even for nonlinear differential equations. After it, eigenseries again converge as fast as Chebyshev series.

If the differential equation is inhomogeneous, then there are a couple of possibilities. One is that the forcing $f(x, t)$ has a rapidly convergent eigenseries. In this case, the inhomogeneous term does not degrade the rate of convergence of the eigenexpansion of $u(x, t)$.

The other possibility is that the eigenseries of $f(x, t)$ converges poorly. For example, if $f(x, t) = \exp(x)$ is added to the diffusion equation, then the Fourier series for u will have only an algebraic rate of convergence for all t, even if the initial condition satisfies all compatibility conditions, because the sine coefficients of the non-periodic function $\exp(x)$ converge only as $O(1/n)$. A Chebyshev series, in contrast, will converge geometrically fast for large times after the initial boundary singularities, if any, have diffused into harmlessness.

It is generally possible to homogenize a differential equation, too. For example, define $E(x, t)$ to be the solution to

$$LE = -f \tag{64}$$

Then, the new unknown

$$w \equiv u - E \tag{65}$$

will solve a homogeneous partial differential equation. The difficulty with this 'homogenization of the PDE' is that it requires solving a boundary value problem for E—one boundary value problem at each time step if f varies with time. This is the same labor as solving the time dependent partial differential equation by an implicit or semi-implicit time-marching algorithm. Furthermore, if f is non-periodic or otherwise has a slowly convergent eigenseries, we must use a Chebyshev basis to get rapid convergence in the computation of E.

Thus, in summary, the method of separation of variables is just as efficient and gives series that converge just as fast as any alternative spectral method when both the differential equation and the boundary conditions are homogeneous. For inhomogeneous problems, it is still possible to use eigenseries, but only with additional work, perhaps considerable extra work.

7. Compatibility conditions for more general spatial operators and nonlinear systems

7.1. Hyperbolic and elliptic equations

The mathematical literature of compatibility conditions has a history that extends back at least half a century and greatly generalizes the simple illustrations of previous sections. In the late 1930s and early 1940s, Bourgin and Duffin [1] and John [24] investigated the continuity of the solution with respect to the continuity of the initial data for the classical vibrating string equation with Dirichlet boundary conditions. For general linear partial differential systems that are scalar, involve only second-order spatial operators, and have time-independent coefficients, the compatibility conditions have been known since the 1950s mostly due to Ladyzenskaya [25,26]. In the 1960s, Ladyzenskaya et al. [27] and Friedman [13] discussed the regularity of the solution with respect to initial data satisfying the compatibility conditions for linear and quasi-linear parabolic systems. During the 1970s, Rauch and Massey [32] and, independently, Sakamoto [33] extended the previous results by giving the necessary and sufficient conditions that the initial data must satisfy to achieve a given order of regularity on the domain for first-order hyperbolic systems with time-dependent coefficients. Their methods can be applied to hyperbolic systems of higher order. In 1980, Smale [35] gave a new proof, similar to that of Section 2 but with much greater generality and rigor, for the heat and wave equation with Dirichlet boundary conditions. Témam [37] obtained the compatibility conditions for semi-linear evolution equations, including the Navier–Stokes equations as an important special case.

Thus, there is a substantial body of work on the theory of compatibility conditions. Unfortunately, it is confined to the mathematical literature, heavy in the jargon of functional analysis and almost impenetrable to most scientists and engineers who actually compute things. Even more unfortunately, we have not been able to locate a textbook or review that ties this theory into a coherent whole. Still, the compatibility conditions for a very wide range of problems can be looked up, if one is diligent.

7.2. Embarrassments

Unfortunately, the theory of compatibility conditions has spread little outside the mathematical community. Corner singularities are most significant for spectral methods, which converge rapidly in the absence of such pathologies. The standard books on spectral methods, however, betray a lamentable ignorance of spectral methods.

For example, consider the following misguided quote from the 1989 book of the senior author [2] (p. 379): '... for a simple problem such as Poisson's equation on a rectangle, $u_{xx} + u_{yy} = f(x, y)$, separating variables gives an eigenvalue problem in x whose solution are the sines and cosines. However, if the boundary conditions are that $u(\pm 1, y) = 0$, rather than periodicity in x, this eigenfunction series will have a slow, algebraic rate of convergence. To obtain high accuracy with non-periodic boundary conditions, we must use Chebyshev polynomials in x. ... Haidvogel and Zang observed that after we have discretizied the double x-derivative in the Poisson equation (or any other separable problem) via Chebyshev polynomials, we are guaranteed that the spectral series (with y-dependent coefficients) will converge exponentially fast.'

Actually, because of the corner singularities, the rate of convergence is NOT exponentially fast, but only high order algebraic—sixth-order for the coefficients, fourth-order for the error [21]. The book contains not a single mention of the existence of singularities in the space–time corner (though this will be corrected in the second edition, which should appear in 1999).

Gottlieb and Orszag state (p. 33) [15]: 'The important conclusion is that *eigenfuction expansions based on Sturm–Liouville problems that are singular at $x = a$ and $x = b$ converge at a rate governed by the smoothness of the function being expanded, not by any special boundary conditions satisfied by the function* [italics theirs].' This assertion is true for ordinary differential equations. For partial differential equations, however, the quote is at best misleading. The method of separation of variables, if applicable, will generate a non-singular eigenproblem. As they show, the series of non-singular eigenfunctions will converge only algebraically fast unless $u(x, t)$ satisfies an infinite set of constraints, the 'special boundary conditions' mentioned in the quote. However, these 'special boundary conditions' are merely the compatibility conditions. If $u(x, t)$ does not satisfy them, expansions in the eigenfunctions of *singular* problems, such as Chebyshev or Legendre polynomials, will *converge just as slowly.*

This thinking still dominates the numerical analysis literature twenty years later. For the radial variable in cylindrical coordinates, Bessel series have been deprecated because they are eigenfunctions of the Laplace operator, which is nonsingular at the outer boundary circle; an ingenious radial basis using functions that solve a singular eigenproblem was recently proposed in [29]. We have seen through the linear, axisymmetric wave equation that for solving partial differential equations, the clever new basis will actually be no better than Bessel functions if the compatibility conditions are ignored.

Canuto et al. [8] prove the convergence of the Chebyshev collocation scheme for the heat equation in one space dimension (one of our examples above) by deriving the estimate (p. 323), where N is the truncation of the Chebyshev series,

$$\left(\int_{-1}^{1}[(u-u^N)]^2 w\,\mathrm{d}x\right)^{1/2}+\left(\int_0^t \mathrm{d}s\int_{-1}^{1}\left[\left(\frac{\partial u}{\partial x}-\frac{\partial u^N}{\partial x}\right)\right]^2 w\,\mathrm{d}x\right)^{1/2}$$
$$\leqslant CN^{1-m}\left\{\int_0^t \mathrm{d}s\left(\left\|\frac{\partial u}{\partial t}\right\|^2_{H_w^{m-2}(-1,1)}+\|u\|^2_{H_w^m(-1,1)}\right)\right]^{1/2} \tag{66}$$

where $w=1/(1-x^2)^{1/2}$. This *seems* to predict an exponential rate of convergence since the integer m can be arbitrarily large. However, the norms on the right-hand side are those of the *exact* solution. If the exact solution is nasty, then the norms on the right-hand side become unbounded for sufficiently large m, and the inequality becomes meaningless. They add 'This inequality proves that the approximation is convergent and the error decays faster than algebraically when the solution is infinitely smooth' [p. 320]. However, they do not state the compatibility conditions which are necessary so that the solution actually *is* infinitely smooth, or note that it is the usual case, rather than the exception, for $u(x,t)$ to have a discontinuous second derivative at $t=0$.

