SIAM REVIEW
Vol. 62, No. 2, pp. 439–462

Eight Perspectives on the Exponentially Ill-Conditioned Equation $\varepsilon y'' - xy' + y = 0$*

Lloyd N. Trefethen†

Abstract. Boundary-value problems involving the linear differential equation $\varepsilon y'' - xy' + y = 0$ have surprising properties as $\varepsilon \to 0$. We examine this equation from eight points of view, showing how it sheds light on aspects of numerical analysis (backward error analysis and ill-conditioning), asymptotics (boundary layer analysis), dynamical systems (slow manifolds), ODE theory (Sturm–Liouville operators), spectral theory (eigenvalues and pseudospectra), sensitivity analysis (adjoints and SVD), physics (ghost solutions), and PDE theory (Lewy nonexistence).

Key words. ill-conditioning, boundary layer analysis, slow manifold, turning point, pseudospectra, Sturm–Liouville operator

AMS subject classification. 34E20

DOI. 10.1137/18M121232X

1. Introduction.

Recently I visited Bernd Krauskopf and Hinke Osinga and their dynamical systems group at the University of Auckland. In our discussions, we were fascinated to find the different perspectives we had on the simple equation

$$\varepsilon y'' - xy' + y = 0, \tag{1.1}$$

where ε is a small positive constant. My interpretations were rooted in numerical analysis and theirs in geometry, and it was remarkable what deep differences in perception this led to. As we examined the equation, we found many other viewpoints coming into the picture too. This paper employs this archetypal equation to make connections among a wide range of mathematical topics. It is aimed at anyone who has studied differential equations and would like to see more of their implications.

Equations like (1.1) were mentioned at least as long ago as the 19th-century book by Forsyth [14, section 58], and it is a variant of the *Hermite equation* $y'' - 2xy' = \lambda y$ related to Hermite polynomials [46, Table 18.8.1]. Its crucial feature is the change of sign of the coefficient x at $x = 0$, which is called a *turning point,* and attention was drawn to the difficulties of such equations in a 1970 paper by Zuckerberg and O'Malley [1]. In the Soviet literature, such problems had been considered by Tikhonov and others beginning in the late 1940s [56, 57, 63], as reviewed in a SIAM book and a *SIAM Review* article in the 1990s by Tikhonov's student Adelaida Vasilieva [61, 62].

*Received by the editors September 7, 2018; accepted for publication (in revised form) January 4, 2019; published electronically May 7, 2020.
https://doi.org/10.1137/18M121232X
Funding: This work was supported by the Margaret and John Kalman Charitable Trust.

†Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK (trefethen@maths.ox.ac.uk).

Our interest will be in the context of a two-point boundary-value problem on $[-1,1]$. To be definite, let us arbitrarily fix the boundary values

$$y(-1) = 2, \quad y(1) = 1. \tag{1.2}$$

The unusual properties of boundary-value problems involving (1.1) have been discussed in many works since [1], including [19, 20, 26, 34, 35, 36, 37, 39, 42, 49, 51, 59, 65], and surveys can be found in [53] and [64]. The present paper differs from most of the existing literature in that it is not so much focused on solving these challenging problems as on examining the deeper significance of what makes them challenging.

One aspect of (1.1) is that it is far from self-adjoint. Its adjoint equation,

$$\varepsilon y'' + xy' + 2y = 0, \tag{1.3}$$

has equally extreme but complementary properties. In the final sections 8–10 of this paper, the focus will be on (1.3) as much as (1.1).

2. Existence, Uniqueness, and Exact Solution. Equation (1.1) takes the form of a linear, second-order, variable coefficient ODE with a nonvanishing leading-order term. As shown in many textbooks, it will accordingly have a two-dimensional vector space of solutions, and in fact, since the coefficients are analytic functions of x, the solutions will be analytic for all x and could be computed in principle by Taylor series. One solution is simply $y(x) = x$. The general solution can be written

$$y(x) = Ay_{\text{odd}}(x) + By_{\text{even}}(x) = Ax + B\left[-e^{x^2/2\varepsilon} + x\varepsilon^{-1}\int_0^x e^{s^2/2\varepsilon}ds\right] \tag{2.1}$$

for arbitrary constants A and B (see [26, eq. (1.3.10)] or [64, eq. (16)]). Note that this representation has the property that the first term is odd and the second is even. Alternative formulas for the analytic solution in terms of parabolic cylinder (= Weber) functions [45] are given in a number of references including [26, 37, 64].

The distinctive feature of (2.1) is the exponential growth of y_{even} when x is large, whether positive or negative. This function consists of the difference of two positive quantities, namely, $e^{x^2/2\varepsilon}$ and $x\varepsilon^{-1}$ times its indefinite integral. For $\varepsilon \ll x^2$, the latter term dominates, making y_{even} exponentially large and positive. We use the word "exponentially" loosely. More precisely, $y_{\text{even}}(x) \sim \varepsilon x^{-2}\exp(x^2/2\varepsilon)$ as $\varepsilon x^{-2} \to 0$.

Now let us bring in the boundary conditions (1.2). This gives a mathematically standard linear two-point boundary-value problem (BVP), which will have a unique solution if and only if the corresponding homogeneous problem,

$$\varepsilon y'' - xy' + y = 0, \quad y(-1) = 0,\ y(1) = 0, \tag{2.2}$$

has only the zero solution. This is indeed the case when ε is small, as we can argue as follows. Since y_{even} is even, zero boundary conditions imply $A = 0$ in (2.1). So $B = 0$ is the unique value satisfying (2.2) if and only if $y_{\text{even}}(1)$ is nonzero. In fact, $y_{\text{even}}(1)$ is zero for $\varepsilon \approx 0.585457706$ and positive for smaller values of ε, growing exponentially as $\varepsilon \to 0$, as explained above.

Figure 2.1 shows the solutions to (1.1)–(1.2) corresponding to $\varepsilon = 0.1$ and 0.02. These curves, computed with a Chebyshev spectral collocation method by Chebfun [11, 59], are correct to plotting accuracy (actually much better). The constants in (2.1) for the exact solution are $A = -1/2$ and $B = (3/2)/y_{\text{even}}(1)$, confirming that for small ε, the solution consists of the straight line $y(x) = -x/2$ plus narrow boundary layers of amplitude 3/2.

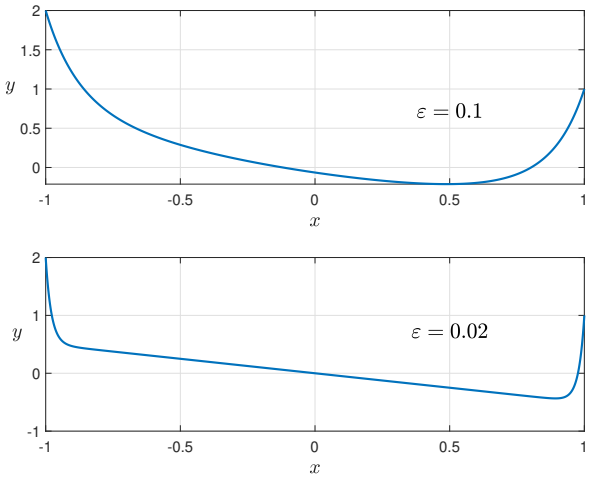

Fig. 2.1 *Solutions to* (1.1)–(1.2) *for two values of* ε *show* $y(x) \approx -x/2$ *in the interior with boundary layers of equal amplitudes at the ends.*

3. Numerical Analysis: Backward Error Analysis and Ill-Conditioning. As soon as one tries smaller values of ε, however, seemingly nonsensical results begin to appear. Figure 3.1 shows the solutions computed by Chebfun and by the MATLAB code `bvp4c` [33] when ε is set to 0.009. These are nowhere near correct. If we try various choices of ε about 0.018 or smaller, we get numerical solutions with the straight segment in the middle oriented with an apparently arbitrary slope.[1] (We chose 0.009 rather than 0.01 for the figure because it happens to give a positive slope, highlighting the contrast with Figure 2.1.)

Computational failures like this, which Bob O'Malley has called "numerical nightmares" [49], must have disturbed a number of people over the years. In fact, during the six months after the publication of *Exploring ODEs* [59], I was contacted independently by two ODE experts about difficulties they were having with this kind of equation in Chebfun (O'Malley in May 2018 and Folkmar Bornemann in July). This added to the motivation to write this paper.

Why do Chebfun and `bvp4c` fail when ε is small? The numerical analyst quickly spots the explanation (and deflects the blame): this is an illustration of the principle of *backward error analysis* made famous by Wilkinson half a century ago [27, 66]. Wilkinson showed that for all kinds of good numerical algorithms, rounding errors and other perturbations can be interpreted as having the effect that one computes the *exact solution to a slightly perturbed problem*. If the problem is well-conditioned, this implies that the computed solution will be close to correct, but if it is highly ill-conditioned—that is, highly sensitive to perturbations—the computed solution may be very far off. This is just what has happened here. For backward error analysis applied specifically to differential equations, some references are given in [27].

To confirm this explanation of Figure 3.1, we can insert the erroneous computed solution y_{comp} into the differential equation. As Wilkinson would have immediately

[1] For $\varepsilon = 0.017$ or $\varepsilon = 0.018$ in our experiments, `bvp4c` gives a warning message: "Unable to meet the tolerance." With smaller values of ε, however, there is no message. For computations in such a regime it may be effective to use a variable precision arithmetic tool such as Maple [64].

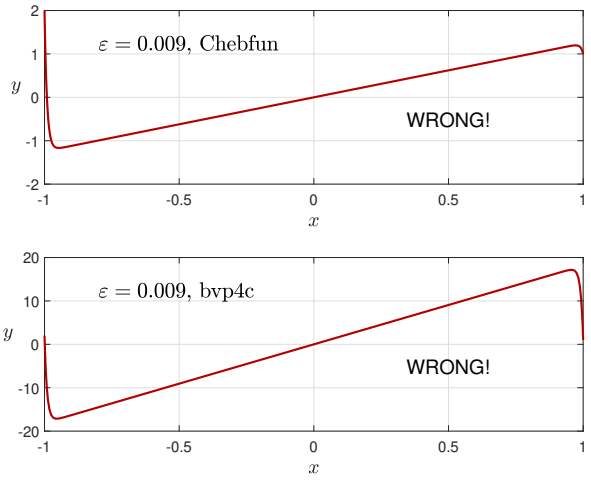

Fig. 3.1 *Erroneous solutions to* (1.1)–(1.2) *for the smaller value* $\varepsilon = 0.009$. *The first is computed by Chebfun and the second by the MATLAB code* `bvp4c`. *Note that the boundary layers do not even have the same sign, though we know from the exact solution* (2.1) *that they ought to be equal in sign and magnitude.*

understood, y_{comp} satisfies the equation to high accuracy, with a maximal error of 1.8×10^{-9} near the boundaries and closer to 10^{-11} in most of the domain. In other words, though the error $y_{\text{comp}} - y_{\text{exact}}$ is large, the residual $L(y_{\text{comp}} - y_{\text{exact}})$, where L is the differential operator $L : y \mapsto \varepsilon y'' - xy' + y$, is very small. The computed function y_{comp} is not *near* the solution of (1.1)–(1.2), but it is *nearly* a solution.

