Numerical error analysis for Evans function computations: a numerical gap lemma, centered-coordinate methods, and the unreasonable effectiveness of continuous orthogonalization

Kevin Zumbrun∗

April 3, 2009

Abstract

We perform error analyses explaining some previously mysterious phenomena arising in numerical computation of the Evans function, in particular (i) the advantage of centered coordinates for exterior product and related methods, and (ii) the unexpected stability of the (notoriously unstable) continuous orthogonalization method of Drury in the context of Evans function applications. The analysis in both cases centers around a numerical version of the gap lemma of Gardner–Zumbrun and Kapitula–Sandstede, giving uniform error estimates for apparently ill-posed projective boundary-value problems with asymptotically constant coefficients, so long as the rate of convergence of coefficients is greater than the “badness” of the boundary projections as measured by negative spectral gap. In the second case, we use also the simple but apparently previously unremarked observation that the Drury method is in fact (neutrally) stable when used to approximate an unstable subspace, so that continuous orthogonalization and the centered exterior product method are roughly equally well-conditioned as methods for Evans function approximation. The latter observation makes possible an extremely simple nonlinear boundary-value method for possible use in large-scale systems, extending ideas suggested by Sandstede. We suggest also a related linear method based on the conjugation lemma of Métivier–Zumbrun, an extension of the gap lemma mentioned above.

Contents

1 Introduction 3
1.1 Computation of the Evans function . . . . . . . . . . . . . . . . . . . . . . . 3
1.2 Three Bad Things: numerical pitfalls and their resolutions . . . . . . . . . . 4
  1.2.1 Potential pitfalls . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4
  1.2.2 Solution one: the centered exterior product method . . . . . . . . . . 6

∗Indiana University, Bloomington, IN 47405; kzumbrun@indiana.edu: Research of K.Z. was partially supported under NSF grants number DMS-0300487, DMS-0505780, and DMS-0801745.
| Section | Title | Page |
|---------|-------|------|
| 1.2.3  | Solution two: the polar coordinate method | 6 |
| 1.3    | Further questions and description of results | 7 |
| 1.4    | Discussion and open problems | 10 |
| 2      | Preliminaries: the gap and conjugation lemmas | 11 |
| 3      | A numerical gap lemma | 13 |
| 3.1    | Continuous problem | 13 |
| 3.2    | Discrete problem | 14 |
| 3.2.1  | Difference scheme | 15 |
| 3.2.2  | Basic assumptions | 15 |
| 3.3    | Discrete conjugation error | 16 |
| 3.4    | Numerical convergence lemma | 18 |
| 3.5    | Convergence of the centered exterior product method | 21 |
| 3.5.1  | Mesh requirements, and computations in the essential spectrum | 22 |
| 4      | Stability of continuous orthogonalization | 23 |
| 4.1    | Asymptotic stability in tangential directions | 23 |
| 4.2    | Asymptotic stability in transverse directions | 24 |
| 4.3    | Convergence of the polar coordinate method | 25 |
| 5      | Boundary-value algorithms | 26 |
| 5.1    | Sandstede’s method | 26 |
| 5.2    | A polar coordinate-based method | 27 |
| 5.3    | A conjugation-based method | 27 |
| 6      | Postscript: initialization of eigenbases at infinity | 28 |
| 6.1    | Kato’s ODE | 28 |
| 6.2    | Numerical implementation | 28 |
| 6.2.1  | Computing $\Pi_j$ | 29 |
| 6.2.2  | First-order integration scheme | 30 |
| 6.2.3  | Second-order scheme | 30 |
| 6.3    | Initialization of Evans function ODE | 30 |
| 6.3.1  | Centered exterior product scheme | 31 |
| 6.3.2  | Polar coordinate scheme | 31 |
| 6.4    | Error control | 31 |
| 6.5    | Finer points: two exceptional cases | 32 |
| 6.5.1  | Behavior near the origin | 32 |
| 6.5.2  | Behavior near a branch singularity | 32 |
1 Introduction

Recently, numerical Evans function computations have received a great deal of attention as a tool for the stability analysis of standing and traveling wave patterns in one and several dimensions; see, e.g., [AS, Br, BrZ, BDG, AlB, HSZ, HuZ1, HuZ2, LPSS, GLZ, HLZ, CHNZ, HLyZ1, HLyZ2]. In the work of the author together with Brin, Humpherys, Sandstede, and others, there has emerged a small list of three computational rules of thumb, without which Evans function computations become hopelessly inefficient, but with which they become in usual situations almost trivial. Indeed, Humpherys has developed a general MATLAB-based package (STABLAB) based on these principles that gives excellent results on essentially all problems up to now considered.

Two of the items on this list are self-evident, but the third, the need to “center” coordinates, does not appear to be well-known and, indeed, at first sight appears to contradict standard stability principles. The purpose of this paper is twofold: first, to share these practical rules of thumb and, second, to give a mathematical justification for the better-than-expected observed results of their implementation, that is, to put these ad hoc principles on a rational and quantitative basis. In the process, we discover a numerical analog of the gap lemma of [GZ, KS], a sort of superconvergence principle; a new stability property of the well-known continuous orthogonalization method of Drury; and, building on ideas of Sandstede [S] and Humpherys–Zumbrun [HuZ1], an extremely simple boundary-value method for possible use in ultra-large scale systems.

1.1 Computation of the Evans function

Let $L$ be a linear differential operator with asymptotically constant coefficients along some preferred spatial direction $x$, and suppose that the eigenvalue equation

$$(L - \lambda)w = 0$$

may be expressed as a first-order ODE in an appropriate phase space:

$$(1.2) \quad W_x = A(x, \lambda)W, \quad \lim_{x \to \pm \infty} A(x, \lambda) = A_\pm(\lambda),$$

with $A$ analytic in $\lambda$ as a function from $\mathbb{C}$ to $C^1(\mathbb{R}, \mathbb{C}^{n \times n})$ and the dimension $k$ of the stable subspace $S_+$ of $A_+$ and dimension $n - k$ of the unstable subspace $U_-$ of $A_-$ summing to the dimension $n$ of the entire phase space. Then, the Evans function is defined as

$$(1.3) \quad D(\lambda) := \det(W_1^+ \cdots W_k^+ W_{k+1}^- \cdots W_n^-)|_{x=0},$$

where $W_1^+, \ldots, W_k^+$ and $W_{k+1}^-, \ldots, W_n^-$ are analytically-chosen (in $\lambda$) bases of the manifolds of solutions decaying as $x \to +\infty$ and $-\infty$. For details of this construction, see, e.g., [AGJ, PW, KS, GZ, Z1, HuZ1, HSZ] and references therein.

Analogous to the characteristic polynomial for a finite-dimensional operator, $D(\cdot)$ is analytic in $\lambda$ with zeroes corresponding in both location and multiplicity to the eigenvalues
of the linear operator $L \{GJ1, GJ2\}$. Taking the winding number around a contour $\Gamma = \partial \Lambda \subset \{\Re \lambda \geq 0\}$, where $\Lambda$ is a set outside which eigenvalues may be excluded by other methods (e.g. energy estimates or asymptotic ODE theory), counts the number of unstable eigenvalues in $\Lambda$ of the linearized operator about the wave, with zero winding number corresponding to stability. See, e.g., [Br, BrZ, BDG, HSZ, HuZ1, BHRZ, HLZ, CHNZ, HLyz1, HLyz2, BHZ]. Alternatively, one may use Mueller’s method or any number of root-finding methods for analytic functions to locate individual roots; see, e.g., [OZ, LS].