Remarkably, knowledge of the compatibility conditions even among mathematicians is rather sketchy. Stephen Smale, who is well known for the Smale Horshoe in dynamical systems theory and a long list of other accomplishments, added the following to his paper [35] after being informed, prior to publication, that some of his work was merely an independent rederivation of earlier studies: 'One can reasonably ask: to what extent is our main theorem a new result in partial differential equations (apart from the methodology introduced here)? I have not seen it explicitly in the literature and the mathematicians in partial differential equations I've talked to were unaware of it.'

7.3. Fluid dynamics

Computational fluid dynamics is something of an exception: many non-mathematicians are aware, at least vaguely, of compatibility conditions and space–time corner singularities for the Navier-Stokes equations. Heywood and Rannacher's articles [22,23] combined compatibility theory with convergence proofs for finite element discretizations and so were read by a wider audience. The influential paper of Deville et al. [10] and especially a series of reviews by Gresho and Sani [17–20] illuminated the importance of compatibility conditions.

To illustrate compatibility conditions for fluids, consider the simplest case: two-dimensional incompressible flow in a channel which is periodic in the down-channel direction (x coordinate). The two unknowns are the streamfunction ψ and the vorticity ζ which solve

$$\zeta_t + J(\zeta,\psi)=\nu\nabla^2\zeta \quad \text{['Vorticity equation']} \tag{67}$$

$$\nabla^2\psi=\zeta \quad \text{['Streamfunction Poisson equation']} \tag{68}$$

where J is the Jacobian operator such that for arbitrary functions a, b, $J(a,b)=a_x b_y - a_y b_x$, and ν is the viscosity coefficient. The fluid velocities are the spatial derivatives of ψ:

$$u=-\psi_y\,,\qquad v=\psi_x \tag{69}$$

The usual no-slip boundary conditions are

$$u=v=0 \qquad \leftrightarrow \psi=0\,,\qquad \psi_y=0 \qquad y=\pm 1 \tag{70}$$

We thus have the rather peculiar situation of two boundary conditions on ψ at each boundary and none on the vorticity.

Since only one of the differential equations is prognostic, only a single initial condition is needed. Since the time-dependent equation is for the derivative of the vorticity, the obvious way to specify a well-posed problem is to choose the initial condition $\zeta(x,y,0)$ and then compute the corresponding initial streamfunction from the streamfunction-Poisson equation. Unfortunately, if we impose the homogeneous Dirichlet condition on ψ, the solution generally will not have a zero y-derivative at the boundaries. This implies that for arbitrary initial vorticity, the initial tangential velocity will be non-zero at the boundaries.

However, numerical algorithms for incompressible flow always impose the no-slip boundary condition. Thus, the numerically-computed flow is forced to have vortex sheets (or the algorithm's best approximation to vortex sheets) where the tangential velocity jumps discontinuously to zero at the walls. This is very reminiscent of the boundary jumps in the second derivative of the solution to the diffusion equation. Indeed, Gresho and Sani [17] note that the similarity is not accidental. Very close to the wall, the normal velocity v is tiny, and therefore the nonlinear advection of the large vorticity gradient perpendicular to the wall is negligibly small. For small times, and in the neighborhood of the wall, the vortex sheet evolves only through diffusion. This implies what they call 'a very important result: the tangential velocity component(s) near a no-slip wall and for small time will respond to the (parabolic) 1D heat equation.'

Just as for the heat equation, the diffusion rapidly smears out the initial vortex sheet into a thin layer which is resolvable by the numerical algorithm, and for larger times, the numerical computation is trustworthy. This defines a 'time of believablity' in Gresho and Sani's terminology: for earlier times, the numerical calculation is inaccurate at the walls, and therefore not to be fully believed.

The zeroth-order compatibility condition, automatically satisfied by the vorticity-streamfunction formulation, is that the initial velocity must be nondivergent. The first order condition can be expressed in explicit form by defining G as the Green's operator for the Poisson equation with homogeneous Dirichlet boundary conditions. The streamfunction is then $\psi = G\zeta$. The first order compatibility condition is that the initial vorticity must be such that the normal derivative of ψ is zero at the walls, or in other words,

$$(G\zeta)_y = 0 \qquad y = \pm 1\,, \quad t = 0 \tag{71}$$

As for other partial differential equations, the compatibility conditions are an infinite set. For example, if we differentiate the streamfunction-Poisson equation with respect to time, it becomes

$$\begin{aligned}\nabla^2\psi_t &= \zeta_t \\ &= -J(\zeta, \psi) + \nu\nabla^2\zeta\end{aligned} \tag{72}$$

where we have applied the vorticity equation to rewrite the right-hand side in the second line. Next, apply the Green's operator G and differentiate with respect to y to obtain

$$\psi_{yt} = -\{G(J(\zeta, \psi) + \nu\nabla^2\zeta)\}_y \tag{73}$$

Since the streamfunction satisfies homogeneous Dirichlet and Neumann conditions for all time, it follows that the streamfunction tendency must also be zero and have zero normal derivative at the walls for all time, too. The vanishing of the left side of Eq. (73) then yields the second-order compatibility condition

$$\{G(-J(\zeta, G\zeta) + \nu\nabla^2\zeta)\}_y = 0 \qquad y = \pm 1\,, \quad t = 0 \tag{74}$$

The argument can obviously be continued to arbitrary order of time derivative.

There is nothing special about two-dimensional incompressible flow except that this case is the easiest to describe. There are similar sets of conditions for all forms of the hydrodynamic equations. Témam [37] has given a very general treatment.

The compatibility conditions for fluids are particularly nasty because they are 'nonlocal' in Heywood and Rannacher's terminology. For ordinary problems, it is only necessary to adjust the boundary values of an initial condition. For fluids, however, one must adjust the vorticity so as to modify the boundary derivatives of a different variable ψ. There is no simple relationship between the initial condition which must be modified and the desired boundary behavior (unless one thinks the inversion of a multidimensional elliptic operator is simple!) Switching to other formulations of fluid mechanics from the vorticity/streamfunction formalism does not make the nonlocal character of the compatibility conditions disappear. Instead, Deville et al. [10] rather gloomily conclude, 'If possible the initial velocity field should be elaborated so that no singularity occurs at the initial time.... finite difference schemes are also sensitive to this [compatibility] condition ... In most applications, however, it will be very difficult to build initial compatible fields.'

8. Remedies

There are many strategies for dealing with compatibility conditions including the following:

(i) Resignation.

(ii) The climate modeller's gambit ('time of believability').

(iii) Implicit filtering by numerical schemes, such as adjusting the two highest Chebyshev coefficients of the vorticity.

(iv) Data assimilation and optimization.

We shall briefly discuss the first three in this section and data assimilation/optimization in the following two sections.

8.1. Resignation or living with it

The simplest strategy for space–time corner singularities is to ignore them. If the initial condition satisfies the zeroth-order compatibility condition, then the corner singularities will produce jumps only in the second derivatives. The effects of these jumps on $u(x, t)$ itself may be almost invisible, a sort of stealth pathology, as illustrated in Figs. 2 and 4. Furthermore, both Fourier and Chebyshev coefficients decrease as $O(1/j^3)$, and this may be acceptable, especially if the truncation N is large.