There are many ways of understanding the ill-conditioning of (1.1)–(1.2), and in fact one could interpret this whole paper as an investigation of this matter from various angles. Here in the numerical section, let us confirm the ill-conditioning numerically by perturbing the problem a little in a regime where ε is large enough for accurate numerical solutions. Figure 3.2 shows that the $\varepsilon = 0.02$ solution of Figure 2.1 changes completely if $Ly = 0$ is replaced by $Ly = f$ with $f(x) = -10^{-7}x$, or by $\tilde{L}y = 0$ with $\tilde{L} : y \mapsto \varepsilon y'' - xy' + 0.9999999y$. Another solution much like the second of these is produced with the perturbed operator $\tilde{L} : y \mapsto \varepsilon y'' - (x + 10^{-6}x^3)y' + y$ (not shown). It is typical for perturbations of the problem (1.1)–(1.2) to lead to solutions like these that are close to zero in the interior of the interval, an effect that is analyzed in [1, 35, 51] and a number of other publications. In particular, a solution close to zero will appear if the coefficient 1 multiplying the y term in (1.1) is changed to anything other than a nonnegative integer. In the literature of turning point problems, cases like (1.1) with nonzero behavior in the interior are regarded as the special ones, characterized by what is known as the *Matkowsky condition* [41].

Examples of backward error analysis involving extreme ill-conditioning raise philosophical questions. Have Chebfun and `bvp4c` simply failed in Figure 3.1, since their results are far from correct, or have they in a deeper sense succeeded, since they have computed the right answer to almost the right problem? There is justice in both of these points of view. One sense in which both algorithms have truly failed is that there are other algorithms which, even in the same 16-digit floating-point arithmetic, can compute the right answer for $\varepsilon = 0.009$ or indeed for lower values of ε. Bornemann has shown me successful results with the ultraspherical spectral method

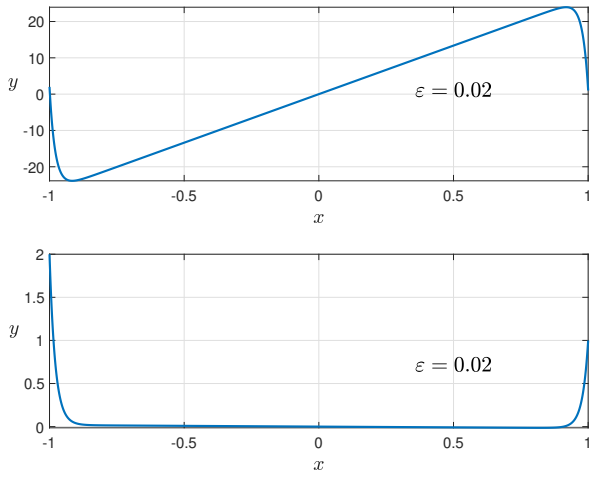

Fig. 3.2 *Repetition of the second panel of Figure* 2.1*, but for two slightly perturbed problems. The first image corresponds to* $Ly = -10^{-7}x$ *instead of* $Ly = 0$ *and the second to* $\tilde{L}y = 0$ *with* $\tilde{L} : y \mapsto \varepsilon y'' - xy' + 0.9999999y$*. Though the perturbations are small, the solutions differ greatly from those of Figure* 2.1*.*

of Olver and Townsend [47] down to $\varepsilon = 10^{-5}$, where the condition number with respect to perturbations of the right-hand side is of order $e^{50,000}$. An algorithm that succeeds under such circumstances must exploit the fact that the right-hand side is known to be exactly zero—analogously to an algorithm of numerical linear algebra imposing a condition of structural sparsity. The same algorithm will still probably fail for a problem $Ly = f$ in which f is a nonzero function represented in floating-point arithmetic.[2] Similarly, it will probably fail for a homogeneous problem with variable coefficients defined by functions represented numerically.

4. Asymptotics: Boundary Layer Analysis. The problem (1.1)–(1.2) happens to have an analytic solution (2.1), but this is unusual. To treat more general problems with small parameters, a powerful method of *boundary layer analysis* was developed in the 20th century [7, 29, 32, 62, 63]. This provides a means to obtain approximate solutions that are accurate for small ε and the become exact in the limit $\varepsilon \to 0$.

Let us apply boundary layer analysis to (1.1)–(1.2); see, e.g., [7, section 9.6]. The first step is to examine the solution in the so-called *outer regions*, outside any layers with rapid transitions. Following the formulation of [59], we apply the following principle:

1. *Outside a boundary or interior layer, terms involving* ε *are negligible.*

Deleting the ε term of (1.1) yields the *outer equation*

$$-xy' + y = 0, \quad \text{i.e.,} \quad \frac{y'}{y} = \frac{1}{x}. \tag{4.1}$$

Solving this ODE, we find that $\log y$ and $\log x$ are equal up to a constant, implying

$$\textit{Outer solution:} \quad y = C_1 x, \ C_1 = \text{constant}. \tag{4.2}$$

[2] Actually, for this problem the ill-conditioning is only excited if f is not even, so if f is even, it would be enough for an algorithm to maintain exact symmetry.

For this problem, (4.2) happens to be an exact solution of the ODE, though in general the outer solution would only be approximately valid. Note that at this stage in the analysis we have one parameter to be determined, the constant C_1.

Next we turn to the transition layers. In general, a problem might have a boundary layer at one or more boundaries, and it might also have interior layers at points where a coefficient changes sign. These are the principles as stated in [59]:

2. *Inside a layer, a low-order derivative is negligible compared with a higher-order one.*

3. *Inside a layer, a variable coefficient can be approximated by a constant, at least if it is locally nonzero.*

Let us first apply these principles at $x = -1$. Equation (1.1) becomes $\varepsilon y'' + y' = 0$, whose solution is

$$y(x) = C_2 + C_3 e^{-(1+x)/\varepsilon}, \tag{4.3}$$

where C_2 and C_3 are arbitrary. What is important here is that the C_3 term involves the exponentially decaying function $e^{-x/\varepsilon}$ rather than an exponentially growing one such as $e^{x/\varepsilon}$. This tells us that a boundary layer may indeed be possible at the left boundary, and its thickness will scale as $O(\varepsilon)$. Moreover, following the procedure known as *matched asymptotic expansions* (which in this case reduces to just adding up the various terms), we can combine (4.2) and (4.3) into a single function

$$y(x) = Cx + (2 + C)e^{-(1+x)/\varepsilon}. \tag{4.4}$$

For any value of C, this satisfies the boundary condition $y(-1) = 2$, and it approximately satisfies the ODE everywhere, with improving accuracy as $\varepsilon \to 0$. Even better approximations could be constructed by taking further terms in an asymptotic series with respect to ε.

So far, this could be a standard application of boundary layer theory. If the coefficient of the y' term in (1.1) did not change sign from one end of the domain to the other, we would apply reasoning like that leading to (4.3) at $x = 1$ and again find a solution involving a term decaying exponentially with x. This would rule out the possibility of a boundary layer at the right, with the consequence that the choice of C would be fixed by the right boundary condition. With this fixed C, we would have a single approximate solution in the form (4.4) for the whole BVP (1.1)–(1.2).

However, the coefficient of the y' term in (1.1) does change sign: the equation has a turning point at $x = 0$. Applying principles 2 and 3 at $x = 1$, we get the equation $\varepsilon y'' - y' = 0$, with solution

$$y(x) = C_4 + C_5 e^{-(1-x)/\varepsilon} \tag{4.5}$$

for constants C_4 and C_5. Now the exponential term decays as x *decreases,* implying that a boundary layer may appear at the right as well as at the left. This is too much freedom! Boundary layer analysis has led us to the approximate global solution

$$y(x) = Cx + (2 + C)e^{-(1+x)/\varepsilon} + (1 - C)e^{-(1-x)/\varepsilon}. \tag{4.6}$$

For any choice of C, this function nearly satisfies (1.1) and the boundary conditions (1.2), with the approximation improving as $\varepsilon \to 0$ (see Figure 4.1). Yet it is only for $C = -1/2$ that it approximates the true solution. The correct value of C

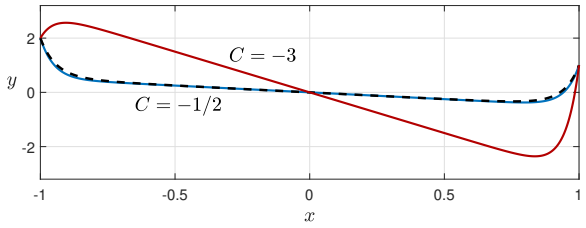

Fig. 4.1 *True solution (dashed) and two boundary layer analysis approximations* (4.6) *(solid) for* $\varepsilon = 0.05$. *Only the curve with* $C = -1/2$ *approximates the true solution, but any choice of* C *is valid in standard boundary layer theory.*

is determined by exponentially weak coupling between the boundaries, and this is missed by the standard boundary layer analysis. Many papers have been written on such effects, including the work mentioned earlier introducing what became known as *Ackerberg–O'Malley resonance* (or just *resonance*) [1]. More advanced techniques including WKBJ analysis can be brought into play, and four of these are outlined in [42]; see also section 2 of [49]. One of the key references is [39], and others include [20] and most of those given in the long list at the end of the introduction.

As usual, philosophical questions arise. We could say that boundary layer analysis has failed, since it has not detected the condition $C = -1/2$. On the other hand, we could say it has correctly identified the deeper "physics" of the problem. There would undoubtedly be applications in which each of these views would be appropriate.

We did not check the possibility of an interior layer at $x = 0$. For this problem there is no interior layer, ultimately because (4.2) is an exact solution of the ODE since the coefficients of the y' and y terms are x and 1. If these coefficients are perturbed, the behavior of solutions near $x = 0$ changes fundamentally, as we saw in Figure 3.2.

5. Dynamical Systems: Slow Manifolds. Differential equations and continuous time dynamical systems are fundamentally the same subject, but the fields have different emphases. In particular, dynamical systems favors analyses of problems based on *geometry*. Let us consider (1.1)–(1.2) from this point of view. An early geometric analysis of turning point problems was published by Kopell [34], and a general geometric theory for singular perturbation ODEs is due to Fenichel [12].

For a second-order ODE, the best known geometric tool is the use of the phase plane, with y on one axis and y' on the other. If the equation is autonomous (i.e., without explicit dependence on x), the trajectories of the phase plane give a complete description of the dynamics. Now (1.1) is nonautonomous, so this principle does not apply, but it is still interesting to plot (y, y') for our solutions, as shown in Figure 5.1. The images show the three stages of left boundary layer, solution in the interior, and right boundary layer. As $\varepsilon \to 0$, the corners get sharper, and the slopes and heights of the boundary layer segments increase at the rate $O(\varepsilon^{-1})$.