Numerical approximation of the Evans function breaks into two steps: (i) the computation of analytic bases for stable (resp. unstable) subspaces of $A_+$ (resp. $A_-$) and (ii) the propagation of these bases by ODE (1.2) on a sufficiently large interval $x \in [M, 0]$ (resp. $x \in [-M, 0]$). In both steps, it is important to preserve the fundamental property of analyticity in $\lambda$, which is extremely useful in computing roots by winding number or other methods. Both problems concern (different aspects of) numerical propagation of subspaces, the first in $\lambda$ and the second in $x$, thus tying into large bodies of theory in both numerical linear algebra [ACR, DDF, DE1, DE2, DF] and hydrodynamic stability theory [Dr, Da, NR1, NR2, NR3, NR4, B].

Problem (i) has been examined in [HSZ, Z2, BHZ]; for completeness, we gather the (existing but dispersed) conclusions here in Section 6. Our main emphasis, however, is on problem (ii) and numerical stability analysis, for which we obtain substantially new results.

1.2 Three Bad Things: numerical pitfalls and their resolutions

We now focus on problem (ii). Our three basic principles for efficient numerical integration of (1.2) are readily motivated by consideration of the simpler constant-coefficient case

\begin{equation}
W_x = AW, \quad A \equiv \text{constant}.
\end{equation}

1.2.1 Potential pitfalls

We note the following three basic pitfalls, two obvious and one perhaps less so.

1. **Wrong direction of integration.** Consider the simplest case that the dimension of the stable subspace of $A$ is one, so that we seek to resolve a single decaying eigenmode, with all other modes exponentially growing with increasing $x$. Evidently, the correct direction of integration is the backwards direction, from $x = +M$ back to $x = 0$, in which the desired mode is exponentially growing, and errors in other modes exponentially decay. Integrating in the forward direction would be numerically disastrous, with exponential error growth $e^{\eta M}$, where $\eta$ is the spectral gap between decaying and growing modes (recall, $M$ is large): the analog for Evans function computations of integrating a backward heat equation. In general, we must always integrate from infinity toward zero.

2. **Parasitic modes** (related to 1). For general systems of equations, the dimension of the stable subspace of $A$ typically involves two or more eigenmodes, with distinct decay rates $\mu_1 < \mu_2 < 0$. Integrating in backward direction as prescribed in part 1 above, we
resolve the fastest decaying $\mu_1$ mode without difficulty. However, in trying to resolve the slower decaying $\mu_2$ mode, we experience the problem that errors in the direction of the $\mu_1$ mode grow exponentially relative to the desired $\mu_2$ mode, at relative rate $e^{(\mu_2-\mu_1)M}$. That is, parasitic faster-decaying modes will tend to take over slower-decaying modes, preventing their resolution.

**Remark 1.1.** Degradation of results from parasitic modes is of the same rough order as that resulting from integrating in the wrong spatial direction, differing only in the fact that the maximum spectral gap between two decaying modes is typically one-half or less of the maximum gap between decay and growing modes. Thus, it is crucial to address this issue.

### 3. Nonequilibrium state

A more subtle problem is that integrating even a single scalar equation $w' = -aw$, $a \geq 0$, over a long interval $[M, 0]$, leads to (sometimes quite large) accumulation of errors, and, more important, a large number of mesh points/computations. The single exception is the equilibrium case $a = 0$, which for essentially all numerical ODE schemes is resolved exactly.

This can be understood more quantitatively by the following heuristic computation, assuming a perfectly adaptive scheme and no machine error. The truncation error $\tau_j$ for a $k$th order scheme at step $j$ is proportional to the $(k+1)$th derivative of the solution times $\Delta x_j$, where $\Delta x_j$ is the size of the $j$th step, or $c_k a^{k+1} e^{-ax_j} \Delta x_j^k$. Taking $\tau_j \sim TOL$ for some fixed tolerance $TOL$, we thus obtain $c_k a^{k+1} e^{-ax_j} \Delta x_j^k \sim TOL$, or

$$\frac{\Delta x_j}{\Delta j} \sim c_k^{-1/k} TOL^{1/k} \times a^{-1-1/k} e^{ax_j/k}.$$

Inverting, and integrating $\frac{\Delta j}{\Delta x_j} \approx \frac{dj}{dx}$ from 0 to $M$, we obtain an estimate

$$(1.5) \quad J \sim c_k^{1/k} TOL^{-1/k} ka^{1/k} \int_0^M (a/k) e^{-ax_j/k} \sim c_k^{1/k} TOL^{-1/k} ka^{1/k}$$

as $M \to +\infty$ for the total number $J$ of mesh blocks, which goes to zero as $a \to 0$ and to infinity as $a \to \infty$.

Though it is tempting to think of this as an example of numerical stiffness, that is not the case, since the problem involves but a single mode. It seems rather to be a secondary, previously unremarked, phenomenon, that in usual circumstances would be negligible. For Evans function computations, however, we observe a difference in computational efficiency of an order of magnitude or more between the cases $a = 1$ and $a = 0$, for $TOL \sim 10^{-6}$.

**Remark 1.2.** It would be interesting to compare (1.5) to results for the standard RK45 scheme with which most Evans computations have been done applied to the constant coefficient scalar problem $w' = -aw$, in particular the $a^{1/4}$ rate. For variable-coefficient systems, additional effects having to do with “conjugation errors” appear as well; see Section 3.

---

1 Direction of integration is symmetric for a single mode.
1.2.2 Solution one: the centered exterior product method

Problem 1 is easily avoided by integrating in the correct direction. Problem 2 may be overcome by working in the exterior product space $W_1^+ \wedge \cdots \wedge W_k^+ \in \mathbb{C}^\binom{n}{k}$ (resp. $W_{k+1}^- \wedge \cdots \wedge W_n^- \in \mathbb{C}^{(n-k)}$), for which the desired subspace appears as a single, maximally stable (resp. unstable) mode, the Evans determinant then being recovered through the isomorphism

$$\det (W_1^+ \cdots W_k^+ W_{k+1}^- \cdots W_n^-) \sim (W_1^+ \wedge \cdots \wedge W_k^+ \wedge (W_{k+1}^- \wedge \cdots \wedge W_n^-));$$

see [AS, Br, BrZ, BDG, AlB] and ancestors [GB, NR1, NR2, NR3, NR4]. This reduces the problem to the case $k = 1$. Problem 3 can then be avoided by factoring out the expected asymptotic decay rate $e^{\mu x}$ of the single decaying mode and solving the “centered” equation

$$Z' = (A - \mu I)Z, \quad Z(+\infty) = r : A_+r = \mu r$$

for $Z := e^{-\mu x}W$, which is now asymptotically an equilibrium as $x \to +\infty$. With these preparations, one obtains excellent results [BDG, HuZ1]; however, omitting any one of them leads to a loss of efficiency of at least an order of magnitude in our experience [HuZ2].