8.2. The climate modeller's gambit or reflections on the 'time of believability'

'Time heals all things' is a proverb that applies to space–time corner singularities, too, if the system is dissipative.

The goal of climate modelling is to compute the time-averaged statistics of atmospheric flows. Climate models are therefore run for many simulated months and then the first part of the output is discarded. It is only after the model has 'forgotten' the initial conditions so that the averaged flow is a function only of *processes* and *forcings*, rather than the accidents of the initial state, that the computer code truly begins to model the climate.

Gresho and Sani [17] have emphasized that there is similarly a 'time of believability' for many engineering calculations, too. Because of space–time corner singularities, the numerical calculation is not entirely believable until $t > t_b$ for some t_b. The time of believability t_b depends on how rapidly the dissipation has smeared out the discontinuities into layers thick enough to be explicitly resolved. An engineer who is trying to model the highway performance of a new automobile engine is interested in the 'climate' of the motor, and so can discard the first few hundred timesteps without worry or remorse.

The rub with this discard-the-spin-up climate modelling approach is that sometimes one wants to know whether it will rain tomorrow, and not merely whether one lives in an arid or rainy part of the globe. The automobile engineer who is trying to cure the ills of an engine that starts erratically is not interested in the climate: the spin-up of the engine *is* the problem. The need to wait until a 'time of believability' has passed may then be as unacceptable as an engineblock dry of oil.

8.3. Implicit modifications of the initial condition: Chebyshev algorithms for two-dimensional incompressible flow

Many Chebyshev algorithms modify the highest two Chebyshev coefficients so as to enforce the boundary conditions, and in the process *implicitly* modify the initial condition so as to satisfy the first-order compatibility conditions. Is this a good method of smoothing?

A comprehensive study of how Chebyshev algorithms adjust to incompatible boundary conditions must await a future study. It is instructive, however, to do one example: two-dimensional incompressible flow in a straight-walled periodic channel. If we expand the vorticity and streamfunction as a Fourier series and then for simplicity consider only a single wavenumber $\exp(ikx)$, the Poisson equation for the streamfunction becomes

$$\psi_{yy} - k^2\psi = \zeta(y) . \tag{75}$$

The usual strategy is to choose the lowest $N-1$ coefficients of the vorticity through the standard Galerkin conditions, and then modify the two coefficients of highest degree in the Chebyshev series, truncated after $(N+1)$ terms, so that

$$\psi_y(\pm 1, t=0) = 0\,. \tag{76}$$

Consider the specific case of a parabolic initial condition:

$$\zeta(y, t=0) \equiv (1-y^2) \tag{77}$$

The exact solution is

$$\psi(y, t=0) = \frac{y^2-1}{k^2} + \frac{2(1-\mathrm{sech}(k)\cosh(ky))}{k^4} \tag{78}$$

which has

$$\psi_y(\pm 1, t=0) = \frac{2}{k^2} - \frac{2\tanh(k)}{k^3}\,. \tag{79}$$

If N is an even integer, then the fix (for this symmetric initial condition) is to modify the coefficient of $T_N(y)$ in ζ so that the modified initial streamfunction has zero y-derivative and therefore zero tangential velocity at the sidewalls. The easiest way to do this is to define the auxiliary function $\sigma_N(y;k)$ as the solution to

$$\sigma_{yy} - k^2\sigma = T_N(y) \tag{80}$$

subject to homogeneous Dirichlet boundary conditions. The exact solution is a weighted sum of exp(ky) and ($-ky$) plus a polynomial in y, but for large N, the solution is rather messy. A good *approximate* solution is ($N \gg 1$, $N \gg k$)

$$\sigma_N(y;k) \approx -\frac{1}{N^2}(1-y^2)T_N(y) + O\left(\frac{k^2}{N^2}\right) + O\left(\frac{1}{N}\right) \tag{81}$$

as follows from the identities

$$\int\!\!\int T_N(x) = \frac{1}{4}\left\{\frac{T_{N+2}}{(N+1)(N+2)} - 2\frac{T_N}{N^2-1} + \frac{T_{N-2}}{(N-1)(N-2)}\right\} \tag{82}$$

$$(1-x^2)T_N(x) = \frac{1}{4}\{-T_{N+2} + 2T_N - T_{N-2}\} \tag{83}$$

A modified streamfunction and vorticity which satisfy the first compatibility condition are

$$\tilde{\zeta} = \zeta - \beta T_N(y) \tag{84}$$

$$\tilde{\psi} = \psi - \beta\sigma_N(y;k) \tag{85}$$

where

$$\beta \equiv \psi_y(\pm 1; t=0)/\sigma_{N,y}(\pm 1; t=0) \tag{86}$$

Fig. 6 shows that the modifications are quite drastic: the modified vorticity is almost fifty times larger than the original vorticity, $\zeta = 1-y^2$!

This abject failure of implicit adjustment suggests that explicit methods to modify the initial conditions are badly needed. In the next section, we describe a general framework for such modifications.

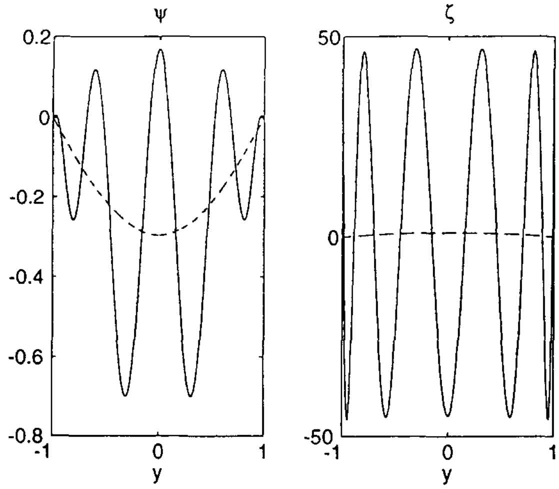

Fig. 6. Dashed curves: unmodified initial streamfunction ψ (left) and vorticity ζ (right). The unmodified ζ is a parabola in y. Solid: The modified initial vorticity ζ and streamfunction $\tilde{\psi}$ after the Nth coefficient has been modified so that $\tilde{\psi}$ satisfies both homogeneous Neumann and Dirichlet boundary conditions at the walls, $y = \pm 1$. (For this schematic, $N = 10$.)

9. Optimization/data assimilation methods for adjusting initial conditions to satisfy compatibility conditions

9.1. Initialization in weather forecasting

Numerical weather forecasting has been forced to abandon the paradigm of an initial value problem as arbitrary-initial-data-plus-evolution. One reason, as noted earlier, is that random observational errors excite spurious high-frequency gravity waves unless the observations are adjusted through an 'initialization' procedure.

A second reason is that much weather data is now asynoptic, that is, collected at different times. In ye old days, forecasts were based on radiosonde observations, released simultaneously at hundreds of stations all around the globe at noon and midnight Greenwich Mean Time (now called 'Universal Mean Time'). This data is called 'synoptic' from the Greek '$\sigma\upsilon\nu$' ['together'] and '$o\pi$' ['see'] because it is collected at a single time. In contrast, satellite observations cannot be made simultaneously, but are a continuous stream of data.

A third reason is that there are vast archives of past forecasts and observations which can in principle be used to tune tomorrow's forecast. The simple initialization schemes of the 1950s have evolved into the modern pursuit of 'data assimilation', which is a catch-all that includes both suppressing spurious gravity waves while simultaneously using asynoptic and historical data to improve the forecast [9,14].