To reduce the problem fully to geometry, we need to introduce a new variable so that the equations are autonomous (though no longer linear). The standard trick is to take t as the independent variable and add a differential equation so that $t = x$. If we define $p = y'$, and use dots instead of primes to denote derivatives with respect to t, then (1.1) becomes the three-variable first-order autonomous system

$$\dot{x} = 1, \quad \dot{y} = p, \quad \dot{p} = \varepsilon^{-1}(xp - y), \tag{5.1}$$

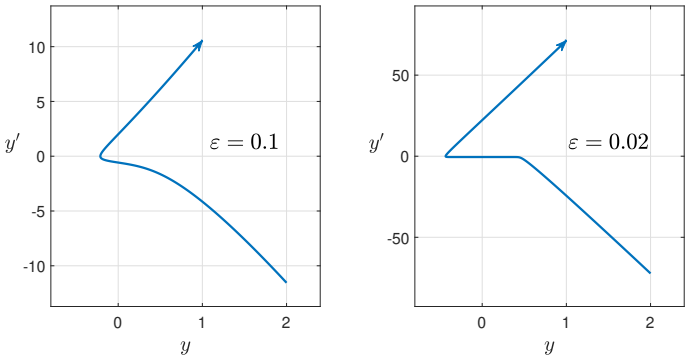

Fig. 5.1 *Repetition of Figure* 2.1 *with the trajectories plotted in the* (y, y') *plane. Since the equation is nonautonomous, this is not a true phase plane encapsulation of the dynamics, but still, it reveals key features. The trajectory begins at the bottom of the picture and quickly moves to the line* $y' \approx -1/2$ *(the left boundary layer). It moves steadily leftward along this line until suddenly leaving it and moving to the top of the picture (the right boundary layer).*

and (1.2) becomes the boundary conditions

$$y(-1) = 2, \quad y(1) = 1, \quad x(-1) = -1. \tag{5.2}$$

This is called a *slow-fast system,* since the variables evolve on different time scales. We have now reduced the problem fully to geometry in the sense that once we work out the field of trajectories in x-p-y space, we will know everything.

Figure 5.2 shows the solution trajectory for $\varepsilon = 0.02$ (the portion of it restricted to $-2 \le p \le 2$). The careful eye can see all the dynamics of the problem in this figure. A unique trajectory passes through each point in phase space, and the collection of these trajectories is the dynamical system. Because of the ε^{-1} factor in (5.1), the motion is fast in most regions. On the *critical manifold* defined by $xp = y$, however, the ε^{-1} term vanishes and the motion is slow. For $x < 0$, trajectories approach the critical manifold exponentially: it is *stable* or *attracting.* For $x > 0$, trajectories depart exponentially: it is *unstable* or *repelling.* Tikhonov, mentioned in the introduction, was one of the first to appreciate this fundamental distinction [56], which he made precise by one of the several results of his now known as Tikhonov's theorem [61, 62].

The exponential sensitivities of the problem can be understood in this framework. Mathematically, given the data (x, p, y) corresponding to any point t_0, the trajectory for all $t \in (-\infty, \infty)$ is uniquely determined. In particular, when and how a curve like the one shown in Figure 5.2 approaches the manifold for $x < 0$ exactly determines when and how it will leave it for $x > 0$. However, the determining information is exponentially squeezed along the way, since any trajectory that is near to the manifold is exponentially attracted to it and thus nearly restricted to a two-dimensional subspace. Thus, it is apparent simply as a matter of geometry that exponentially small perturbations may change the fate of a trajectory completely.

The critical manifold is an exact solution of (5.1), so any trajectory that touches the manifold is confined to it for all $t \in (-\infty, \infty)$. It follows that no trajectory can *cross* the critical manifold. The curves of Figure 3.1, however, must correspond to trajectories that do cross the manifold, since their boundary layers have opposite signs. This confirms from a new point of view that these curves must be erroneous.

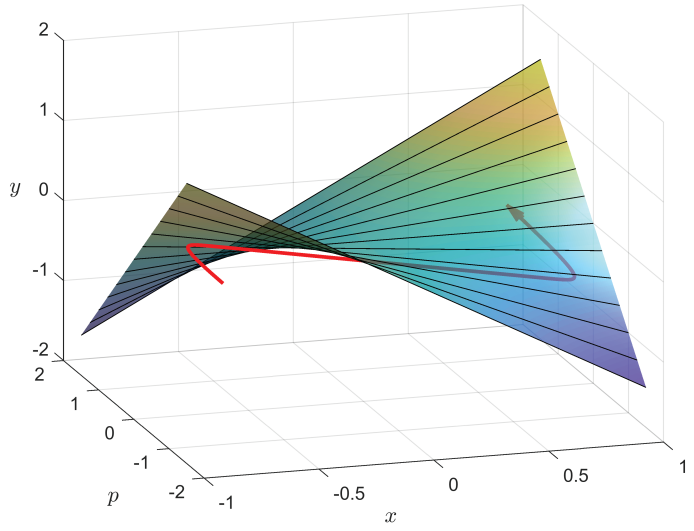

Fig. 5.2 *Solution trajectory with $\varepsilon = 0.02$ in three-dimensional x-p-y phase space. The surface is the critical manifold obtained by setting $\varepsilon = 0$, which for this problem can also be regarded as the slow manifold for $\varepsilon \ll 1$. The trajectory starts away from the manifold but quickly moves exponentially close to it. It then lies almost exactly along the straight line $y = px$ for a constant value of p (solid lines) until eventually zooming away again. In this image the trajectory has been artificially lowered a distance 0.08 to emphasize that it lies below the manifold.*

We have spoken of the critical manifold for (5.1), which is a geometric representation of the solutions of the outer equation (4.1) corresponding to $\varepsilon = 0$. For our problem, solutions of the outer equation also happen to be solutions of the full equation for any $\varepsilon > 0$. For general problems with multiple time scales, however, this is not the case, and one must be more careful. This is where the idea of a *slow manifold* comes in [12, 16, 40]. Roughly speaking, a slow manifold in an n-dimensional slow-fast system is an m-dimensional subset of the phase space, $m < n$, along which the motion is at speed $O(1)$ rather than $O(\varepsilon^{-1})$. Although it would be nice to know the shape of such a surface throughout n-space, more realistically one attempts to approximate it computationally over a finite range. The required definitions are nontrivial and contain some arbitrariness that sharpens up as $\varepsilon \to 0$. Algorithms for such problems have been developed by Krauskopf and Osinga and their collaborators and applied to applications in three and four dimensions, generally more complex than (1.1)–(1.2) [8, 21, 25, 44]. This group uses Doedel's software tool AUTO for continuation and bifurcation analysis as the computational engine [2, 9].

6. ODE Theory: Sturm–Liouville Operators. The theme of this article is that (1.1)–(1.2) is unusual, yet from one point of view it appears entirely standard: it is a *Sturm–Liouville problem*. Upon multiplying (1.1) by the integrating factor $e^{-x^2/2\varepsilon}$, we find that it is equivalent to the equation $\varepsilon(e^{-x^2/2\varepsilon}y')' + e^{-x^2/2\varepsilon}y = 0$ or, after a further multiplication by $e^{x^2/2\varepsilon}$,

$$\varepsilon e^{x^2/2\varepsilon}(e^{-x^2/2\varepsilon}y')' + y = 0. \tag{6.1}$$

This form amounts to a rewriting of the differential operator L associated with (1.1),

$$L\colon y \mapsto \varepsilon y'' - xy' + y, \tag{6.2}$$

in the equivalent form

$$L\colon y \mapsto \varepsilon w^{-1}(wy')' + y \tag{6.3}$$

with $w(x) = e^{-x^2/2\varepsilon}$. Sturm–Liouville theory tells us that L is a self-adjoint operator whose eigenvalues are a sequence of real, distinct numbers decreasing to $-\infty$, and that the corresponding eigenfunctions can be taken to be real (as opposed to complex) and form a complete orthogonal set, with the jth eigenfunction having $j-1$ simple zeros in $(-1,1)$. The idea is that any Sturm–Liouville operator has essentially the same behavior as the basic operator $y \mapsto y''$, whose eigenfunctions are sines and cosines.

All this sounds good, but when we look at exactly what the words mean, we find there is a catch of exponentially large magnitude. The operator L is self-adjoint not in the usual L^2 inner product, but in the weighted inner product defined by

$$\langle u, v\rangle = \int wuv, \tag{6.4}$$

where the integral is over $[-1,1]$.[3] Self-adjointness means that if u and v are smooth functions vanishing at $x = \pm 1$, then $\langle u, Lv\rangle = \langle Lu, v\rangle$. To verify this condition, we compute $\langle u, Lv\rangle = \int wu(\varepsilon w^{-1}(wv')') + wuv = \int \varepsilon u(wv')' + wuv$. Integration by parts gives $\langle u, Lv\rangle = -\int \varepsilon u'(wv') + wuv = -\int \varepsilon (wu')v' + wuv$, and a second integration by parts gives $\langle u, Lv\rangle = \int \varepsilon (wu')'v + wuv = \int w(Lu)v = \langle Lu, v\rangle$.

For small ε, the inner product (6.4) is unrecognizably different from the ordinary inner product of $L^2([-1,1])$. With $\varepsilon = 0.01$, for example, $w(x)$ takes the value 1 at $x = 0$ and $e^{-50} \approx 10^{-22}$ at $x = \pm 1$. Such enormous disparities in magnitude imply that in measuring the inner product of two functions u and v, (6.4) essentially doesn't care about their behavior near $x = \pm 1$.

We can see the consequences by looking at eigenfunctions, a central topic of Sturm–Liouville theory. The eigenvalue equation for L is

$$\varepsilon y'' - xy' + y = \lambda y, \quad y(-1) = y(1) = 0, \tag{6.5}$$

and if this is satisfied for a nonzero function y, then λ is an *eigenvalue* of L and y is a corresponding *eigenfunction*. The theory asserts that eigenfunctions y_j and y_k corresponding to distinct eigenvalues are orthogonal in the inner product (6.4),

$$\int_{-1}^{1} e^{-x^2/2\varepsilon} y_j(x) y_k(x)\, dx = 0, \quad j \neq k. \tag{6.6}$$

Figure 6.1 shows the first four eigenfunctions for $\varepsilon = 0.2$ and $\varepsilon = 0.02$. The first set look something like the sines and cosines one might expect for a Sturm–Liouville problem, but the second set are completely different, approximating the monomials 1, x, x^2, and x^3. The monomials are famously nonorthogonal in the L^2 inner product on $[-1,1]$, with the condition number of the set $1, x, \ldots, x^n$ growing exponentially as

[3] In linear algebra, the analogue would be a matrix A that is not symmetric but can be symmetrized by a diagonal similarity transformation, $B = DAD^{-1}$. Since all the effects at issue in this paper are real, we dispense with complex conjugate bars in defining inner products.