1.2.3 Solution two: the polar coordinate method

Unfortunately, the dimension $\binom{n}{k}$ of the phase space for the exterior product grows exponentially with $n$, since $k \sim n/2$ in typical applications. This limits its usefulness to $n \leq 10$ or so, whereas the Evans system arising in compressible MHD is size $n = 15$, $k = 7$ [BHZ], giving a phase space of size $\binom{n}{k} = 6,435$: clearly impractical. A more compact, but nonlinear, alternative is the polar coordinate method of [HuZ1], in which the exterior products of the columns of $W_\pm$ are represented in “polar coordinates” $(\Omega, \gamma)_\pm$, where the columns of $\Omega_+ \in \mathbb{C}^{n \times k}$ and $\Omega_- \in \mathbb{C}^{(n-k) \times k}$ are orthonormal bases of the subspaces spanned by the columns of $W_+ := (W_1^+ \cdots W_k^+)$ and $W_- := (W_{k+1}^- \cdots W_n^-)$, $W_j^\pm$ defined as in (1.3), i.e., $W_+ = \Omega_+ \alpha_+$, $W_- = \Omega_- \alpha_-$, and $\gamma_\pm := \det \alpha_\pm$, so that

$$W_1^+ \wedge \cdots \wedge W_k^+ \wedge = \gamma_+(\Omega_1^+ \wedge \cdots \wedge \Omega_k^+),$$

where $\Omega_j^\pm$ denotes the $j$th column of $\Omega_\pm$, and likewise $W_{k+1}^- \wedge \cdots \wedge W_n^- = \gamma_-(\Omega_1^- \wedge \cdots \wedge \Omega_n^-)$.

This yields the block-triangular system

$$\Omega' = (I - \Omega^* \Omega)A \Omega, \quad (\log \tilde{\gamma})' = \text{trace} (\Omega^* A \Omega) - \text{trace} (\Omega^* A \Omega)(\pm \infty),$$

$$\tilde{\gamma} := \tilde{\gamma} e^{\text{trace} (\Omega^* A \Omega)(\pm \infty)}x,$$

for which the “angular” $\Omega$-equation is exactly the continuous orthogonalization method of Drury [Dr, Da], and the “radial” $\tilde{\gamma}$-equation, given $\Omega$, may be solved by simple quadrature. Ignoring the numerically trivial radial equation, we see that problem 2 by fiat does not occur. Likewise, for constant $A$, it is easily verified that invariant subspaces $\Omega$ of $A$ are equilibria of the flow, so problem 3 does not occur. Indeed,
for $A$ constant, solutions of the $\tilde{\gamma}$-equation are constant, so that any first-order or higher numerical scheme resolves $\log \tilde{\gamma}$ exactly; thus, the $\tilde{\gamma}$-equation may for simplicity be solved together with the $\Omega$-equation, with no need for a final quadrature sweep. The Evans function is recovered, finally, through the relation

$$\text{(1.9)} \quad D(\lambda) = \det (W_1^+ \cdots W_k^+ W_{k+1}^- \cdots W_n^-) |_{x=0} = \tilde{\gamma}_+ \tilde{\gamma}_- \det(\Omega^+, \Omega^-) |_{x=0}.$$

1.3 Further questions and description of results

The discussion of the previous subsection leaves open some important questions. First, can we justify this heuristic discussion with rigorous, quantitative error bounds for the actual, variable coefficient problems that occur in practice? In particular, we note that centering equations as in Section 1.2.2 goes counter to the intuition afforded by standard two-point boundary-value theory on intervals $[0, M]$ as $M \to \infty$ [Be1], which asserts convergence error of order $e^{-\eta M}$ where $\eta$ is the minimum spectral gap of decaying (resp. growing) modes from zero to the solution on $[0, +\infty)$. Applied blindly to the centered equations, this would predict nonconvergence rather than the good behavior observed in practice.

Second, there is a well-known problem of instability of the continuous orthogonalization method with respect to perturbations disturbing the assumed orthonormal structure of $\Omega$ [Dr, BrRe]. In the language of [BrRe], the Stiefel manifold $S := \{\Omega : \Omega^* \Omega = I_k\}$ of orthonormal matrices is preserved by the flow of (1.8), but is typically neither attracting nor repelling. In view of Remark 1.1, this should lead to terrible results for Evans function computations. This issue was discussed at length in [HuZ1], with numerous different solutions discussed, from artificial stabilization to geometric integration. Yet, surprisingly, the method that performed best was the original Drury algorithm with no stabilization, implemented by a standard RK45 scheme. This yielded results quite similar to those of the exterior product scheme, which seems to contradict the conclusions of Remark 1.1.

**Result I.** Our first main result is to establish a numerical version of the gap lemma of [GZ, KS], which states that, ignoring machine error, provided the coefficient matrix $A(x, \lambda)$ is uniformly exponentially convergent as $x \to +\infty$, with rate $|A - A_+| \leq Ce^{-\theta x}$ for $x \geq 0$, and provided the gap between $\mu$ minimum real part of the eigenvalues of $A_+$ is strictly greater than $-\theta$, then the solution of problem (1.7) on $[0, M]$ initialized as $Z(M) = r$ converges as $M \to \infty$ to the solution on $[0, +\infty)$ at rate $C(\tilde{\theta}, \theta)e^{-\tilde{\theta}M}$ for any $0 < \tilde{\theta} < \theta$. This resolves the first issue, explaining the observed convergence of the centered exterior product method.

Completing the analogy to [GZ], we establish a corresponding result for general centered two-point boundary problems (1.7) with projective boundary conditions on $[0, M]$, under the assumption that the gap $\gamma$ between $\mu$ minimum real part of the eigenvalues of $A_+$ associated with the projective boundary condition at $M$ is strictly greater than $-\theta$, obtaining convergence at the same rate $C(\tilde{\theta}, \theta)e^{-\tilde{\theta}M}$ as $M \to +\infty$ for any $0 < \tilde{\theta} < \theta$. This could be viewed as a type of superconvergence, as the standard theory [Be1] predicts convergence at

---

2 See Section 6.3.2 for numerical prescriptions $(\Omega, \tilde{\gamma})_+$ of $(\Omega, \tilde{\gamma})$ at $\pm\infty$.
rate $e^{-\gamma x}$, with a nonpositive spectral gap $\gamma \leq 0$ corresponding to ill-conditioned boundary conditions. For detailed statements and proofs, see Section 3.

**Result II.** Our second main result is to explain the apparent contradiction between observed good results for the polar coordinate method [HuZ1] and the well-known instability of continuous orthogonalization [Da, BrRe]. The simple resolution is that, though continuous orthogonalization is in general unstable, it is in the present context stable!

Heuristically, this is quite simple to see. Intuitively, it is is clear that the stable manifold of $A_+$ is asymptotically attracting in backward $x$ under the flow of (1.8) for $\Omega$ confined to the Stiefel manifold $S = \{ \Omega : \Omega^*\Omega = I_k \}$, and in fact this is well known (see Section 4.1). Thus, we need only verify that the Stiefel manifold, likewise, is attracting in backward $x$.