From a forecaster's perspective, the compatibility conditions are just one more complication in a problem that is already far removed from a simple 'load-and-go' initial value problem. Indeed, weather forecasting is almost too complex to be a good paradigm for most engineering applications. We shall therefore ignore the statistical and asynoptic aspects of numerical weather prediction and focus upon the narrower problem of adjusting the initial state to be consistent with the compatibility conditions.

The basic tool of data assimilation is to perform an optimization problem which requires that the modified initial condition u_M is as close as possible to the original initial condition while simultaneously satisfying various constraints. Weather models impose a constraint of (generalized) geostrophic balance; we are interested in imposing the compatibility conditions.

9.2. Optimization: general case

Define the a_j as the spectral coefficients of the original, unmodified initial condition and the $\tilde{a}_j$ are the coefficients of the adjusted initial condition. The boundary constraints, n_b in number, are expressed in spectral form as

$$\sum_{j=1}^{N} B_{kj}\tilde{a}_j = 0 \tag{87}$$

For example, if one compatibility condition is that the second derivative should vanish at a boundary, then the corresponding B_{kj} are simply the values of the second derivative of the basis function at that boundary. Let S_j denote some 'smoothing weights' which will be defined and explained later. Define the coefficients of the adjustment function $d(x)$ by

$$d_j \equiv \tilde{a}_j - a_j \tag{88}$$

Then, the adjusted initial condition is defined as the set of spectral coefficients that minimizes the cost function J defined by

$$J \equiv (1/2)\sum_{j=0}^{N} d_j^2 + \sum_{k=1}^{n_b} \lambda_k \sum_{j=1}^{N} B_{kj}(a_j + d_j) + \frac{1}{2}\sum_{j=1}^{N} S_j d_j^2 \tag{89}$$

where the λ_k are Lagrange multipliers. These are determined along with the adjusted spectral coefficients as the solution to the unconstrained minimization of J.

The first term in the cost function forces the new spectral coefficients to be as close as possible to the old so that the initial condition is only slightly changed when forced to be compatible. The second term enforces the compatibility constraints. Note that a necessary condition for the minimum is that the gradient of J with respect to the unknowns (coefficients plus Lagrange multipliers) is zero; the subset of conditions $\partial J/\partial \lambda_k = 0$, $k = 1, \ldots, n_b$ are just the boundary constraints. The smoothness weights S_j increase rapidly with degree so that the coefficients of the adjustment function $d(x)$ are penalized for excessive wiggles through the third term in J.

The optimization problem is a very general and flexible approach to the adjustment of initial conditions because it gives us control of the adjustment. We are free to impose as few or as many compatibility conditions as we wish; the larger n_b, the higher the order of differentiability of the solution.

For simplicity, we have assumed that the compatibility conditions are linear, but the cost function is easily generalized merely by replacing the coefficients of the Lagrange multiplier by the appropriate nonlinear boundary constraint. It goes almost without saying that solving the minimization problem for the modified spectral coefficients is a non-trivial undertaking in some applications.

However, if the compatibility conditions are linear, then the conditions that the derivatives of J with respect to the spectral coefficients are zero gives

$$d_j = -\frac{1}{1+S}\sum_{k=1}^{n_b} \lambda_k B_{kj} \tag{90}$$

By substituting this into the constraint equations, the optimization problem can be reduced to n_b linear equations in the n_b unknowns, the Lagrange multipliers λ_k. Define the vectors

$$\begin{aligned} \vec{B}_k &\equiv (B_{k1}B_{k2}\cdots B_{kN})^{\mathrm{T}} \\ \hat{B}_k &\equiv \left(\frac{B_{k1}}{1+S_j}\,\frac{B_{k2}}{1+S_j}\cdots\frac{B_{kN}}{1+S_j}\right)^{\mathrm{T}}, \end{aligned} \tag{91}$$

In words, $\vec{B}_k$ is the kth row of the matrix whose elements are B_{kj} and $\hat{B}_k$ is the same vector with each element divided by $1 + S_j$ where the S_j are the smoothness weights. The Lagrange multipliers are then the solution to the $n_b \times n_b$ matrix problem

$$\vec{M}\vec{\lambda} = \vec{f} \tag{92}$$

where the kth element of $\vec{\lambda}$ is the kth Lagrange multiplier λ_k and

$$M_{ij} \equiv \vec{B}_i \bullet \hat{B}_j\,, \qquad f_i \equiv \vec{B}_k \bullet \vec{a} \tag{93}$$

where the $\bullet$ denotes the usual inner product of two vectors.

Often, symmetry allows this remaining problem to be split into two subproblems of smaller size. Thus, when the compatibility conditions are linear, the optimization problem is almost trivial as illustrated by example in the next subsection.

9.3. Adjustment-by-minimization: one-dimensional example

Suppose that the compatibility condition is that all even derivatives are zero at the endpoints, as appropriate for the one-dimensional examples discussed earlier. Suppose further that it is sufficient for our purposes to impose the first-order compatibility conditions that $u_{xx}(0,t) = u_{xx}(\pi,t) = 0$ in addition to homogeneous Dirichlet boundary conditions.

Two identities are helpful for a Chebyshev spectral method. The first is a transformation from the interval $x \in [a,b]$ to the Chebyshev argument $y \in [-1,1]$ where

$$y = \frac{2}{b-a}\left(x - \frac{b+a}{2}\right), \quad \leftrightarrow \quad x = \frac{b+a}{2} + \frac{b-a}{2}y \tag{94}$$

The second is the formula for the derivatives of the Chebyshev polynomials at the endpoints

$$\left.\frac{\mathrm{d}^p T_n}{\mathrm{d}y^p}\right|_{y=\pm 1} = (\pm 1)^{n+p} \prod_{k=0}^{p-1} \frac{n^2-k^2}{2k+1} \tag{95}$$

Another simplification is that there is no coupling between the Chebyshev coefficients of even degree and odd degree. This reduces the problem to two simpler problems, one for each parity, in half as many unknowns. (Exploiting parity is described at length in [2, Chap. 7].)

We shall assume for simplicity that $u(x)$ is a symmetric function. (General, unsymmetric u can be treated by splitting u into its symmetric and antisymmetric parts by means of $u = \mathrm{Sym} + \mathrm{AntiSym}$ where $\mathrm{Sym} = (u(x) + u(-x))/2$ and $\mathrm{AntiSym} = (u(x) - u(-x))/2$. One can then solve each problem separately as illustrated below and then add the corrections.)

The cost function has the generic form of Eq. (89). However, because $u(x)$ is assumed to be symmetric, the coefficients of the adjustment are all of even degree:

$$d_j \equiv a_{2j-2} - \tilde{a}_{2j-2}, \quad j = 1, \ldots, N \tag{96}$$

The boundary constraint matrix B has two rows and N columns where

$$B_{1j} \equiv 1, \qquad B_{2j} = \frac{4}{3}(j-1)^2(4(j-1)^2 - 1), \quad j = 1, \ldots, N \tag{97}$$

The first row is a vector that stores the boundary values of the Chebyshev polynomials of even degree while the second row stores the second derivatives of these polynomials at the endpoints. The vanishing of the N derivatives of the cost function with respect to the Chebyshev coefficients gives

$$\vec{d} = -\lambda_1 \hat{B}_1 - \lambda_2 \hat{B}_2 \tag{98}$$

where $\hat{B}_1$ and $\hat{B}_2$ are vectors which are the first and second rows of the compatibility conditions matrix B with each element divided by $1/(1+S_j)$. The final step is to solve the 2×2 matrix equation, (92) above.