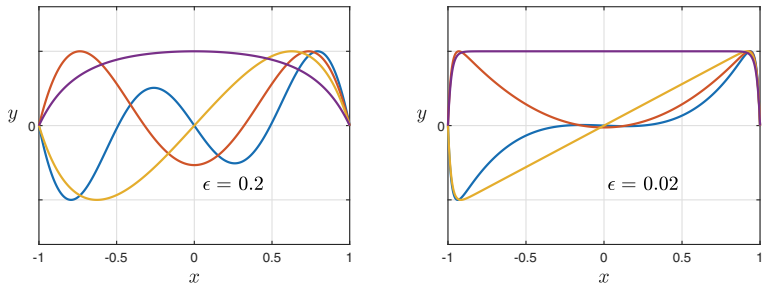

Fig. 6.1 *The first four eigenfunctions of the operator L of* (6.2)*, orthogonal with respect to the inner product* (6.4)*. With large ε, on the left, the inner product is reasonably close to that of $L^2([-1,1])$ and the eigenfunctions show the familiar oscillations of sines and cosines. With small ε, on the right, the inner product effectively ignores behavior near $x = \pm 1$, and the eigenfunctions approximate the monomials $1, x, x^2, x^3$ except in the boundary layers. The associated eigenvalues approximate the numbers* 1*,* 0*,* -1*, and* -2 *to about* 10*,* 9*,* 7*, and* 6 *digits, respectively.*

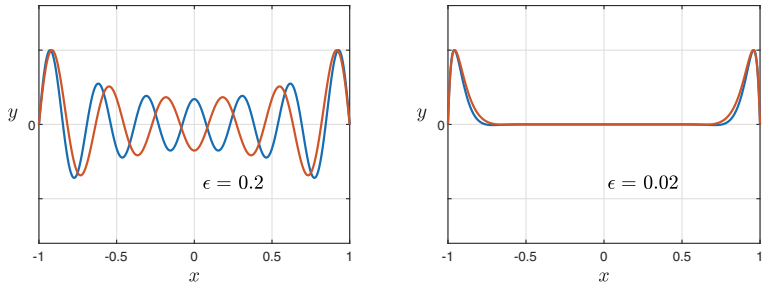

Fig. 6.2 *Repetition of Figure* 6.1 *for eigenfunctions* 11 *and* 13 *(i.e., even eigenfunctions* 6 *and* 7*). With $\varepsilon = 0.2$ we recognize the eigenfunctions as relatives of sines and cosines, but with $\varepsilon = 0.02$ the shape is very different and the eigenfunctions are obviously far from orthogonal in L^2.*

$n \to \infty$ at a rate of approximately $(1+\sqrt{2})^n$ [15].[4] For example, the set $\{1, x, \ldots, x^{30}\}$ has condition number about 1.1×10^{11}. If one examines the first $n+1$ eigenfunctions of L for various ε, one finds condition numbers of this magnitude for $\varepsilon \ll n^{-1}$. For example, the first 31 eigenfunctions of L have condition number about 8×10^4 for $\varepsilon = 0.02$, 3×10^8 for $\varepsilon = 0.01$, and between 10^{10} and 10^{11} for $\varepsilon = 0.005$. In the presence of such ill-conditioning, it would be hard to argue that eigenfunction expansions are the right tool for the investigation of (1.1).

Figure 6.2 repeats Figure 6.1 but for eigenfunctions 11 and 13, that is, even eigenfunctions 6 and 7. The image for $\varepsilon = 0.02$ highlights how far from orthogonal these functions are in L^2 (their L^2 inner product is about 0.9924). Higher eigenfunctions continue the pattern of having almost all their amplitude in the boundary layers.

[4] Another reflection of the nonorthogonality of monomials is the *Müntz–Szasz theorem*, which asserts that a continuous function on $[0,1]$ can be approximated arbitrarily closely by linear combinations of the functions $1, x^{\alpha_1}, x^{\alpha_2}, \ldots$ for any choice of exponents satisfying $\sum \alpha_k^{-1} = \infty$ [52].

We can explain as follows why small values of ε lead to eigenfunctions approximating monomials. For small ε, the inner product (6.4) essentially ignores the boundary conditions, and we end up approximately with polynomials orthogonal with respect to $e^{-x^2/2\varepsilon}$ over $[-1,1]$, or equivalently, with respect to e^{-x^2} over $[-(2\varepsilon)^{-1/2},(2\varepsilon)^{-1/2}]$. These are essentially the Hermite polynomials $H_0, H_1, \ldots$, which closely approximate the monomials if they are rescaled from $[-(2\varepsilon)^{-1/2},(2\varepsilon)^{-1/2}]$ back to $[-1,1]$ with magnitude 1. The corresponding eigenvalues are close to $1, 0, -1, -2, \ldots$.

7. Spectral Theory: Eigenvalues and Pseudospectra. We have just noted that Sturm–Liouville theory ensures that the eigenvalues of L are real and distinct, decreasing toward $-\infty$. We now consider the significance of these eigenvalues.

As we saw in Figure 6.1, for small ε, the functions $1, x, x^2, \ldots$ approximately satisfy the eigenvalue equation (6.5) except in the boundary layer. This is readily explained by boundary layer analysis, where following (4.1), we find the outer equation for the eigenvalue problem, $-xy' + y = \lambda y$, that is, $y'/y = (1-\lambda)/x$. This equation is satisfied by the monomial $y = x^{1-\lambda}$ for any integer $\lambda \le 1$.

One of the eigenvalues of L is particularly important: the one that is exponentially close to zero and becomes exactly zero in the limit $\varepsilon = 0$, which we shall call λ_0. For $\varepsilon = 0.1$, 0.05, and 0.02 we have $\lambda_0 \approx -0.13$, -0.0027, and -3.7×10^{-9}, respectively, and computations guided by the estimates of [20] suggest that for $\varepsilon \to 0$ the behavior is

$$\lambda_0 \sim C\varepsilon^{-3/2} e^{-1/2\varepsilon}, \quad C \approx -0.8. \tag{7.1}$$

Now when an operator has a small eigenvalue, that implies that its inverse is large in norm. In particular, if L^{-1} denotes the solution operator for the problem $Lu = f$ with boundary conditions $y(\pm 1) = 0$, then we have

$$\|L^{-1}\| \ge |\lambda_0|^{-1}. \tag{7.2}$$

Thus we see that although L is mathematically nonsingular for all small ε, its inverse will be huge. The solution y of $Ly = f$ for almost any choice of f will be close to a multiple of the eigenfunction y_0 associated with λ_0, which has the shape of a sawtooth, as can be seen in the second image of Figure 6.1. Figure 7.1 shows this function again for $\varepsilon = 0.03$ and $\varepsilon = 0.01$. We call y_0 a *pseudonull function,* since it very nearly satisfies the condition $Ly = 0$. It is y_0 that explains why the curves of the first four figures of this paper all have approximately the same form. Figure 7.2 shows the effect explicitly by plotting a smooth random function f produced by the Chebfun `randnfun` command together with the corresponding solution of $Ly = f$ with zero boundary conditions.

What about the other eigenvalues and eigenfunctions of L? Sturm–Liouville analysis suggests that these should be important for understanding the behavior of the operator. We have already argued via ill-conditioning of the set of eigenfunctions, however, that this expectation is mistaken when ε is small. We shall now derive the same conclusion via the theory of pseudospectra.

An operator with a complete set of orthogonal eigenfunctions is said to be *normal.* (The definition generalizes to other classes of operators, but we won't need this.) For small ε, L is far from this situation, and we say that it is *highly nonnormal.* The *resolvent* of L is the operator $(\lambda I - L)^{-1}$ depending on a complex variable λ, and to assess nonnormality we look at its 2-norm $\|(\lambda I - L)^{-1}\|$ as a function of λ. Let $\sigma(L)$ denote the spectrum of L, which for our operator is simply the set of eigenvalues. If

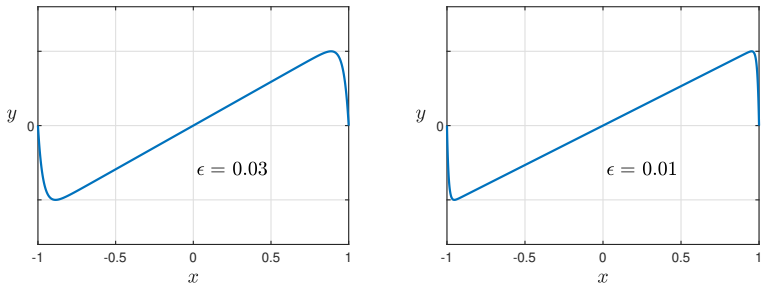

Fig. 7.1 *Pseudonull functions of L for two values of ε. These are the eigenfunctions y_0 of L associated with the eigenvalue λ_0 that falls very close to zero. Solutions to $Ly = f$ will tend to have approximately this shape, with very large amplitude, except for special choices of f.*

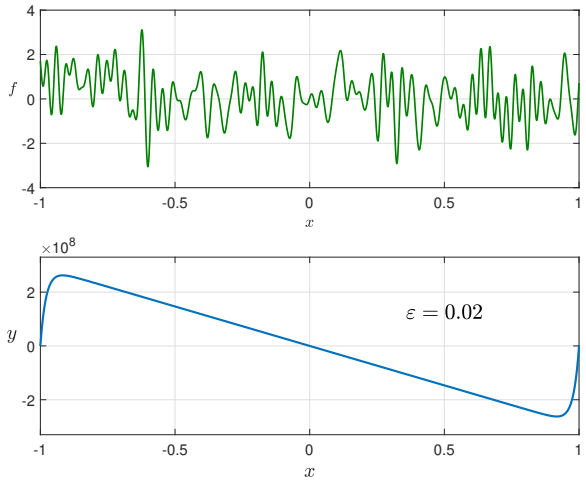

Fig. 7.2 *Above: the smooth random function f computed by the Chebfun commands* `rng(0), f = randnfun(.03)`. *Below: the solution y of the problem $Ly = f$ with $\varepsilon = 0.02$. Despite the complicated forcing function, we see that y closely approximates a large multiple of the eigenfunction y_0 of L.*

$\lambda \notin \sigma(L)$, then $\|(\lambda I - L)^{-1}\|$ is a positive number that diverges to ∞ as λ approaches $\sigma(L)$. For any $\delta > 0$, we define the *δ-pseudospectrum of L* to be the complex set

$$\sigma_\delta(L) = \{\lambda \in \mathbf{C} : \|(\lambda I - L)^{-1}\| > \delta^{-1}\}, \tag{7.3}$$

with the convention $\|(\lambda I - L)^{-1}\| = \infty$ if $\lambda \in \sigma(L)$. Full details and many examples are presented in [60].