Defining the Stiefel error $E(\Omega) := \Omega^*\Omega - I$, we obtain after a brief computation the error equation

\[(1.10) \quad E' = -E(\Omega^*A\Omega) - (\Omega^*A\Omega)^*E\]

of [HuZ1]. Linearizing about the exact solution $\bar{\Omega} \to \Omega_+, \bar{E} \to 0$ and replacing coefficients by their asymptotic limits, we obtain a linear equation $E' = A_+E$, where $A\mathcal{E} := -E\bar{A}_+ - (\bar{A}_+)^*E$ is a Sylvester operator, $A_+ := (\Omega^*A_+\Omega^*_k)$, with eigenvalues and eigenmatrices $a_j + a_k^*, r_jr_k^*$, where $a_j$ and $r_j$ are eigenvalues and eigenvectors of $\bar{A}_+$. Noting that the eigenvalues of $\bar{A}_+$ are exactly the eigenvalues of $A_+$ restricted to its stable subspace, i.e., the stable eigenvalues of $A_+$, we find that $A_+$ has positive real part eigenvalues, and so $E$ decays in backward $x$.

That is, the Stiefel manifold is attracting under the backward flow of continuous orthogonalization (repelling under the forward flow) if $\Omega$ is a stable subspace of $A$, an observation that previously seems to have gone unremarked. This confirms that, as suggested by numerical results of [HuZ1], continuous orthogonalization is roughly equally well-conditioned as the centered exterior product method. For details and further discussion, see Section 4.

**Remark 1.3.** An important consequence is that geometric integrators like those suggested in [BrRe] for general Orr–Sommerfeld applications are probably not worth the trouble for Evans function computations, since the Drury method is stable and much simpler to code.

**Remark 1.4.** The generalized inverse method, or Davey method [Da], $\Omega' = (I - \Omega\Omega^*)A\Omega$, where $\Omega^* := (\Omega^*\Omega)^{-1}\Omega^*$ denotes the generalized inverse, exhibits neutral error growth $E' = 0$, so is often used as a stabilization of the basic continuous orthogonalization method (1.8) of Drury. In the context of Evans function computations, the behavior is slightly worse, however [HuZ1]. This can now be understood from the fact that the Drury method actively damps errors in the Evans function context. The fact that the Drury method outperformed the Davey method (and all others) was a mysterious aspect left unresolved in [HuZ1].

**Result III.** For equations arising in complicated physical systems or through transverse discretization of a multi-dimensional problem on a cylindrical domain [LPSS], the dimension $n$ can be very large. For example, for the multidimensional systems considered in [LPSS], $n = 8M \sim 48$ (see p. 1447, [LPSS]) where $M \sim 6$ is the number of transverse Fourier
modes being computed. The development of numerical methods suitable for efficient Evans function computations for large systems has been cited by Jones and others as one of the key problems facing the traveling-wave community in the next generation [J].

For large dimensions, the accumulation of errors associated with shooting methods appears potentially problematic, and so various other options have been considered. For example, one may always abandon the Evans function formulation and go back to direct discretization/Galerkin techniques, hoping to optimize perhaps by multi-pole type expansions on a problem-specific basis. However, this ignores the useful structure, and associated dimensionality reduction, encoded by existence of the Evans function.

Alternatively, Sandstede [S] has suggested to work within the Evans function formulation, but, in place of the high-dimensional shooting methods described above, to recast (1.2) as a boundary-value problem with appropriate projective boundary conditions, which may be solved in the original space $\mathbb{C}^n$ for individual modes by robust and highly-accurate boundary-value/continuation techniques.

Problems with this scheme as conventionally implemented in uncentered coordinates are two. First, solutions exponentially decay as $x \to \pm \infty$, so that the direct connection to data at $\infty$ of (1.7) is lost; as a consequence, up to now, it is not known how to recover analyticity of the Evans function by such a scheme. Second, since the uncentered problem involves decaying modes as well as growing modes, there must be provided boundary conditions at $x = 0$ as well as at $x = \pm M$; indeed, it is the boundary conditions at $x = 0$ that mainly determine the decaying modes we seek. Since behavior of (1.2) is only known near its asymptotic limits as $x \to \pm \infty$, there appears to be no analytic way to prescribe a priori well-conditioned projective boundary conditions at $x = 0$ (or, as mentioned already, to relate these to a desired asymptotic behavior as $x \to \pm \infty$), and so apparently these must be adjusted “on the fly” by trial and error, a process that requires error checks and additional complications in program structure.

As pointed out in [HuZ1] (last sentence of introduction), using the polar coordinate method– or any centered scheme– as the basis of a boundary-value scheme eliminates immediately the first problem, of preserving analyticity, since in the centered format date is explicitly described as $x \to \pm \infty$. The second problem in general remains. However, by the remarkable stability property recorded in result II, the polar coordinate method involves only modes that are neutral or growing as $x \to \pm \infty$, and not decaying; that is, it is both centered and one-sided. The centered exterior product method though dimensionally unsuitable shares this property as well; indeed, it is precisely the one-sided property that makes these schemes suitable for shooting. For such schemes, appropriate projective boundary conditions by result I are full Dirichlet conditions as $x \to \pm \infty$, with no boundary conditions at $x = 0$, and thus the second, essentially logistic, problem does not either arise.

We therefore propose the polar coordinate method with Dirichlet conditions at $x = \pm M$ as a promising candidate for boundary-value-based Evans function computations, noting in particular that it is essentially trivial to program given an existing shooting code. The only disadvantage that we immediately see is the nonlinearity of the scheme. We propose at the

---

3 See however [GLZ], which points out a useful intermediate structure based on Fredholm determinants.
same time an alternative linear, centered but two-sided, scheme based on the conjugation lemma of \[ \text{MeZ1} \; \text{[Z1]} \], an extension of the gap lemma that is our main tool in the analysis.

### 1.4 Discussion and open problems

We gather in this paper a complete prescription together with rigorous error bounds for efficient numerical Evans function computations by the shooting methods of \[ \text{HuZ1} \; \text{[HLZ]} \; \text{[HLyZ1]} \; \text{[BHZ]}, \] etc., of systems up to the intermediate size \( n \sim 20 \) or so encountered in one- and multi-dimensional problems of continuum mechanics, and propose some promising boundary-value methods for further exploration in computations for ultra-large systems of size \( n \sim 50 \) and up. In the process, we rehabilitate the continuous orthogonalization method of Drury, explaining its unexpectedly good performance in the context of Evans function computations. Finally, and most important, we point out the importance of centered coordinates, in contrast to standard numerical intuition and practice in the study of boundary-value problems on unbounded domains \[ \text{[Be1]} \], establishing the related stability/superconvergence principle embodied by our numerical gap lemma.

The latter result seems to suggest larger implications in the construction of numerical boundary-value schemes. At the same time, it serves to clarify some up-to-now rather confusing existing results. For example, in the seminal work \[ \text{[Er]} \], Erpenbeck performed a numerical Evans function analysis of stability of ZND detonation waves by a method obeying principle 1 and 2 of Section 1.2.1. Much later, Lee and Steward \[ \text{[LS]} \] introduced what is now the effective standard method in detonation literature, obeying principle 2 but violating principle 1, reporting an apparently counter-intuitive improvement in speed of computation. The explanation of this paradox is that Erpenbeck carried out his computations in uncentered coordinates, whereas Lee and Steward, by mapping to a bounded interval and applying a singular integral solver effectively centered their equations, factoring out the principal dynamics. Thus, there is a cancellation of errors involved, that cannot be seen without reference to principle 3. A centered version of Erpenbeck’s original method appears to outperform both schemes by an order of magnitude; see \[ \text{[HuZ2]} \] for further discussion/simplifications.