To illustrate the role of the smoothness weights, we first look at the solution when $S_j = 0$, that is, when all coefficients are equally weighted in the cost function. We also assume that the unmodified initial condition is the parabola $x(\pi - x)$ so that the Dirichlet boundary conditions are satisfied before the correction, and the matrix element $f_1 = 0$ in Eq. (92).

Fig. 7 shows the adjustment function $d(x;N)$. The magnitude of d depends on the magnitude of the second derivative of the unmodified initial condition at the endpoints. The *shape* of the correction is a function only of N. However, this shape is roughly the same for all sufficiently large N: delta-like spikes at the endpoints in the second derivatives and rapid oscillations (plus roots at the endpoints) for d itself. Fig. 9 shows the corresponding coefficients.

This correction is fairly successful in the sense that the maximum value of d, when the curvature was chosen to match that of our earlier one-dimensional examples, is less than 10^{-5}. We have not graphed u_M because it is graphically indistinguishable from the original, parabolic initial condition.

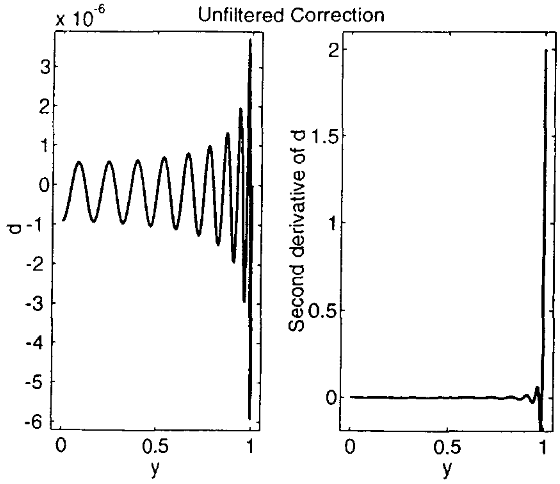

Fig. 7. The correction $d(x)$ (left panel), whose Chebyshev coefficients are the elements of the vector $\vec{d}$, and its second derivative (right panel). The axis is the rescaled coordinate y which is the argument of the Chebyshev polynomials.

However, the second derivative of d is a discrete approximation to the Dirac delta-function—zero at all grid points except that at the boundary. This lack of smoothness in the correction is likely to have a lamentable effect on the evolution of the solution.

9.4. Smoothing the adjustment

The cost function J allows one to filter or smooth the correction as little or as much as one pleases, so there can be no unique choice for the smoothness weights S_j. However, to reduce the higher coefficients of the correction, the weights S_j should rapidly increase with j. The price for smoothing is that the magnitude of the correction, that is, $\max|d(x)|$, will almost certainly increase.

The *unconstrained* smoothing problem provides some guidance for the choice of weights. If the optimization problem is the same as above except that the boundary constraints are omitted from the cost function, then the usual condition that the gradient of the cost function is zero implies

$$\tilde{a}_j = \frac{1}{1+S_j} a_j \tag{99}$$

where the a_j are the coefficients of the original unsmoothed function and the $\tilde{a}_j$ are those of the smoothed spectral series. Thus, the unconstrained smoothing problem is equivalent to mulitplying the spectral coefficients by the weights, which we shall dub the 'filter' weights,

$$\sigma_j \equiv \frac{1}{1+S_j} \tag{100}$$

There *is* an existing theory for choosing the filter weights σ as reviewed in [4].

We must emphasize that there is no proof that weights that are good for *unconstrained* filtering are necessarily the best when smoothing a function which is constrained to satisfy certain boundary or compatibility conditions. Nevertheless, unconstrained filter theory provides at least a starting point.

In [4], it is shown that a simple but effective choice is the so-called 'Erfc-Log' filter which is defined by

$$\sigma_j(p) \equiv \begin{cases} 1, & j < N - \nu + 1 \\ \dfrac{1}{2}\operatorname{erfc}\left\{ 2p^{1/2}\theta\sqrt{\dfrac{-\log(1-4\theta^2)}{4\theta^2}} \right\} & \text{otherwise} \end{cases} \tag{101}$$

where

$$\theta \equiv \frac{j + \nu - (N+1)}{\nu} - \frac{1}{2} \tag{102}$$

The parameter p is a user-choosable 'order' parameter whose role is discussed further in [4]. The parameter ν, allows the filter to be applied only to the 'tail' of the truncated spectral series so that the lowest few coefficients of the correction are unweighted ($\sigma_j = 1$, $S_j = 0$). Somewhat arbitrarily, we choose $p = 10$ and $\nu = 13$ to filter the first twenty even degree Chebyshev polynomials.

Fig. 8 shows the smoothed correction indeed is much less oscillatory than the unsmoothed correction; the coefficients are compared in Fig. 9. The price for the smoothness is that maximum pointwise value (L_∞ norm) is roughly *tripled*. However, this maximum is still smaller than the maximum of the unmodified initial condition $u_{\text{unmod}}(x) \equiv x(x - \pi)$ by a factor of 130 000! (The second derivative, not shown, is smoothed by the filtering, too, but still has a large spike at the end point but very small amplitude at all other gridpoints.)

Other filters can be applied. For example, the 'Gaussian filter'

$$\sigma_j \equiv \exp(-\Xi j^2) \tag{103}$$

much more strongly damps the higher coefficients. Fig. 10 shows that with $\Xi = 1/4$, the correction is very smooth indeed; its Chebyshev coefficients for $j > 5$ (tenth degree) are graphically indistinguishable from zero. Such smoothness enormously increases the L_∞ norm of the correction, but the maximum of $|d|$ is still smaller than the maximum of the unmodified initial condition by a factor of 400!

For the unsmoothed correction and both smoothed corrections, we cannot usefully graph the modified initial conditions because these are graphically indistinguishable from the unmodified initial condition $x(\pi - x)$.

9.5. Alternative correction methods

The machinery of constrained optimization is not the only option. Another strategy is the following. First, add a highly localized correction to the original initial condition so as to enforce compatibility conditions up to some desired order. For incompressible flow in a channel, one can add a vortex sheet at each wall. For the one-dimensional example discussed earlier in this section, one could define the second derivative of the correction to be proportional to the discrete delta-function—zero at all Chebyshev–Lobatto points except the endpoint—compute its coefficients by interpolation, and then integrate term-by-term twice using Eq. (82) to obtain $d(x)$ itself, choosing the two arbitrary constants of integration so that d is zero at both endpoints. (Recall that we assume that the unmodified initial condition already satisfies the Dirichlet boundary conditions.) The

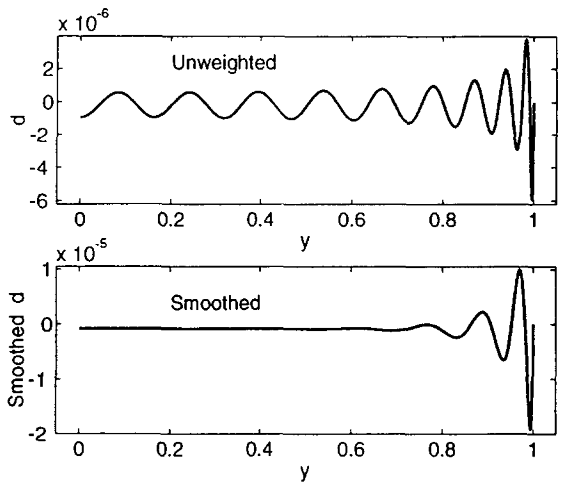

Fig. 8. The correction $d(x)$ to a parabolic initial condition both without smoothing (top) and with the Erfc-Log filter weights (bottom).