For a normal operator, $\sigma_\delta(L)$ is the union of the open δ-balls around the points of the spectrum. For a nonnormal operator, however, it may be much bigger. The norm $\|(\lambda I - L)^{-1}\|$ may be much greater than the reciprocal of the distance of λ to $\sigma(L)$, and indeed, $\|(\lambda I - L)^{-1}\|$ may take huge values at points λ that are nowhere near $\sigma(L)$. Figure 7.3 shows pseudospectra of L for $\varepsilon = 0.1$ and 0.01, confirming that for small ε, the operator is highly nonnormal.

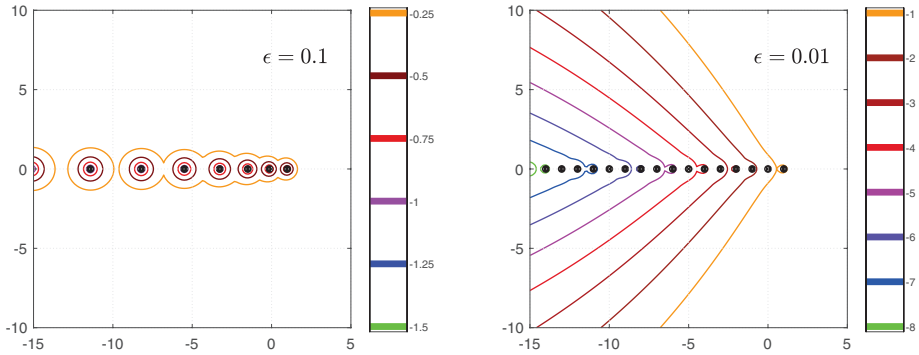

Fig. 7.3 *Spectra and δ-pseudospectra in the complex plane of the operator L as computed by EigTool* [67]*; the level curves correspond to $\delta = 10^a$ for the values of a indicated by the colorbars. On the left, for large ε, L is close to normal and the pseudospectra approximate balls around the eigenvalues. On the right, for small ε, L is highly nonnormal and the pseudospectra have little to do with the eigenvalues. These curves tell us that although the rightmost eigenvalues of L may have some physical significance, the eigenvalues further left have little. Instead, L will behave as if its spectrum included large continuous regions of the left half-plane. The corresponding pseudoeigenfunctions have the shape of wave packets localized at the boundaries.*

To see the implications, for concreteness, consider the 10^{-6}-pseudospectrum of L for $\varepsilon = 0.01$, shown in Figure 7.3. This is the region in the left half λ-plane bounded by the purple-blue line, an infinite triangular sector extending along the negative real axis to about $\lambda = -8$. Each point λ in this sector is a 10^{-6}-pseudoeigenvalue of L, which means that for each such point, there is a pseudoeigenfunction y that satisfies (6.5) with a 2-norm error of less than 10^{-6}. In a physical experiment, unless it is accurate to six digits of precision, such a value would likely be indistinguishable from a mathematically true eigenvalue. This is the theme of the study of nonnormality: if an operator is far from normal, this doesn't just mean that its eigenvalues are highly sensitive to perturbations, it means its eigenvalues and eigenfunctions lose their significance.

The pseudoeigenfunctions of L are shaped like wave packets localized at the boundary, like the eigenfunctions themselves as seen, for example, in the second panel of Figure 6.2. Similarly, the adjoint operator L^* we are about to consider has pseudoeigenfunctions in the form of wave packets localized at $x = 0$. A theory that explains such structures is presented in [58] and [60, Chapter 11], where a key role is played by the so-called *twist condition* (for L^*) and *antitwist condition* (for L); the twist condition is a special case of the *commutator condition* introduced by Hörmander in connection with the subject of section 10 [30, 68]. These conditions are stated in terms of a variable coefficient symbol function $\varphi(x, k)$ describing the local behavior of the operators when applied to a wave $\exp(iks/\varepsilon)$ for small ε and $s \approx x$, all part of the circle of ideas known in different communities as WKBJ theory, semiclassical analysis, and microlocal analysis [69].

8. Sensitivity Analysis: Adjoints and SVD. We have seen that for small ε, the problem (1.1)–(1.2) is highly sensitive to perturbations. The study of how outputs of models depend on perturbations of the inputs is called *sensitivity analysis,* and when

differential equations are involved, one of the important tools is the use of adjoints. Two stimulating publications in this area are [17] and [22].

For our problem the linear operator is

$$L : y \mapsto \varepsilon y'' - xy' + y, \tag{8.1}$$

which in this section we take to act on smooth functions with $y(\pm 1) = 0$. The *adjoint* of L is the operator

$$L^* : v \mapsto \varepsilon v'' + xv' + 2v \tag{8.2}$$

acting on the same functions. The standard characterization of L^* is the property $(v, Ly) = (L^*v, y)$, where $(u, v) = \int_{-1}^{1} u(x)v(x)dx$, but to understand why adjoints are important, it may be more enlightening to write an equivalent condition in terms of the inverse operators L^{-1} and L^{-*} acting on smooth functions f and g (with no boundary conditions). Setting $Ly = f$ and $L^*v = g$ converts the characterization to

$$(L^{-*}g, f) = (g, L^{-1}f), \tag{8.3}$$

and we can interpret this equation as follows. The number $(g, L^{-1}f)$ can be thought of as a scalar measurement, via the inner product with a given function g, of the solution y of the differential equation $Ly = f$. Equation (8.3) asserts that if we know the function $v = L^{-*}g$, we can obtain the same number simply as the inner product of v and f, without solving $Ly = f$. If there are a thousand right-hand sides f_j of interest, then all we need is a thousand inner products (v, f_j), not a thousand solutions $Ly_j = f_j$. This is the starting point of adjoint-based sensitivity analysis. For linear problems, small perturbations behave in the same way as large ones, but for nonlinear problems, the adjoint is defined via linearization and captures the relationship between infinitesimal perturbations of f and their effects on the measured output.

Now let us look at the behavior of the adjoint (8.2). The crucial difference from (8.1) is the change of sign from $-xy'$ to xy', and the result is that instead of boundary layers at each end of the interval, we get wave packets in the middle. To explore this effect, Figure 8.1 recapitulates many of the figures of this paper, but based on (8.2) instead of (8.1). We will not repeat the analyses of the past sections, but in each case the same themes reappear with interesting variations. For example, the outer equation (4.1) now becomes $y'/y = -2/x$, with solution $y = C/x^2$. Obviously this solution cannot be valid near $x = 0$, where an interior layer analysis is required; nor is it exactly valid anywhere when ε is nonzero. On the other hand, the signs are such that boundary layers are precluded at both $x = -1$ and $x = 1$. Whereas before, boundary layer analysis gave us an approximate solution (4.6) with too much freedom, an extra undetermined parameter, now it gives too little freedom, not enough parameters, and again a deeper analysis would be required to match the true solution.

To see the implications of Figure 8.1 for sensitivity analysis, perhaps the best image to focus on is the curve in the middle on the right, corresponding to the solution of a problem $L^*v = g$ with a smooth random function g. We can think of g as defining an arbitrary linear functional one might choose to measure solutions of problems $Ly = f$. The amplitude $O(10^9)$ of this curve for $x \approx 0$ shows that the influence of function values of f for $x \approx 0$ will be amplified by $O(10^9)$ in their effects on $L^{-1}f$ as measured by g. By contrast, the values of f for $x \approx \pm 1$ have much smaller effect.

Figure 8.1 does not include an image corresponding to Figure 7.3, because the eigenvalues and pseudospectra of any operator are the same as those of its adjoint. It

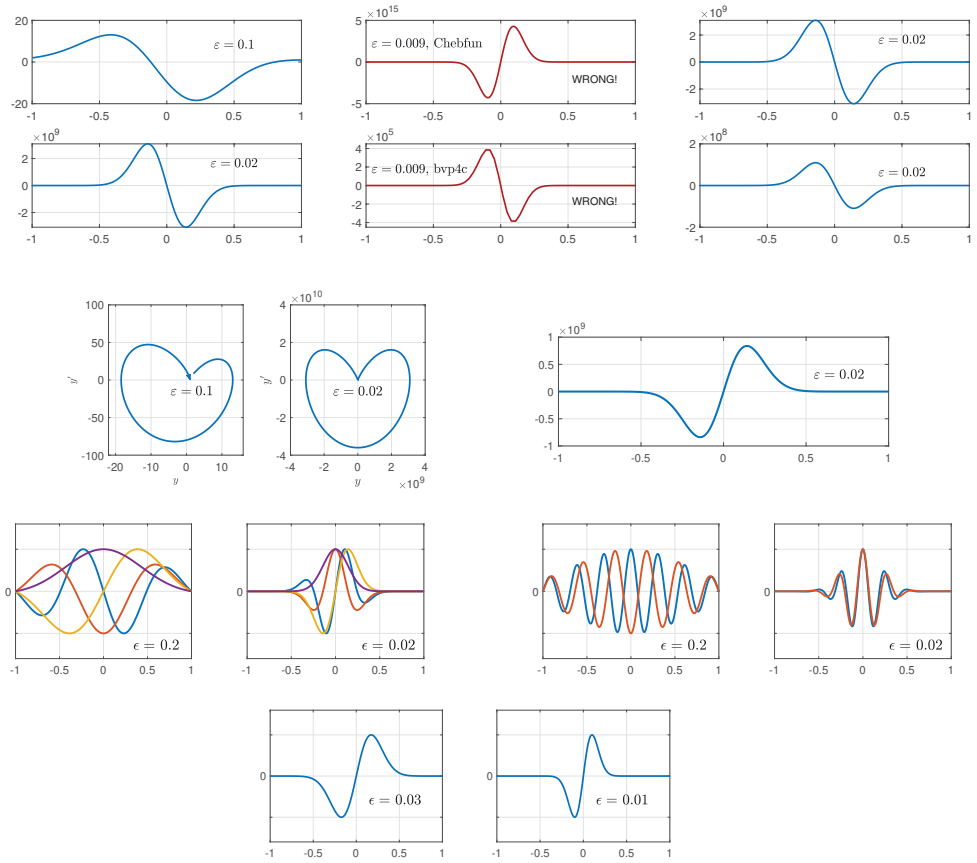

Fig. 8.1 *Repetition of earlier figures with* (8.1) *replaced by its adjoint* (8.2). *The top double row repeats Figures* 2.1 *(true solutions),* 3.1 *(erroneous solutions with* $\varepsilon = 0.09$*), and* 3.2 *(perturbed problems). Instead of boundary layers at* ± 1*, we see waves at* $x = 0$*. Note the huge amplitudes in these and other images. (The "WRONG!" solutions have the right shape but the wrong amplitudes.) The second row repeats Figures* 5.1 *(*(y, y') *plane) and* 7.2 *(solution to* $L^* v = g$ *for a smooth random function* g*). The third row repeats Figures* 6.1 *and* 6.2*, showing eigenfunctions as wave packets near* $x = 0$*. The final row repeats Figure* 7.1 *(null functions).*

is the eigenfunctions and pseudoeigenfunctions that differ, changing from structures localized near ± 1 for L to structures localized near $x = 0$ for L^*. In numerical linear algebra such effects are familiar as the difference between the left and the right eigenvectors of a matrix. On the other hand, an analogue of Figure 5.2 can be found in Figure 8.2. Note that the scale is on the order of 10^{10}. The geometry is more complicated than before, because the critical manifold is no longer an invariant manifold for $\varepsilon > 0$.