The main mathematical interest of our numerical convergence results is their application to two-sided boundary-value schemes posed on the entire interval \([0, M]\). Indeed, for shooting methods, a routine translation into the discrete setting of the continuous gap lemma yields the result, whereas the general case involves a more subtle argument based on approximate conjugation to constant-coefficients; see Remark 3.12. The determination of realistic mesh requirements for centered boundary-value schemes, and the question of whether or not centered boundary-value schemes yield in practice the same good performance observed for centered shooting schemes, remain important open problems.
2 Preliminaries: the gap and conjugation lemmas

We begin by recalling the standard gap and conjugation lemmas of [GZ] and [KS]. Consider a general family of first-order ODE

\[ W' - \Lambda(x, \Lambda)W = 0 \]

indexed by a parameter \( \Lambda \in \Omega \subset \mathbb{C}^m \), where \( W \in \mathbb{C}^N \), \( x \in \mathbb{R} \) and \( \partial/\partial x \) denotes \( d/dx \). Assume

(h0) Coefficient \( \Lambda(\cdot, \Lambda) \), considered as a function from \( \Omega \) into \( C^0(x) \) is analytic in \( \Lambda \). Moreover, \( \Lambda(\cdot, \Lambda) \) approaches exponentially to limits \( \Lambda_{\pm} \) as \( x \to \pm \infty \), with uniform exponential decay estimates

\[ |(\partial/\partial x)^k(\Lambda - \Lambda_{\pm})| \leq Ce^{-\theta|x|}, \quad \text{for } x \geq 0, \quad 0 \leq k \leq K, \]

\( C, \theta > 0 \), on compact subsets of \( \Omega \).

**Lemma 2.1** (The gap lemma [GZ, ZH]). Assuming (h0), if \( V^{-}(\Lambda) \) is an eigenvector of \( \Lambda_{-} \) with eigenvalue \( \mu(\Lambda) \), both analytic in \( \Lambda \), then there exists a solution of (2.1) of form

\[ W(\Lambda, x) = V(x, \Lambda)e^{\mu(\Lambda)x}, \]

where \( V \) is \( C^1 \) in \( x \) and locally analytic in \( \Lambda \) and, for any fixed \( \theta < \Theta \), satisfies

\[ V(x, \Lambda) = V^{-}(\Lambda) + O(e^{-\theta|x|}V^{-}(\Lambda))), \quad x < 0. \]

**Proof.** Setting \( W(x) = e^{\mu x}V(x) \), we may rewrite \( W' = \Lambda W \) as

\[ V' = (\Lambda_{-} - \mu I)V + \Theta V, \quad \Theta := (\Lambda - \Lambda_{-}) = O(e^{-\theta|x|}), \]

and seek a solution \( V(x, \Lambda) \to V^{-}(x) \) as \( x \to -\infty \). Choose \( \tilde{\theta} < \Theta_1 < \Theta \) such that there is a spectral gap \( |\Re(\sigma_{\Lambda_{-}} - (\mu + \Theta_1))| > 0 \) between \( \sigma_{\Lambda_{-}} \) and \( \mu + \Theta_1 \). Then, fixing a base point \( \Lambda_0 \), we can define on some neighborhood of \( \Lambda_0 \) to the complementary \( \Lambda_{-} \)-invariant projections \( P(\Lambda) \) and \( Q(\Lambda) \) where \( P \) projects onto the direct sum of all eigenspaces of \( \Lambda_{-} \) with eigenvalues \( \tilde{\mu} \) satisfying \( \Re(\tilde{\mu}) < \Re(\mu) + \Theta_1 \), and \( Q \) projects onto the direct sum of the remaining eigenspaces, with eigenvalues satisfying \( \Re(\tilde{\mu}) > \Re(\mu) + \Theta_1 \). By basic matrix perturbation theory (eg. [Kat]) it follows that \( P \) and \( Q \) are analytic in a neighborhood of \( \Lambda_0 \), with

\[ e^{(\Lambda_{-} - \mu I)x}P \mid_{x > 0} \leq C(e^{\Theta_1 x}), \quad e^{(\Lambda_{-} - \mu I)x}Q \mid_{x < 0} \leq C(e^{\Theta_1 x}), \]

It follows that, for \( M > 0 \) sufficiently large, the map \( T \) defined by

\[ TV(x) = V^{-} + \int_{-\infty}^{x} e^{(\Lambda_{-} - \mu I)(x-y)}P\Theta(y)V(y)dy \]

\[ -\int_{x}^{-M} e^{(\Lambda_{-} - \mu I)(x-y)}Q\Theta(y)V(y)dy \]
is a contraction on $L^\infty(-\infty,-M]$. For, applying (2.6), we have

$$|TV_1 - TV_2|_x \leq C|V_1 - V_2|_\infty \left( \int_{-\infty}^x e^{\theta_1(x-y)}e^{\theta y}dy + \int_x^{-M} e^{\theta_1(x-y)}e^{\theta y}dy \right)$$

(2.8)

$$= C|V_1 - V_2|_\infty \frac{e^{\theta_1x}e^{-(\theta - \theta_1)M}}{\theta - \theta_1} < \frac{1}{2}|V_1 - V_2|_\infty.$$

By iteration, we thus obtain a solution $V \in L^\infty(-\infty,-M]$ of $V = TV$ with $V \leq C_3|V^-|$; since $T$ clearly preserves analyticity $V(\Lambda, x)$ is analytic in $\Lambda$ as the uniform limit of analytic iterates (starting with $V_0 = 0$). Differentiation shows that $V$ is a bounded solution of $V = TV$ if and only if it is a bounded solution of (2.5). Further, taking $V_1 = V$, $V_2 = 0$ in (2.8), we obtain from the second to last inequality that

$$|V - V^-| = |T(V) - T(0)| \leq C_2e^{\tilde{\theta}x}|V| \leq C_4e^{\tilde{\theta}x}|V^-|,$$

giving (2.4). Analyticity, and the bounds (2.4), extend to $x < 0$ by standard analytic dependence for the initial value problem at $x = -M$.

**Remark 2.2.** The title “gap lemma” alludes to the fact that we do not make the usual assumption of a spectral gap between $\mu(\Lambda)$ and the remaining eigenvalues of $A_-$, as in standard results on asymptotic behavior of ODE [Ca]; that is, the lemma asserts that exponential decay of $A_-$ can substitute for a spectral gap.

**Remark 2.1.** In the case $\Re\sigma(A_+) > -\theta$, we may take $Q = \emptyset$ and $\tilde{\theta} = \theta$, improving (2.4).