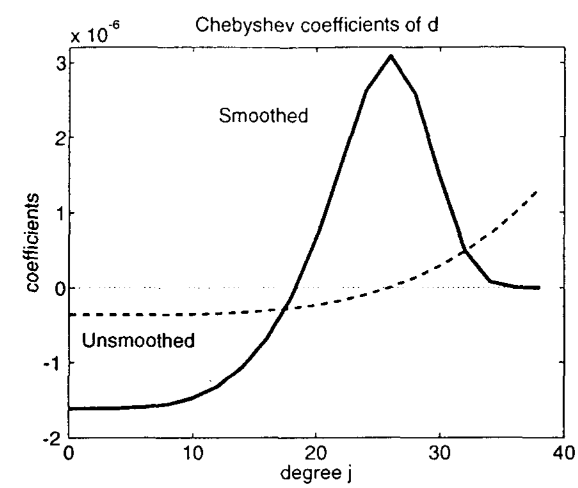

Fig. 9. Chebyshev coefficients of the correction $\vec{d}$. The thick, solid curve is for the Erfc-Log-smoothed correction; the thinner, dashed curve is the unsmoothed adjustment.

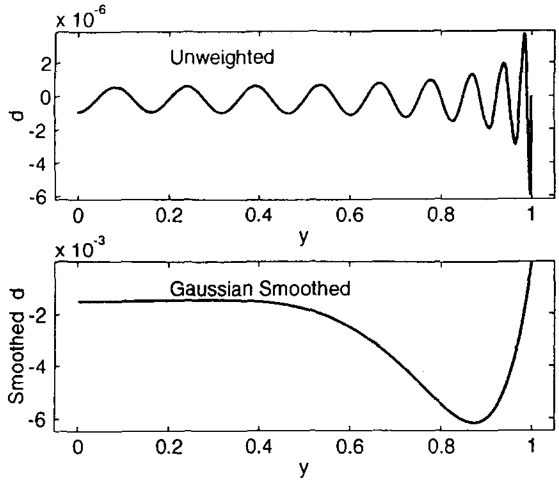

Fig. 10. The correction $d(x)$ to a parabolic initial condition both without smoothing (top) and with the Gaussian weight, $\sigma_j = \exp(-j^2/4)$ (bottom).

second step is to apply a filter to smooth the correction so that the adjustment of the initial condition does not add a lot of spurious noise to the higher spectral coefficients.

The complication is that blind application of standard filters will destroy the compatibility conditions. Oops! However, it turns out to be quite easy to modify the filtering so that all boundary/compatibility conditions are preserved. The trick is to change the basis set from the usual Chebyshev polynomials (or whatever) to new basis functions which are linear combinations of the old, chosen so that each basis function individually satisfies the boundary and/or compatibility conditions [7].

When the compatibility conditions are linear, these alternative procedures are straightforward. When the conditions are nonlocal and nonlinear, as in fluid mechanics, then there is no easy way to compute initial conditions that are both compatible and smooth. However, these difficulties have been overcome for initialization in weather forecasting where the physical constraints are also nonlinear and nonlocal. Adjustment-by-optimization and other technologies of control theory are a very general framework for making initial conditions that are compatible.

10. Summary

For the solution to an initial value problem to be analytic for all times, the initial conditions must satisfy an infinite set of compatibility conditions at the spatial boundary. Failure to satisfy these conditions implies that $u(x, t)$ will be singular at the corners of the space–time domain, even for such simple problems as the homogeneous one-space-dimensional wave equation and diffusion equation. These singularities destroy the high order convergence of spectral and spectral element methods. Worse, there is little awareness of this arithmurgical stealth bomb among numerical modellers.

However, there is some good news:

(i) The initial singularities are slowly healed (smeared out) by dissipation.

(ii) Mathematicians have derived the compatibility conditions for a wide range of systems including the Navier–Stokes equations.

(iii) The singularities are weak, typically second derivative discontinuities, and can be ignored for some applications and for second-order numerical algorithms.

(iv) The singularities can be removed by a variety of strategies that slightly modify the initial conditions.

Weather forecasting has already shown that such adjustments are often necessary for reasons other than compatibility conditions. To avoid spurious gravity waves and precipitation, meteorologists insert an 'initialization' procedure between the observed initial condition and the time-marching calculation. After being adjusted for nonlinear balance and consistency with asynoptic and historical data through this 'initialization' step, the modified initial conditions are a more faithful representation of the real atmosphere than the original uncorrected measurements.

The trouble with unadjusted meteorological observations is that they lack a sense of history. The atmosphere has had five billion years to evolve to geostrophic balance; the errors in radiosondes and anemometers are random and therefore unbalanced. The forecast is greatly improved by incorporating this history into the initial conditions by balancing the winds and temperature.

Engineers do not brood about geostrophic balance, but the forecasting paradigm applies whenever a partial differential equation imposes compatibility conditions—which is always. As the wind tunnel settles into a steady state, the air is adjusting to the compatibility conditions. An initial condition which has been machine-modified to reflect this history is more physical than an experimentally-measured-and-interpolated initial state which violates the compatibility conditions.

The mathematical theory of compatibility conditions is in reasonably good shape. However, even here, there are many PDEs or PDE systems for which the compatibility conditions are unknown, or are known only through non-rigorous arguments.

The engineering of initial conditions to satisfy compatibility conditions is largely unexplored. We have offered a few halting, tentative experiments in the previous two sections. A thorough study with comparison of different methods is a task for the future. Knowing that one *ought* to do the right thing is not the same as knowing *how* to do the right thing. This article explains why the compatibility conditions are important. The handbook of initial condition adjustments is still unwritten.

Mathematics published only in mathematical journals, and disseminated erratically even within the mathematical community, is as useful to engineers as a really good archive of Etruscan literature.[1] We hope that this article will increase awareness of the compatibility conditions among those who actually solve things. To paraphrase what Gorky said of war: 'One may not be interested in compatibility conditions, but space–time singularities are certainly interested in you.'

Acknowledgments

This work was supported by the National Science Foundation through OCE9521133 and by NASA through NGT-51409. We wish to thank Jeffrey Rauch for several valuable discussions. We are also appreciative to the reviewers for helpful comments. We thank George Em Karniadakis for editing this article.

Appendix A. Symmetries of imbricate series

THEOREM 3 (Symmetries of Imbricate series)

(i) *Let u_S denote a periodic function of period L which is defined by an alternating imbricate series with a pattern function G which is symmetric with respect to $x = L/4$:*

$$u_S \equiv \sum_{m=-\infty}^{\infty} (-1)^m G(x - mL/2) \tag{A.1}$$

Then u_S must be antisymmetric with respect to the origin.

(ii) *Let u_A denote a regular (non-alternating) imbricate series of period $L/2$ with a pattern function G which is antisymmetric with respect to $x = L/4$:*

$$u_A \equiv \sum_{m=-\infty}^{\infty} G(x - mL/2) \tag{A.2}$$

Then, u_A is also antisymmetric with respect to $x = 0$.