Along with adjoints, the title of this section mentions the *SVD,* i.e., the *singular value decomposition.* This is the appropriate tool if one seeks the function f with $\|f\| = 1$ (2-norm) that maximizes $\|L^{-1} f\|$. Such a function f is the minimal left singular function of L, and we can write $L^{-1} f = g/\sigma_{\min}$, where g is the corresponding right singular function and $\sigma_{\min}$ the associated minimal singular value. With the

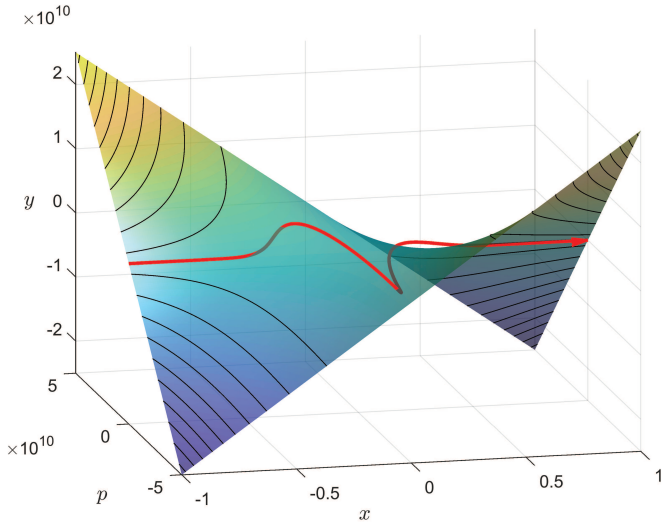

Fig. 8.2 *Analogue of Figure 5.2 for the adjoint operator L^*. In this case the outer solutions $y(x) = C/x^2$ for $\varepsilon = 0$ are not exact solutions of the differential equation with $\varepsilon > 0$, so the critical manifold $y = -px/2$ plotted is not exactly a slow manifold. The region $x < 0$ is unstable, and the solution zooms away from this sheet to amplitude $O(10^{10})$ near $x \approx 0$, then returns in the stable part of the manifold, $x > 0$, to match the boundary condition at $x = 1$. Note the huge scales of the p and y axes needed to capture this trajectory.*

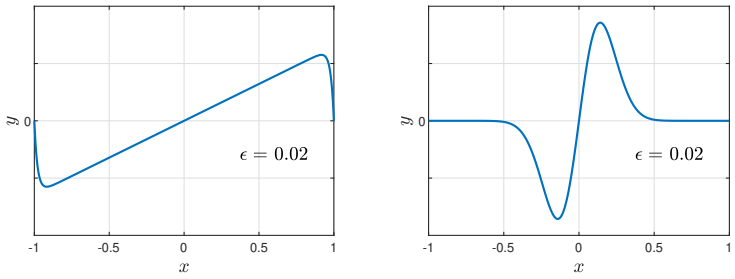

Fig. 8.3 *Left and right minimal singular functions of L for $\varepsilon = 0.02$, with corresponding singular value $\sigma_{\min} \approx 7 \times 10^{-10}$.*

SVD, the inequality (7.2) improves to an equality,

$$\|L^{-1}\| = \sigma_{\min}^{-1}. \tag{8.4}$$

Figure 8.3 shows minimal singular functions for $\varepsilon = 0.02$. Because $\sigma_{\min}$ is so close to 0, they are indistinguishable in a plot from the pseudonull functions of Figures 7.1 and 8.1.

9. Physics: Ghost Solutions. Domokos and Holmes published a fascinating paper in 2003, "On Nonlinear Boundary-Value Problems: Ghosts, Parasites and Discretizations" [10]. They pointed out that sometimes a physical experiment may generate results that are far from the exact solution of the associated mathematical model for reasons of just the kind we have been examining: though far from the true solution, the observed results nearly satisfy the ODE BVP.[5]

[5] *Shadowing* in dynamical systems is a related idea [23].

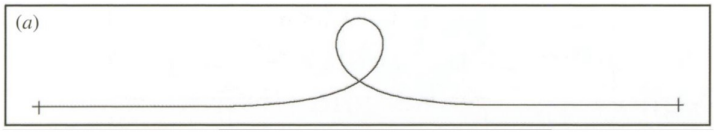

Fig. 9.1 *A loop in an elastica with loads at both ends as sketched in Figure* 1 *of Domokos and Holmes* [10]. *This corresponds to a locally unique solution of an ODE BVP. In the laboratory, however, the loop is observed to be equally stable if it is translated a certain distance right or left of center. Such shapes are pseudosolutions of the equation, or "ghosts." A mathematically equivalent example involves a rigid pendulum starting nearly vertically above its pivot, waiting nearly motionless for a while, then suddenly swinging around to nearly vertical again. (Image courtesy of the Royal Society of London.)*

This possibility is implicit in the curves of Figure 8.1 corresponding to the adjoint equation (1.3). Suppose, for example, our problem is $\varepsilon y'' + xy' + 2y = 0$ with $y(0) = y(1) = 0$. The mathematically correct unique solution is the function $y = 0$, but if ε is small, the pseudonull function of the last row of Figure 8.1 will satisfy the boundary conditions exactly and the ODE nearly exactly, with error of order 10^{-9} for $\varepsilon = 0.02$. Other similar curves would satisfy the ODE exactly and the boundary conditions nearly (this is the combination Domokos and Holmes focus on). All these nearly-solutions can be called "ghosts." In an experiment conducted in a laboratory with its inevitable imperfections, it is hard to see why the exact solution should be more likely to be observed than the pseudosolutions.

Domokos and Holmes illustrate their idea with a compelling example (see Figure 9.1). Suppose a long flexible rod, known as an *elastica,* is loaded at both ends. Equations of elasticity going back to Euler model this by the nonlinear ODE $\varepsilon y'' = \sin(y)$ with $y'(\pm 1) = 0$, which, for small enough ε, has a locally unique solution in the form of a loop in the center of the domain. As ε decreases, however (think of the elastica getting longer), the boundary coupling that constrains the loop to lie exactly in the middle grows exponentially weaker. In practice, one can move the loop some distance to either side and the elastica will be observed to remain stable. These off-center pseudo-solutions—ghosts—are as valid experimentally as the true solutions. An alternative interpretation of the same equation is that we have a rigid pendulum on a long time interval, beginning nearly vertical above its pivot point with speed zero. For a long time the weight may move imperceptibly until finally it gathers speed and swings over once up to the equal and opposite nearly vertical position. Although in exact mathematics the swing must occur in the middle of the time interval, an earlier or later swing fails the boundary conditions by only an exponentially small amount.

As discussed in [58], ghosts can appear with both linear and nonlinear equations, and are a special case of the phenomenon of pseudoeigenfunctions. Our equation (1.3) is linear, which means that no particular amplitude for an observed ghost is selected. Nonlinearity gives extra interest to the Domokos–Holmes example, because it fixes the size of the ghosts as well as their shape.

Notions of ghosts in dynamical systems have appeared in other contexts too. Another fascinating one concerns the snap-through instability exhibited by (among other mechanisms) the Venus flytrap and the "hopper popper" toy [18, 55].

10. PDE Theory: Lewy Nonexistence. Like any ODE BVP, (1.1)–(1.2) can be interpreted as the steady-state equation for a time-dependent PDE. If $u(x,t)$ evolves

according to the PDE BVP

$$u_t = \varepsilon u_{xx} - x u_x + u, \quad u(-1,t) = 2, \quad u(1,t) = 1, \tag{10.1}$$

and approaches a steady solution $y(x)$ as $t \to \infty$, then y will satisfy (1.1)–(1.2). This PDE is an advection-diffusion equation with a variable coefficient, and the exponentially weak coupling between the boundaries at $x = \pm 1$ can be interpreted as resulting from the fact that the advection is leftward in $[-1, 0]$ and rightward in $[0, 1]$, so that information can travel between the boundaries only by diffusing against the flow. Indeed, this sweeping of information toward the boundaries is the physical explanation of why solutions to most equations close to (1.1) are nearly zero in the interior; Ackerberg–O'Malley resonance alludes to the exceptional circumstances needed for the interior values to be significantly nonzero. Further examples of ODEs as steady-state equations of time-dependent PDEs are presented in [59, Chapter 22]; see also [60, Chapter 12]. Sometimes a steady solution is unstable with respect to perturbations and thus hardly likely to appear as $t \to \infty$, and that is the case here because of the y term in (1.1). Time-dependent simulations actually show not convergence to a steady state, but rapid formation of boundary layers at both ends combined with overall exponential growth.

Here in this final section, however, we want to mention a different connection of (1.1)–(1.2) with PDE theory that has more of a functional analysis flavor. In 1957 Hans Lewy caused a stir among mathematicians by showing that a linear PDE with smooth coefficients and right-hand side (C^∞, though not analytic) may fail to have a solution, even locally [38]. His discovery has been extended in many ways in the years since then by a *Who's Who* of eminent mathematicians including Hörmander, C. Fefferman, Nirenberg, Treves, Sjöstrand, and Dencker [30, 68]. It turns out that we can interpret this phenomenon in terms of the pseudospectra of exponentially nonnormal operators related to (1.3). This discussion follows Chapter 13 of [60], which was sparked by Zworski's paper [68].

First let us make an observation about the dual BVPs (1.1) and (1.3), both with boundary conditions $y(-1) = a$ and $y(1) = b$. For small ε, each problem is well-posed in the sense that there exists a unique solution that depends continuously on the data a and b. However, the condition numbers are exponentially large, and the two problems come close to being ill-posed in interesting dual ways. With (1.1), we have a phenomenon of "near-nonuniqueness," since as we have seen in (4.6), there is a one-dimensional space of functions that come exponentially close to satisfying the ODE and the boundary conditions. With (1.3), on the other hand, there is a phenomenon of "near-nonexistence." The issue here is that unless $a = b$, the solution y will be huge, diverging rapidly to infinity as $\varepsilon \to 0$. For example, with $\varepsilon = 0.02$ we get $\max |y(x)| \approx 46.9$ if $a = b = 1$, but $\max |y(x)| = 3.1 \times 10^6$ if $a = 1$ and $b = 0.999$. With $\varepsilon = 0.01$, the solution with these boundary data is bigger than 10^{17}. These BVPs involving L and L^* for small ε accordingly have the flavor of Fredholm operators of index $+1$ and -1, respectively [4, 31].