**Corollary 2.3** (The conjugation lemma [McZ]). Given (h0), there exist locally to any given $\Lambda_0 \in \Omega$ invertible linear transformations $P_+(x,\Lambda) = I + \Theta_+(x,\Lambda)$ and $P_-(x,\Lambda) = I + \Theta_-(x,\Lambda)$ defined on $x \geq 0$ and $x \leq 0$, respectively, $\Phi_\pm$ analytic in $\Lambda$ as functions from $\Omega$ to $C^0[0, \pm\infty)$, such that:

(i) For any fixed $0 < \tilde{\theta} < \theta$ and $0 \leq k \leq K + 1$, $j \geq 0$,

$$|(\partial/\partial \Lambda)^j(\partial/\partial x)^k\Theta_\pm| \leq C(j, k)e^{-\tilde{\theta}|x|} \quad \text{for} \quad x \geq 0.$$

(2.10)

(ii) The change of coordinates $\mathbb{W} =: P_\pm Z$, $\mathbb{F} =: P_\pm G$ reduces (2.1) to

$$Z' - A_\pm Z = G \quad \text{for} \quad x \geq 0.$$

(2.11)

**Remark 2.2.** Equivalently, solutions of (2.1) may be factored as

$$\mathbb{W} = (I + \Theta_{\pm})Z_\pm,$$

where $Z_\pm$ satisfy the limiting, constant-coefficient equations (2.11) and $\Theta_{\pm}$ satisfy (2.10).
Proof. Substituting \( W = P Z \) into (2.1), equating to (2.11), and rearranging, we obtain the defining equation

\[
P' = \mathbb{A}_- P - P \mathbb{A}, \quad P_- \rightarrow I \quad \text{as} \quad x \rightarrow -\infty.
\]

Viewed as a vector equation, this has the form \( P' = A P_- \), where \( A \) approaches exponentially as \( x \rightarrow -\infty \) to its limit \( A_- \), defined by

\[
A_- P := \mathbb{A}_- P - P \mathbb{A}.
\]

The limiting operator \( A_- \) evidently has analytic eigenvalue, eigenvector pair \( \mu \equiv 0, \ P_- \equiv I \), whence the result follows by Lemma 2.1 for \( j = k = 0 \). The \( x \)-derivative bounds \( 0 < k \leq K + 1 \) then follow from the ODE and its first \( K \) derivatives, and the \( \Lambda \)-derivative bounds from standard interior estimates for analytic functions. Finally, invertibility of \( P_- \) follows for \( x \) large and negative from (2.10) and for \( x \leq 0 \) by global existence of a solution to

\[
(P^{-1})' = -P^{-1}P'P^{-1} = A_+P^{-1} - P^{-1}A.
\]

A symmetric argument gives the result for \( P_+ \).

\[\square\]

3 A numerical gap lemma

We now establish the main result of the paper, a numerical analog of Lemma 2.1

3.1 Continuous problem

Consider similarly as in (2.1) a first-order ODE

\[
W' - A(x)W = 0, \quad W \in \mathbb{C}^N
\]
on the half-line \( x \in [0, +\infty) \), assuming exponential convergence

\[
|(\partial/\partial x)^k (A - A_+)| \leq Ce^{-\theta x}, \quad \text{for} \ x \geq 0, \ 0 \leq k \leq K,
\]

\( C, \theta > 0 \), of \( A \) to \( A_+ \) and asymptotic stationarity

\[
V_+ \in \text{Ker} A_+.
\]

Suppose further that there holds the following gap condition.

Assumption 3.1. \( \Sigma_+ \) is a \( k \)-dimensional invariant subspace of \( A_+ \) containing all eigenmodes associated with nonnegative real part eigenvalues, and no eigenmodes with real part \( \leq -\theta \), with associated eigenvector \( \Pi_+ \).

Let \( \Pi_0 \) be an arbitrary projection of rank \( n - k \).
Lemma 3.1. For generic $\Pi_0$, specifically those satisfying (3.10) below, there exists for any $\alpha \in \text{Range } \Pi_0$ a unique solution of (3.1) under the projective boundary conditions
(3.4) $\Pi_0 W(0) = \alpha, \quad \lim_{x \to +\infty} \Pi_+ (W(x) - V_+ ) = 0,$
satisfying for any $0 < \tilde{\theta} < \theta$
(3.5) $|\partial_x^k (W(x) - V_+ )| \leq C(\theta, \tilde{\theta}) e^{-\tilde{\theta} x}, \quad \text{for } x \geq 0, \ 0 \leq k \leq K.$

Proof. By Lemma 2.3 there exists an invertible coordinate transformation
(3.6) $W = (I + \Theta) Z$
with
(3.7) $|\partial_x^k \Theta| \leq C(k) e^{-\tilde{\theta} x} \quad \text{for } x \geq 0,$
converting (3.1), (3.4) to
(3.8) $Z' - A_+ Z = 0 \quad \text{for } x \geq 0$
and
(3.9) $\tilde{\Pi}_0 Z(0) = \tilde{\alpha}, \quad \lim_{x \to +\infty} \Pi_+ (Z(x) - V_+ ) = 0,$
where $\tilde{\Pi}_0 := (I + \Theta(0))^{-1} \Pi_0 (I + \Theta(0))^{-1}$ is again arbitrary and $\tilde{\alpha} := (I + \Theta(0))^{-1} \alpha$. By inspection, there exists a unique solution if and only if Ker$\tilde{\Pi}_0$ is transverse to $\Sigma_+$, i.e., there holds the generically satisfied Evans/Lopatinski condition
(3.10) $\det \tilde{\Pi}_0 \tilde{\Sigma}_+ \neq 0,$
where $\tilde{\Sigma}_+$ is the complementary (rank $n - k$) invariant subspace of $\Sigma_+$, consisting of the sum of the constant solution $Z \equiv V_+$ and the solution of the Cauchy problem $Z(0) = \tilde{\alpha} - \tilde{\Pi}_0 V_+$ under the flow of the constant-coefficient equation (3.8) restricted to the subspace $\tilde{\Sigma}_+$ of exponentially decaying solutions. \hfill \Box

3.2 Discrete problem

We now consider a discretized version of (3.1), (3.4) on the truncated domain $x \in [0, M]$ with the corresponding projective boundary conditions
(3.11) $\Pi_0 W(0) = \alpha, \quad \Pi_+ (W(M) - V_+ ) = 0,$
and examine convergence as $M \to +\infty$ and mesh size goes to zero of the approximate to the exact solution.

Remark 3.2. In view of the discussion of the introduction, our main interest in the context of Evans function computations is the “Dirichlet” case $\Sigma_+ = \mathbb{C}^n$ arising for one-sided schemes suitable for shooting methods: in particular, the centered exterior-project method or the polar coordinate method linearized about a desired exact solution. In this more favorable case, there is no boundary condition at $x = 0$, and no arbitrary projection $\Pi_0$, and the projective boundary conditions (3.11) reduce to the simple Dirichlet condition $W(M) = V_+.$
3.2.1 Difference scheme

We assume a general linear difference scheme of form

\[(3.12)\]
\[SW = 0,\]

with boundary conditions

\[(3.13)\]
\[\Pi_0 W_0 = \alpha, \quad \Pi_+(W_J - V_+) = 0,\]

where \(W = (W_1, \ldots, W_J)\) are approximations of \(W\) at mesh points \(x_j, j = 0, \ldots, J\) and \(S\) is a linear difference operator with finite stencil \(\{j - \ell, j + \ell\}\), i.e., the value of \((SW)_j\) depends on \(W\) only through \(\{W_{j-\ell}, \ldots, W_{j+\ell}\}\). Moreover, we assume that \((SW)_j\) depends on \(A\) linearly and only through the restriction of \(A\) to the interval \([x_{j-\ell}, x_{j+\ell}]\). We define boundary and truncation errors \(\varepsilon_0, \varepsilon_J\) and \(\tau = (\tau_0, \ldots, \tau_J)\) as usual by

\[(3.14)\]
\[SW = h_j \tau, \quad \Pi_0 W_0 - \alpha = \varepsilon_0, \quad \Pi_+(W_J - V_+) = \varepsilon_J,\]

where \(W_J := W(x_J)\) denotes the solution of the continuous problem (3.1), (3.4) on \([0, +\infty)\) sampled at mesh points \(x_j\), and \(h_j\) is the \(j\)th mesh length, defined as the maximum difference between points \(x_k\) involved in the evaluation of \(S_j\).