PROOF (*i*). The first step is to group the copies of the pattern function in pairs by writing

[1] Many Etruscan inscriptions have survived, but their script has never been deciphered.

$$\sum_{m=-\infty}^{\infty} (-1)^m G(x - m(L/2)) = \sum_{n=-\infty}^{\infty} \Lambda_n(x) \tag{A.3}$$

where Λ_n is the sum of the $m = -(2n + 1)$ term with the $m = 2n$ term:

$$\Lambda_n \equiv -G(x + (2n + 1)(L/2)) + G(x - 2n(L/2)) \tag{A.4}$$

Now, $G(x + (2n + 1)(L/2))$ is a copy of the pattern function which is centered at $x = -(2n + 1)(L/2)$. It follows, because of the assumed symmetry of the pattern $G(x)$, that this copy of the pattern function is also symmetric with respect to a point $L/4$ to the right of the point where it is centered. This symmetry with respect to $x = -(2n + 1/2)(L/2)$ implies

$$\begin{aligned} G(x + (2n + 1)(L/2)) &= G([x + (2n + 1/2)(L/2)] + L/4) \\ &= G([-x - (2n + 1/2)(L/2)] + L/4) \end{aligned} \tag{A.5}$$

This implies that, without approximation,

$$\Lambda_n \equiv -G(-x - 2n(L/2)) + G(x - 2n(L/2)) \tag{A.6}$$

It follows that replacing the argument of Λ_n by its negative,

$$\Lambda_n(-x) = -G(x - 2n(L/2)) + G(-x - 2n(L/2)) \tag{A.7}$$

which is just the negative of Λ_n. It follows that each pair individually satisfies the antisymmetry condition

$$\Lambda_n(x) = -\Lambda_n(-x) \tag{A.8}$$

Therefore, the series as a whole must satisfy this same condition.

PROOF (*ii*). Group the copies of the pattern function in pairs by writing

$$\sum_{m=-\infty}^{\infty} G(x - m(L/2)) = \sum_{n=-\infty}^{\infty} \Lambda_n(x) \tag{A.9}$$

where Λ_n is the sum of the $m = -(2n + 1)$ term with the $m = 2n$ term:

$$\Lambda_n \equiv G(x + (2n + 1)(L/2)) + G(x - 2n(L/2)) \tag{A.10}$$

which differs from the previous definition of Λ_n only through the omission of the minus sign in front of the first term. Now $G(x + (2n + 1)(L/2))$ is a copy of the pattern function which is centered at $x = -(2n + 1)(L/2)$. It follows that it is antisymmetric with respect to a point $L/4$ to the right of the point where it is centered. This antisymmetry with respect to $x = -(2n + 1/2)(L/2)$ implies

$$\begin{aligned} G(x + (2n + 1)(L/2)) &= G([x + (2n + 1/2)(L/2)] + L/4) \\ &= G([-x - (2n + 1/2)(L/2)] + L/4) \end{aligned} \tag{A.11}$$

This implies that, without approximation,

$$\Lambda_n \equiv -G(-x - 2n(L/2)) + G(x - 2n(L/2)) \tag{A.12}$$

which is identical with Eq. (A.7) above. The rest of the proof is then line-by-identical to that for the first part of the theorem.

Appendix B. Bessel function solution of the axisymmetric wave equation

To illustrate the key ideas, it is sufficient to specialize to Dirichlet boundary conditions. The problem is

$$u_{tt} = u_{rr} + \frac{1}{r} u_r \tag{B.1}$$

subject to the boundary conditions

$$u(1, t) = 0\,, \qquad u(0, t) = \text{bounded} \tag{B.2}$$

and the initial conditions

$$u(r, 0) = u_0(r)\,, \qquad u_t(r, 0) = v_0(r) \tag{B.3}$$

The solution is

$$u(r, t) = \sum_{n=1}^{\infty} J_0(k_n r)\{a_n \cos(k_n t) + b_n \sin(k_n t)\} \tag{B.4}$$

$$a_n = \frac{2}{J_1^2(k_n)} \int_0^1 r u_0(r) J_0(k_n r)\,\mathrm{d}r \tag{B.5}$$

$$b_n = \frac{2}{J_1^2(k_n)} \frac{1}{k_n} \int_0^1 r v_0(r) J_0(k_n r)\,\mathrm{d}r \tag{B.6}$$

Bessel coefficients can be integrated-by-parts by using the following observations:

(i)

$$J_0(k_n r) = -\frac{1}{k_n^2} \nabla^2 J_0(k_n r) \tag{B.7}$$

because the basis functions are eigenfunctions of the Laplace operator;

(ii) the Laplace operator is self-adjoint with respect to the inner product defined by

$$(f, g) \equiv \int_0^1 r\,\mathrm{d}r\, f(r) g(r) \tag{B.8}$$

for any two functions $f(r)$, $g(r)$ in the sense that for any two functions that satisfy the boundary conditions

$$(f, \nabla^2 g) = (\nabla^2 f, g) \tag{B.9}$$

as can be proved by integration by parts.

These two observations make it possible to introduce the Laplace operator inside the coefficient integrals (operating on the Bessel function), and then transfer it using the self-adjointness theorem to operating on the initial condition $u_0(r)$ instead.

Every time that we apply this trick, we rewrite the integral as the product of a power of $1/k_n$, raised by two with each application, times an integral which can be bounded by a constant independent of n. This gives rate-of-convergence bounds for Bessel series. To express these in terms of n rather than k_n, use

$$k_n \sim (n - 1/4)\pi\,, \quad n \to \infty\,. \tag{B.10}$$

If we can apply this trick an arbitrarily large number of times, then the coefficients of the Bessel series must decay faster than any finite power of n.

Unfortunately, the self-adjointness property applies only if both functions in the inner product are zero at the boundaries. If the result of repeated applications yields a function which does not vanish at the boundary—unfortunately true of an initial condition that fails the compatibility conditions at some order—then the integration-by-parts yields a boundary term which does not vanish, and we can shift the Laplace operator no further.

For the special case

$$u_0(r) = 1 - r^2\,, \qquad v_0(r) = 0 \tag{B.11}$$

illustrated in the main text, the coefficient integral transforms to

$$
\begin{aligned}
a_n &= -\frac{2}{J_1^2(k_n)k_n^2}\int_0^1 r\{\nabla^2 u_0(r)\}J_0(k_n r)\,\mathrm{d}r \\
&= -\frac{2}{J_1^2(k_n)k_n^2}\int_0^1 r\{-4\}J_0(k_n r)\,\mathrm{d}r \\
&= \frac{8}{k_n^3 J_1(k_n)}
\end{aligned}
\tag{B.12}
$$

without approximation where we have used the identities $\nabla^2(1-r^2) = -4$ and $\int_0^1 rJ_0(k_n r)\,\mathrm{d}r = 1/k_n J_1(k_n)$ to obtain the exact coefficients for this special initial condition.

By applying the asymptotic approximations for the Bessel functions and their zeros, one finds

$$
a_n \sim \frac{4\sqrt{2}}{\pi^2}\frac{(-1)^{n+1}}{n^{5/2}}, \quad n \gg 1 \tag{B.13}
$$

This seems to suggest that the Bessel series has convergence order 5/2, which is slower than the $O(1/n^3)$ convergence of the Chebyshev coefficients for this example. However, this is a mirage. The argument of the Bessel functions, being proportional to k_n and therefore asymptotically to n, grows larger and larger as n increases for any fixed r. It follows that for fixed radius, we can always approximate the high degree, infinitely long 'tail' of the series by replacing the Bessel functions by their asymptotic approximations:

$$
J_0(k_n r) \sim \frac{\sqrt{2}}{\pi}\frac{1}{\sqrt{n}\sqrt{r}}\cos\left\{\left(n-\frac{1}{4}\right)\pi r - \frac{\pi}{4}\right\}, \quad n \to \infty, \quad \text{FIXED } r \tag{B.14}
$$

Thus, the terms of the Bessel series are decreasing as $O(1/n^3)$ for large n, neither faster nor slower than that of the Chebyshev series for the same solution.