Without giving details, we now outline how related phenomena are connected with Lewy nonexistence. As the simplest possible example, consider the *Mizohata equation* [43]

$$\frac{\partial u}{\partial x} + ix \frac{\partial u}{\partial y} = f(x, y), \tag{10.2}$$

which we regard as a PDE in two complex variables x and y. The coefficient ix is analytic, but we shall suppose that f is C^∞ but not analytic. We now take the Fourier

transform of the adjoint of (10.2) with respect to y to obtain the equation

$$\frac{d\hat{u}}{dx} + kx\hat{u} = \hat{f}(x,k), \tag{10.3}$$

or equivalently, setting $\varepsilon = k^{-1}$ and $\hat{g}(x,\varepsilon) = \varepsilon\hat{f}(x,k)$,

$$\varepsilon\frac{d\hat{u}}{dx} + x\hat{u} = \hat{g}(x,\varepsilon). \tag{10.4}$$

Instead of a PDE, we now have a family of ODEs with the parameter ε, and each one has a turning point at $x = 0$ of the same type as (1.3). Solutions to such equations will be problematic when ε is small. For any individual choice of ε, there is a unique solution, but by constructing a function $\hat{g}$ with a suitable dependence on ε, we can contrive things so that g is C^∞ yet the adjoint of (10.4) has a nonunique solution. By the Fredholm alternative, this implies nonexistence for the problem (10.2) itself. Clearly we are omitting details of this argument!—see [60, Chapter 13].

In a word, thanks to the Fourier transform, a linear PDE can be thought of as a parametrized linear ODE with all parameter values present at once. That is why unbounded ill-conditioning for the ODEs translates into ill-posedness of the PDE.

11. Conclusion. Mathematics comprises so many lands; none of us knows them all. I hope the reader has enjoyed the adventure of taking $\varepsilon y'' - xy' + y = 0$ as a calling card on a mathematical world tour.

"Eight perspectives" may sound a bit inflated, but in fact, in the course of my working out these connections, quite a few more perspectives also turned up and made claims to attention. We have not discussed Green's functions [54], which take on surprising shapes in cases of extreme non-self-adjointness [60, Chapter 12]. We have not discussed the theory of exponential dichotomy, an established tool for quantifying exponential effects in ODEs [3, 6, 24, 50]. We have not discussed stochastic effects, an area of ever-growing importance in mathematics and science [28, 42], nor the mechanism of resonance in turning point problems [53], nor (hardly) any physical applications; on all these matters, see the recent paper by Matkowsky [42]. Numerical linear algebra analogues have been around every corner but only hinted at here. As for the theory of dynamical systems, the brief presentation of section 5 barely begins to do justice to this flourishing field. Readers of this article may notice further relevant areas of mathematics that seem to have been inexplicably omitted.

Equation (1.1) has remarkable properties, but it is not unique in this respect. Slight changes in the equation lead to qualitative changes in the solutions (jumps, oscillations, corners, near-zero regions, ... [26, 51, 64]), and lessons could have been drawn from variants like these too. But I have resisted all temptation to generality, resolutely keeping the focus on just one beautiful equation and its adjoint.

Appendix. Notes on Numerical Computations. The numbers and figures presented in this paper were computed with MATLAB and Chebfun [11, 59]. Here are a few details for readers who may wish to try their own experiments.

2. *Existence, uniqueness, and exact solution.* Chebfun users can explore numerical solutions of (1.1)–(1.2) and its variants with code like this:

```
L = chebop(@(x,y) ep*diff(y,2)-x*diff(y)+y);
L.lbc = 2; L.rbc = 1; y = L\0; plot(y)
```

Experiments indicate that results for values of `ep` as low as $\varepsilon = 0.02$ are trustworthy. The structure `L` is called a chebop, and the solution `y` is a chebfun—a finite Chebyshev series, that is, a polynomial represented in the Chebyshev basis. The polynomials of Figure 2.1 have degrees 36 and 64, respectively.

Another Chebfun method for exploring ODEs (BVPs, initial-value problems, and eigenvalue problems) is the graphical user interface `chebgui`, which executes the appropriate Chebfun code segments automatically. I relied on `chebgui` at every step to make sure I was getting things right.

The following code segment generates a chebfun for y_{even} on $[0, 1]$:

```
x = chebfun('x',[0 1]); ex = exp(x^2/(2*ep));
yeven = -ex + (x/ep)*cumsum(ex);
```

3. *Backward error analysis and ill-conditioning.* Concerning nonzero right-hand sides in $Lu = f$, it is interesting to try smooth random functions [13], like this:

```
f = randnfun(0.1); y = L\f;
```

For $\varepsilon = 0.02$, the resulting solutions have slopes on the scale of 10^8; see Figure 7.2. It is also interesting to break the symmetry of (1.1)–(1.2), e.g., by changing $[-1, 1]$ to $[-1, 0.9]$, in which case a boundary layer (assuming $\varepsilon \ll 0.1$) appears at just one end.

5. *Slow manifolds.* To draw a phase plane image like Figure 5.1, one can compute the solution in the form of a chebfun `y` and then execute `plot(y,diff(y))`. An arrow can be put at the end with the command `arrowplot(y,diff(y))`, using the `'ystretch'` parameter to adjust the orientation. Figure 5.2 was plotted with the Chebfun2 `surf` command [11, Chapter 14]), setting `alpha 0.7` in MATLAB to give the surface some transparency. More dramatic and detailed images, computed with the BVP solver of AUTO, can be found in the works of Krauskopf and Osinga and their collaborators [8, 21, 25, 44].

6. *Sturm–Liouville operators.* To compute the first k eigenvalues and eigenfunctions of an operator L in Chebfun, one can represent it in the usual manner as a chebop `L` and then execute `[V,D] = eigs(L,k)`. The object V is an "$\infty \times k$" Chebfun *quasi-matrix,* i.e., an object with k columns each of which is a function defined on $[-1, 1]$. Details of such "continuous linear algebra" can be found in [5]. The condition number of the set $\{1, x, \ldots, x^{30}\}$ on $[-1, 1]$ can also be computed by constructing a quasi-matrix, now with 31 columns: `x = chebfun('x'); A = vander(x,31); cond(A)`. The `cond` command computes the SVD of the quasi-matrix and returns the ratio of its first and final singular values.

7. *Eigenvalues and pseudospectra.* Figure 7.3 was produced by EigTool, the standard tool for computing pseudospectra, due to Wright [67, 60]. As input, EigTool requires a matrix approximating the differential operator L with zero Dirichlet boundary conditions, for which we used the 100×100 Chebyshev spectral discretization matrix generated by the Chebfun commands

```
n = 100; D = diffmat(n); D2 = diffmat(n,2);
x = chebpts(n); X = diag(x); I = eye(n);
ep = .1; L = ep*D2 - X*D + I; L = L(2:n-1,2:n-1);
W = diag((1-x(2:n-1)).^(1/4)); L = W*L/W;
```

The need for the weighting matrix W is explained on p. 411 of [60].

8. *Adjoints and SVD.* If `L` is a chebop, then `adjoint(L)` or `L'` is the chebop corresponding to its adjoint. The first k singular values and functions can be computed with `[U,S,V] = svds(L,k)`.

10. *Lewy nonexistence.* Movies of solutions of one-dimensional time-dependent PDEs such as (10.1) can be computed by `chebgui`, which calls the Chebfun code `pde15s`.

Acknowledgments. The impulse to write this paper came from a visit to New Zealand as Michael Erceg Senior Visiting Fellow in November–December 2017, made possible by generous support from the Margaret and John Kalman Charitable Trust. My hosts Bernd Krauskopf and Hinke Osinga and their dynamical systems group provided an extraordinarily stimulating environment. In addition I am grateful for comments on drafts of this article to Uri Ascher, Andrei Bogatyrev, Folkmar Bornemann, Jon Chapman, Gabor Domokos, Toby Driscoll, Abinand Gopal, Alain Goriely, Matthew Juniper, Nancy Kopell, Derek Moulton, Sheehan Olver, Bob O'Malley, John Pryce, Brennan Sprinkle, Alex Townsend, and Maciej Zworski—and to an anonymous referee.

REFERENCES

[1] R. C. Ackerberg and R. E. O'Malley, Jr., *Boundary layer problems exhibiting resonance*, Stud. Appl. Math., 49 (1970), pp. 277–295. (Cited on pp. 439, 440, 442, 445)

[2] E. L. Allgower and K. Georg, *Numerical Continuation Methods: An Introduction*, Springer, 2012. (Cited on p. 447)

[3] U. M. Ascher, R. M. M. Mattheij, and R. D. Russell, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, SIAM, 1995, https://doi.org/10.1137/1.9781611971231. (Cited on p. 458)

[4] J. L. Aurentz and L. N. Trefethen, *Block operators and spectral discretizations*, SIAM Rev., 59 (2017), pp. 423–446, https://doi.org/10.1137/16M1065975. (Cited on p. 457)

[5] Z. Battles and L. N. Trefethen, *An extension of MATLAB to continuous functions and operators*, SIAM J. Sci. Comput., 25 (2004), pp. 1743–1770, https://doi.org/10.1137/S1064827503430126. (Cited on p. 459)

[6] A. Ben-Artzi and I. Gohberg, *Inertia theorems for nonstationary discrete systems and dichotomy*, Linear Algebra Appl., 120 (1989), pp. 95–138. (Cited on p. 458)

[7] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, Springer Science & Business Media, 2013. (Cited on p. 443)

[8] M. Desroches, B. Krauskopf, and H. M. Osinga, *Mixed-mode oscillations and slow manifolds in the self-coupled FitzHugh–Nagumo system*, Chaos, 18 (2008), art. 015107-1–8. (Cited on pp. 447, 459)

[9] E. J. Doedel et al., *AUTO-07P: Continuation and Bifurcation Software for Ordinary Differential Equations*, Technical report, Dept. of Computer Science and Software Engineering, Concordia University, 2007; see http://cmvl.cs.concordia.ca/auto/. (Cited on p. 447)

[10] G. Domokos and P. Holmes, *On nonlinear boundary-value problems: Ghosts, parasites and discretizations*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci., 459 (2003), pp. 1535–1561. (Cited on pp. 455, 456)

[11] T. A. Driscoll, N. Hale, and L. N. Trefethen, *Chebfun Guide*, Pafnuty Press, 2014; see also www.chebfun.org. (Cited on pp. 440, 458, 459)

[12] N. Fenichel, *Geometric singular perturbation theory for ordinary differential equations*, J. Differential Equations, 31 (1979), pp. 53–98. (Cited on pp. 445, 447)