This could be a boundary-value scheme, with \(S\) realized as a large \(Jn \times Jn\) banded matrix. Or, in the case of our main interest \(\Sigma = \mathbb{C}^n\) that boundary conditions are imposed only at one end \(x = M\), it could be a backward Cauchy solver such as RK45, with \(S\) realized as a series of \(J - \ell\) successive \((2\ell + 1)n \times (2\ell + 1)n\) matrix multiplications.

3.2.2 Basic assumptions

We do not specify the details of the scheme, other than to make certain mild assumptions.

Assumptions 3.2.  (i) \(k\)-th order consistency: As mesh size \(h_j \to 0\),

\[(3.15)\]
\[|\tau_j| \leq C h_j^k \sup_{z \in [x_{j-\ell}, x_{j+\ell}]} |\partial_x^{k+1} W(z)| \to 0.\]

(ii) Strict constant-coefficient stability: In the constant-coefficient case \(A \equiv A_+\), if \(\Re \sigma(A_+) \leq \mu\), then, for any \(\tilde{\mu} > \mu\), Cauchy information \(SW = \tau h\) and \(W_0 = \varepsilon_0\) imply (under appropriate mesh restrictions)

\[(3.16)\]
\[|\tilde{\Delta}^r W_j| \leq C \sup_{j-r \leq m \leq j} |h_m|^r (\varepsilon_0 e^{\tilde{\mu} x_j} + \sum_{0 \leq k \leq j} e^{\tilde{\mu}(x_j - x_k)} \tau_k h_k)\]

for \(0 \leq r \leq 2\), where \(\tilde{\Delta}\) is the backward difference operator \((\tilde{\Delta} f)_j := f_j - f_{j-1}\). Likewise, if \(\Re \sigma(A_+) \geq \mu\), then, for any \(\tilde{\mu} < \mu\), \(SW = \tau h\) and \(W_j = \varepsilon_J\) imply

\[(3.17)\]
\[|\tilde{\Delta}^r W_j| \leq C \sup_{j-r \leq m \leq j} |h_m|^r (\varepsilon_J e^{\tilde{\mu}(x_j - x)} + \sum_{j \leq k \leq j} e^{\tilde{\mu}(x_j - x_k)} \tau_k h_k).\]
In the case $\Sigma = \mathbb{C}^n$ of our main interest, we require only (3.17) and only for $\mu \leq 0$.

As is customary, we ignore machine error.

**Remark 3.3.** Condition (i) is of course standard. In the situation of our main interest of a backward Cauchy solver in the case $\Sigma = \mathbb{C}^n$, condition (ii) (in this case, the single condition (3.17) for $r = 0$) is equivalent by Duhamel’s principle/variation of constants to the homogeneous stability condition

$$|W_j| \leq C\varepsilon_k e^{\tilde{\mu}(x_j - x_k)}$$

for $x_j < x_k$, $\tilde{\mu} \leq 0$, when $SW = 0$ and $W_k = \varepsilon_k$, which is related to a circle of ideas including $A$-stability and one-sided Lipshitz bounds [D1, D2, HNW1, HNW2] regarding approximation of stiff ODE. For general schemes it may impose impractically severe limitations on the mesh size; however, it is satisfied without such conditions for RK45 and other “nice” schemes, as may be seen by applying a constant linear coordinate transformation $A \rightarrow \tilde{A} := SAS^{-1}$ for which $\Re \tilde{A} := (1/2)(\tilde{A} + \tilde{A}^*) \geq \tilde{\mu} > \mu$ with $\tilde{\mu}$ arbitrarily close to $\mu$, then applying the results of [D1, D2, HNW1, HNW2]. For some simple examples, see Remark 3.8. Likewise, for Cauchy solvers (ii) ($r = 1, 2$) follows using the difference scheme from (ii) ($r = 0$). For general boundary-value schemes, (ii) must be checked on a scheme-by-scheme basis. For general theory on stability of boundary-value schemes, see, e.g., [K1, Be1, Be2].

**Remark 3.4.** Evidently, (ii) implies also the dual bounds

$$|W_j| \leq C(\varepsilon_0 e^{\tilde{\mu}x_j} + \sup_{j-r \leq m \leq j} \sum_{0 \leq k \leq j} h_m |e^{\tilde{\mu}(x_j - x_k)}\tau_k h_k|)$$

and

$$|W_j| \leq C(\varepsilon_J e^{\tilde{\mu}(x_j - x_J)} + \sup_{j-r \leq m \leq j} \sum_{j \leq k \leq J} e^{\tilde{\mu}(x_j - x_k)}\tau_k h_k)$$

for differenced data $SW = \Delta^r(\tau h)$, $0 \leq r \leq 2$, where $\Delta$ denotes the forward difference operator $(\Delta f)_j := f_{j+1} - f_j$.

### 3.3 Discrete conjugation error

We now make the key observation that the conjugating transformation for the continuous problem, up to a small commutation error, is a conjugator for the discrete problem as well.

**Example 3.5.** Consider the first-order (forward explicit/backward implicit) Euler scheme

$$(SW)_j := W_{j+1} - W_j - h_j A_j W_j = 0.$$

\[4\] This is all that is needed in any case for (3.17), but this seems to give no advantage in the general case since there remains the more restrictive assumption (3.16) in the other direction.
Substituting $W_j =: P_j Z_j$, $P_j := P(x_j)$, where $P = I + \Theta$ is the continuous conjugator of (3.6), we obtain after a brief calculation

$$(\tilde{S} Z)_j := Z_{j+1} - Z_j - h_j A_{j+1} Z_j = (\Psi Z)_j,$$

where $\tilde{S}$ is the realization of the same first-order Euler scheme to the constant-coefficient case $A = A_+$ and $\Psi$ is the commutator error

$$(\Psi Z)_j := -h_j P_j^{-1} \left( P_{j+1} - P_j - h_j(A_j P_j - P_{j+1} A_{j+1}) \right) Z_j.$$

Noting that $P_{j+1} - P_j = h_j(A_j P_{j+1} - P_{j+1} A_{j+1})$ is a discretization of the ODE $P' = AP - PA_+$ defining $P$, (2.13), we find similarly as in (3.15) that the exact solution $P$ has truncation error

$$\hat{\tau}_j := h_j^{-1} \left( P_{j+1} - P_j - h_j(A_j P_j - P_{j+1} A_{j+1}) \right) = O(|h_j| \sup_{x \in [x_j-\ell, x_{j+\ell}]} |P'|) = O(|h_j| e^{-\theta x_j}),$$

we find that

(3.20) \quad |(\Psi Z)_j| \leq C|h_j|^2 e^{-\theta x_j} |Z_j|.

**Remark 3.6.** Note that the righthand side of (3.20) is smaller by a factor $h_j$ than the corresponding estimate

$$(\tilde{S} W)_j = O(|h_j| \sup_{j-\ell \leq m \leq j+\ell} |A_m - A_{m+1}|) \sup_{j-\ell \leq m \leq j+\ell} |W_m| \leq C|h_j| e^{-\theta x_j} \sup_{j-\ell \leq m \leq j+\ell} |W_m|,$$

obtained by separating out constant-coefficient and exponentially decaying parts in the original equation for $W$. This additional factor is crucial in the contraction argument of (3.4).