References

[1] D.G. Bourgin and R. Duffin, The Dirichlet problem for the vibrating string equation, Am. Math. Soc. Bull. 45 (1939) 851–858.

[2] J.P. Boyd, Chebyshev and Fourier Spectral Methods (Springer-Verlag, New York, 1989).

[3] J.P. Boyd, A lag-averaged generalization of Euler's method for accelerating series, Appl. Math. Comput. 72 (1995) 143–166.

[4] J.P. Boyd, The Erfc-Log filter and the asymptotics of the Euler and Vandeven sum accelerations, in: A.V. Ilin and L.R. Scott, eds., Proc. Third Int. Conf. on Spectral and High Order Methods (Houston J. Mathematics, Houston, TX, 1996) 267–276.

[5] J.P. Boyd, New directions in solitons and nonlinear periodic waves: Polycnoidal waves, imbricated solitons, weakly non-local solitary waves and numerical boundary value algorithms, in: T.-Y. Wu and J.W. Hutchinson, eds., Advances in Applied Mechanics 27 (Academic Press, New York, 1989) 1–82.

[6] J.P. Boyd, Construction of Lighthill's unitary functions: The imbricate series of unity, Appl. Math. Comput. 86 (1997) 1–10.

[7] J.P. Boyd, Two comments on filtering (artificial viscosity) for Chebyshev and Legendre spectral and spectral element methods: Preserving boundary conditions and interpretation of the filter as a diffusion, J. Comput. Phys. 143 (1998) 283–288.

[8] C. Canuto, M.Y. Hussaini, A. Quarteroni and T.A. Zang, Spectral Methods in Fluid Dynamics (Springer-Verlag, New York, 1987).

[9] R. Daley, Atmospheric Data Analysis (Cambridge University Press, New York, 1991).

[10] M. Deville, L. Kleiser and F. Montigny-Rannou, Pressure and time treatment for Chebyshev spectral solution of a Stokes problem, Int. J. Numer. Methods Fluids (1984) 1149–1163.

[11] B.A. Finlayson, The Method of Mean Weighted Residuals and Variational Principles (Academic Press, New York, 1972) 412.

[12] B. Fornberg, A Practical Guide to Pseudospectral Methods (Cambridge University Press, New York, 1996).

[13] A. Friedman, Partial Differential Equations of Parabolic Type (Prentice-Hall, Englewood Cliffs, NJ, 1964).

[14] M. Ghil and P. Malanotte-Rizzoli, Data assimilation in meteorology and oceanography, Adv. Geophys. 33 (1991) 141–266.

[15] D. Gottlieb and S.A. Orszag, Numerical Analysis of Spectral Methods (SIAM, Philadelphia, 1977).

[16] D. Gottlieb and C.-W. Shu, On the Gibbs Phenomenon and its resolution, SIAM Rev. 39 (1997) 644–668.

[17] P.M. Gresho and R.L. Sani, On pressure boundary conditions for the incompressible Navier–Stokes equations, Int. J. Numer. Methods Fluids 7 (1987) 1111–1145.

[18] P.M. Gresho, Incompressible fluid dynamics: some fundamental formulation issues, Ann. Rev. Fluid Mech. 23 (1991) 413–453.

[19] P.M. Gresho, Some current CFD issues relevant to the incompressible Navier–Stokes equations, Comput. Methods Appl. Mech. Engrg. 87 (1991) 201–252.

[20] P.M. Gresho, Some interesting issues in incompressible fluid dynamics, both in the continuum and in numerical simulation, in: T.-Y. Wu and J.W. Hutchinson, eds., Advances in Applied Mechanics 28 (Academic Press, New York) 46–140.

[21] D.B. Haidvogel and T.A. Zang, The accurate solution of Poisson's equation by expansion in Chebyshev polynomials, J. Comput. Phys. 30 (1979) 167–180.

[22] J.C. Heywood and R. Rannacher, Finite element approximation of the nonstationary Navier–Stokes problem. I: Regularity of the solutions and second-order error estimates for the spatial discretization, SIAM J. of Numer. Anal. 19 (1982) 275–311.

[23] J.C. Heywood and R. Rannacher, Finite element approximation of the nonstationary Navier–Stokes problem. II: Stability of solutions and error estimates uniform in time, SIAM J. Numer. Anal. 23 (1986) 750–777.

[24] F. John, The Dirichlet problem for a hyperbolic equation, Am. J. Math. 63 (1941) 141–154.

[25] O. Ladyzenskaja, On the convergence of Fourier series defining a solution of a mixed problem for hyperbolic equations, Dokl. Akad. Nauk. SSSR 85 (1952) 481–484.

[26] O. Ladyzenskaja, On the solvability of the fundamental boundary problems for equations of parabolic and hyperbolic type, Dokl. Akad. Nauk. SSSR 87 (1954) 395–398.

[27] O. Ladyzenskaja, V. Solonnikov and N. Ural'ceva, Linear and quasi-linear equations of parabolic type, Vol. 23, Translations of Math. Monog. (AMS, Providence, RI, 1968).

[28] P. Lynch, Richardson's barotropic forecast: A reappraisal, Bull. Am. Met. Soc. 73 (1992) 35–47.

[29] T. Matsushima and P.S. Marcus, A spectral method for polar coordinates, J. Comput. Phys. 120 (1995) 365–374.

[30] B. Mercier, An Introduction to the Numerical Analysis of Spectral Methods (Springer-Verlag, Berlin, 1989).

[31] H.K. Moffat, Viscous and resistive eddies near a sharp corner, J. Fluid Mech. 18 (1964) 1–18.

[32] J.B. Rauch and F.J. Massey, Differentiability of solutions to hyperbolic initial-boundary value problems, Trans. Am. Math. Soc. 189 (1974) 303–318.

[33] R. Sakamoto, Hyperbolic Boundary Value Problems (Cambridge University Press, Cambridge, 1982) (first published in Japanese in 1978) 204.

[34] W.W. Schultz, N.Y. Lee and J.P. Boyd, Chebyshev pseudospectral method of viscous flows with corner singularities, J. Sci. Comput. 4(1) (1989) 1–24.

[35] S. Smale, Smooth solutions of the heat and diffusion equations, Comment. Math. Helvetici 55 (1980) 1–12.

[36] D.A. Smith and W.F. Ford, Numerical comparisons of nonlinear convergence accelerators, Math. Comput. 38 (1982) 481–499.

[37] R. Témam, Behaviour at time $t = 0$ of the solutions of semi-linear evolution equations, J. Diff. Eqns. 43 (1982) 73–92.

[38] E.J. Weniger, Nonlinear sequence transformations for the acceleration of convergence and the summation of divergent series, Comput. Phys. Rep. 10 (1989) 189–371.

[39] E.J. Weniger, On the derivation of iterated sequence transformations for the acceleration of convergence and the summation of divergent series, Comput. Phys. Comm. 64 (1991) 19–45.

[40] D.V. Widder, The Heat Equation (Academic Press, New York, 1975).