[13] S. Filip, A. Javeed, and L. N. Trefethen, *Smooth random functions, random ODEs, and Gaussian processes*, SIAM Rev., 61 (2019), pp. 185–205, https://doi.org/10.1137/17M1161853 . (Cited on p. 459)

[14] A. R. Forsyth, *A Treatise on Differential Equations*, BoD–Books on Demand, 2013, originally published in 1885. (Cited on p. 439)

[15] W. Gautschi, *The condition of polynomials in power form*, Math. Comp., 33 (1979), pp. 343–352. (Cited on p. 449)

[16] C. W. Gear, T. J. Kaper, I. G. Kevrekidis, and A. Zagaris, *Projecting to a slow manifold: Singularly perturbed systems and legacy codes*, SIAM J. Appl. Dyn. Syst., 4 (2005), pp. 711–732, https://doi.org/10.1137/040608295. (Cited on p. 447)

[17] M. B. Giles and N. A. Pierce, *An introduction to the adjoint approach to design*, Flow Turbulence Combust., 65 (2000), pp. 393–415. (Cited on p. 453)

[18] M. Gomez, D. E. Moulton, and D. Vella, *Critical slowing down in purely elastic "snap-through" instabilities*, Nature Phys., 13 (2017), pp. 142–145. (Cited on p. 456)
[19] J. Grasman and B. J. Matkowsky, *A variational approach to singularly perturbed boundary value problems for ordinary and partial differential equations with turning points*, SIAM J. Appl. Math., 32 (1977), pp. 588–597, https://doi.org/10.1137/0132047. (Cited on p. 440)
[20] P. P. N. de Groen, *The nature of resonance in a singular perturbation problem of turning point type*, SIAM J. Math. Anal., 11 (1980), pp. 1–22, https://doi.org/10.1137/0511001. (Cited on pp. 440, 445, 450)
[21] J. Guckenheimer, B. Krauskopf, H. M. Osinga, and B. Sandstede, *Invariant manifolds and global bifurcations*, Chaos, 25 (2015), art. 097604.(Cited on pp. 447, 459)
[22] M. D. Gunzburger, *Perspectives in Flow Control and Optimization*, SIAM, 2002, https://doi.org/10.1137/1.9780898718720. (Cited on p. 453)
[23] S. M. Hammel, J. A. Yorke, and C. Grebogi, *Do numerical orbits of chaotic dynamical processes represent true orbits?*, J. Complexity, 3 (1987), pp. 136–145. (Cited on p. 455)
[24] P. Hartman, *Ordinary Differential Equations*, SIAM, 2002, https://doi.org/10.1137/1.9780898719222. (Cited on p. 458)
[25] C. R. Hasan, B. Krauskopf, and H. M. Osinga, *Saddle slow manifolds and canard orbits in* $\mathbf{R}^4$ *and application to the full Hodgkin–Huxley model*, J. Math. Neurosci., 8 (2018), art. 5. (Cited on pp. 447, 459)
[26] P. W. Hemker, *A Numerical Study of Stiff Two-Point Boundary Problems*, Math. Centre Tracts 80, Mathematisch Centrum, Amsterdam, 1977. (Cited on pp. 440, 458)
[27] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, 2002, https://doi.org/10.1137/1.9780898718027. (Cited on p. 441)
[28] D. Holcman and Z. Schuss, *Asymptotics of Elliptic and Parabolic PDEs*, Springer, 2018. (Cited on p. 458)
[29] M. H. Holmes, *Introduction to Perturbation Methods*, 2nd ed., Springer, 2013. (Cited on p. 443)
[30] L. Hörmander, *Differential equations without solutions*, Math. Ann., 140 (1960), pp. 169–173. (Cited on pp. 452, 457)
[31] T. Kato, *Perturbation Theory for Linear Operators*, 2nd ed., Springer, 1976. (Cited on p. 457)
[32] J. Kevorkian and J. D. Cole, *Perturbation Methods in Applied Mathematics*, Springer, 2010. (Cited on p. 443)
[33] J. Kierzenka and L. F. Shampine, *A BVP solver based on residual control and the MATLAB PSE*, ACM Trans. Math. Softw., 27 (2001), pp. 299–316. (Cited on p. 441)
[34] N. Kopell, *A geometric approach to boundary layer problems exhibiting resonance*, SIAM J. Appl. Math., 37 (1979), pp. 436–458, https://doi.org/10.1137/0137035. (Cited on pp. 440, 445)
[35] H. O. Kreiss and S. V. Parter, *Remarks on singular perturbations with turning points*, SIAM J. Math. Anal., 5 (1974), pp. 230–251, https://doi.org/10.1137/0505025. (Cited on pp. 440, 442)
[36] W. D. Lakin, *Boundary value problems with a turning point*, Stud. Appl. Math., 51 (1972), pp. 261–276. (Cited on p. 440)
[37] J.-Y. Lee and L. Greengard, *A fast adaptive numerical method for stiff two-point boundary value problems*, SIAM J. Sci. Comput., 18 (1997), pp. 403–429, https://doi.org/10.1137/S1064827594272797. (Cited on p. 440)
[38] H. Lewy, *An example of a smooth linear partial differential equation without solution*, Ann. Math., 66 (1957), pp. 155–158. (Cited on p. 457)
[39] A. D. MacGillivray, *A method for incorporating transcendentally small terms into the method of matched asymptotic expansions*, Stud. Appl. Math., 99 (1997), pp. 285–310. (Cited on pp. 440, 445)
[40] R. S. MacKay, *Slow manifolds*, in Energy Localisation and Transfer, T. Dauxois, A. Litvak-Hinenzon, R. Mackay, and A. Spanoudaki, eds., World Scientific, 2004, pp. 149–192. (Cited on p. 447)
[41] B. J. Matkowsky, *On boundary layer problems exhibiting resonance*, SIAM Rev., 17 (1975), pp. 82–100, https://doi.org/10.1137/1017005. (Cited on p. 442)
[42] B. J. Matkowsky, *Singular perturbations in noisy dynamical systems*, European J. Appl. Math., 29 (2018), pp. 570–593. (Cited on pp. 440, 445, 458)
[43] S. Mizohata, *Solutions nulles et solutions non analytiques*, J. Math. Kyoto Univ., 1 (1962), pp. 271–302. (Cited on p. 457)
[44] J. Mujica, B. Krauskopf, and H. M. Osinga, *Tangencies between global invariant manifolds and slow manifolds near a singular Hopf bifurcation*, SIAM J. Appl. Dyn. Syst., 17 (2018), pp. 1395–1431, https://doi.org/10.1137/17M1133452. (Cited on pp. 447, 459)

[45] F. W. J. Olver, *Asymptotics and Special Functions*, AK Peters/CRC Press, 1997. (Cited on p. 440)
[46] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, eds., *NIST Handbook of Mathematical Functions*, Cambridge University Press, 2010. (Cited on p. 439)
[47] S. Olver and A. Townsend, *A fast and well-conditioned spectral method*, SIAM Rev., 55 (2013), pp. 462–489, https://doi.org/10.1137/120865458. (Cited on p. 443)
[48] R. E. O'Malley, Jr., *On boundary value problems for a singularly perturbed differential equation with a turning point*, SIAM J. Math. Anal., 1 (1970), pp. 479–490, https://doi.org/10.1137/0501041. (Not cited)
[49] R. E. O'Malley, Jr., *Singularly perturbed linear two-point boundary value problems*, SIAM Rev., 50 (2008), pp. 459–482, https://doi.org/10.1137/060662058. (Cited on pp. 440, 441, 445)
[50] K. J. Palmer, *Exponential dichotomies and Fredholm operators*, Proc. Amer. Math. Soc., 104 (1988), pp. 149–156. (Cited on p. 458)
[51] C. E. Pearson, *On a differential equation of boundary layer type*, J. Math. Phys., 47 (1968), pp. 134–154. (Cited on pp. 440, 442, 458)
[52] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, 1987. (Cited on p. 449)
[53] K. K. Sharma, P. Rai, and K. C. Patidar, *A review on singularly perturbed differential equations with turning points and interior layers*, Appl. Math. Comput., 219 (2013), pp. 10575–10609. (Cited on pp. 440, 458)
[54] I. Stakgold and M. J. Holst, *Green's Functions and Boundary Value Problems*, Wiley, 2011. (Cited on p. 458)
[55] S. H. Strogatz, *Nonlinear Dynamics and Chaos with Applications to Physics, Biology, Chemistry and Engineering*, Westview Press, 2015. (Cited on p. 456)
[56] A. Tikhonov, *On the dependence of the solutions of differential equations on a small parameter*, Mat. Sb., 64 (1948), pp. 193–204. (Cited on pp. 439, 446)
[57] A. Tikhonov, *Systems of differential equations containing small parameters in the derivatives*, Mat. Sb., 73 (1952), pp. 575–586. (Cited on p. 439)
[58] L. N. Trefethen, *Wave packet pseudomodes of variable coefficient differential operators*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 461 (2005), pp. 2099–3122. (Cited on pp. 452, 456)
[59] L. N. Trefethen, A. Birkisson, and T. A. Driscoll, *Exploring ODEs*, SIAM, 2018. (Cited on pp. 440, 441, 443, 444, 457, 458)
[60] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra*, Princeton University Press, 2005. (Cited on pp. 451, 452, 457, 458, 459)
[61] A. B. Vasil'eva, V. F. Butuzov, and L. V. Kalachev, *The Boundary Function Method for Singular Perturbation Problems*, SIAM, 1995, https://doi.org/10.1137/1.9781611970784. (Cited on pp. 439, 446)
[62] A. B. Vasilieva, *On the development of singular perturbation theory at Moscow State University and elsewhere*, SIAM Rev., 36 (1994), pp. 440–452, https://doi.org/10.1137/1036100. (Cited on pp. 439, 443, 446)
[63] M. I. Vishik and L. A. Lyusternik, *Regular degeneration and boundary layer for linear differential equations with small parameter*, Usp. Mat. Nauk, 12 (1957), pp. 3–122. (Cited on pp. 439, 443)
[64] N. Wang and R. E. O'Malley, *On the Asymptotic Solution of Singularly Perturbed Boundary Value Problems with Turning Points*, manuscript, 2018. (Cited on pp. 440, 441, 458)
[65] A. M. Watts, *A singular perturbation problem with a turning point*, Bull. Austr. Math. Soc., 5 (1971), pp. 61–73. (Cited on p. 440)
[66] J. H. Wilkinson, *Rounding Errors in Algebraic Processes*, Prentice-Hall, 1963. (Cited on p. 441)
[67] T. G. Wright, *EigTool*, https://github.com/eigtool/eigtool. (Cited on pp. 452, 459)
[68] M. Zworski, *A remark on a paper of E. B. Davies*, Proc. AMS, 129 (2001), pp. 2955–2957. (Cited on pp. 452, 457)
[69] M. Zworski, *Semiclassical Analysis*, AMS, 2012. (Cited on p. 452)