**Example 3.7.** Consider the second-order (forward/backward implicit) midpoint scheme

(3.21) \quad (SW)_{j+1} := W_{j+1} - W_{j-1} - (x_{j+1} - x_{j-1}) A_j W_j = 0.

Substituting $W_j =: P_j Z_j$, $P = I + \Theta$ as in (3.6), we obtain

$$(\tilde{S} Z)_j := Z_{j+1} - Z_{j-1} - (x_{j+1} - x_{j-1}) A_{j+1} Z_j = (\Psi Z)_j + \Delta(\Psi^1 Z)_j$$

where $\tilde{S}$ is the constant-coefficient realization of $S$, $\Delta$ is the forward difference operator $(\Delta f)_j := f_{j+1} - f_j$, and

$$(\Psi Z)_j := -(x_{j+1} - x_{j-1}) P_j^{-1} \left( P_{j+1} - P_j - (x_{j+1} - x_{j-1})(A_j P_j - P_{j+1} A_{j+1}) \right) Z_j + P_j^{-1} \left( \frac{P_{j+1} + P_{j-1}}{2} - P_j \right) (x_{j+1} - x_{j-1}) A_j Z_j + \left( \frac{1}{2} \right) \Delta^2 \left( P_j^{-1}(P_{j+1} - P_{j-1}) \right) Z_{j-1},$$

$$(\Psi^1 Z)_j := \left( P_j^{-1}(P_{j+1} - P_{j-1}) \right) (Z_j - Z_{j-1}).$$
Arguing as in the previous example, we obtain similarly as in (3.20) the bound
\[
|\langle \Psi Z \rangle_j| \leq C|h_j|^2 e^{-\tilde{\theta}x_j} \sup_{j-1 \leq m \leq j} |Z_m|.
\]
and, likewise,
\[
|\langle \Psi^1 Z \rangle_j| \leq C|h_j| e^{-\tilde{\theta}x_j} \sup_{j-1 \leq m \leq j} |Z_m|.
\]

Remark 3.8. The implicit backward Euler scheme of Example 3.5 is \(A\)-stable as a Cauchy solver in the backward \(x\) direction, hence satisfies Assumption 3.2 in the one-sided case \(\Sigma_+ = \mathbb{C}^n, \hat{\mu} = 0\), with no assumption on the mesh size. For example, in the scalar case
\[
W_j+1 = W_j + h_j a W_j \text{ yields immediately } W_j = (I + h_j a)^{-1} W_j+1, \text{ for any } h_j > 0.
\]
The implicit Midpoint method viewed as a two-step backward scheme has characteristic roots \(-ah_j \pm \sqrt{(ah_j)^2 + 1}\) that remain strictly \(\geq 1\) independent of \(h_j > 0\), but is not backward \(A\)-stable as a Cauchy solver.

Generalizing the results of the examples, we make the following final assumption on the scheme, which appears to be satisfied in most if not all cases of interest.

Assumption 3.3. Under the change of coordinates \(W_j =: P_j Z_j\), \(P\) defined as in (3.6), equation (3.1) becomes
\[
\tilde{S} Z = \Psi^0 Z + \Delta(\Psi^1 Z),
\]
where \(\tilde{S}\) is the same discretization scheme used for \(W\) applied to the constant-coefficient case \(A = A_+\), \(\Delta\) is the forward difference operator, and
\[
|\langle \Psi^r Z \rangle_j| \leq C(\theta, \tilde{\theta})|h_j|^{1-r} e^{-\tilde{\theta}x_j} \sup_{j-\ell-1 \leq m \leq j+\ell} |Z_m|.
\]

3.4 Numerical convergence lemma

With these preparations, it is straightforward to establish our following main result.

Theorem 3.9. Assuming (3.2), (3.3), (3.10), and Assumptions 3.1, 3.2, and 3.3 for fixed \(0 < \tilde{\theta} < \theta\), for \(\sup_j |h_j|\) sufficiently small and \(M > 0\) sufficiently large, there exists a unique solution \(W\) of (3.12), (3.13), which, moreover, satisfies for \(C > 0\) independent of \(M, \tau\),
\[
|W_j - \bar{W}(x_j)| \leq C(e^{-\hat{\theta}M} + \sup_j |\tau_j|) \text{ for all } j = 0, \ldots, J,
\]
where \(\bar{W}\) is the solution of (3.1), (3.4) guaranteed by Lemma 3.1 and \(\theta\) is as in (3.2). For \(\Sigma_+ = \mathbb{C}^n\), the mesh condition \(\sup_j |h_j| \ll 1\) can be relaxed to \(\sup_j |h_j| e^{(-\tilde{\theta} - \tilde{\nu})x_j} \ll 1\), where \(-\tilde{\theta} < \tilde{\nu} \leq 0\) is less than the smallest real part of the eigenvalues of \(A_+\).
That is, we assert existence and convergence so long as the spectral gap \( \min \Re \sigma (A_+|_{\Sigma_+}) \) associated with the boundary projection \( \Pi_+ \) at \(+\infty\) is greater than \(-\theta\), where \( \theta \) as in (3.2) is the exponential rate of convergence of \( A \) to \( A_+ \) as \( x \to +\infty \). On the other hand, standard theory \( [Be1] \) requires a positive spectral gap \( \beta \) and concludes convergence at rate \( Ce^{-\beta M} \) for \( 0 < \beta < \beta_0 \). In other words, though it fails the positive gap condition of standard theory, the problem remains numerically well-conditioned so long as “badness” of the boundary condition as measured by negativity of the spectral gap is less than the rate of exponential convergence, in exact analogy with the gap lemma of continuous theory \( [GZ, KS, ZH] \).

Proof. By Assumption 3.3, \( \tilde{S}Z = \sum_{r=0}^{2} \Delta^r \Psi^r Z \) and \( \tilde{S}\tilde{Z} = \sum_{r=0}^{2} \Delta^r \Psi^r \tilde{Z} + F \), where \( W_j =: P_j Z_j, \tilde{W}_j = W(x_j) =: P_j \tilde{Z}_j, \) and \( h_j \tau_j =: F_j \). Defining the convergence error \( E := Z - \tilde{Z} \), we thus obtain the error equation

\[
\tilde{S}E = \sum_{r=0}^{2} \Delta^r \Psi^r E - F,
\]

with boundary conditions

\[
\tilde{\Pi}_0 E_0 = 0 \quad \text{and} \quad \tilde{\Pi}_+(E_j) = -\tilde{\Pi}_+(\tilde{Z}_j - \tilde{V}_+) = O(e^{-\beta M}).
\]

Denoting by \( \tilde{\mathcal{E}} = T(E) \) the solution of \( \tilde{S}\tilde{\mathcal{E}} = \sum_{r=0}^{2} \Delta^r \Psi^r \mathcal{E} - F, \) (3.29), we thus obtain the fixed-point formulation

\[
\mathcal{E} = T(E)
\]

for the solution of (3.28)-(3.29), and thus for the solution \( Z = \tilde{Z} + \mathcal{E} \) of the original problem.

Applying bounds (3.18)-(3.19) of Remark 3.4 together with (3.26), (3.29), and (3.15), and the principle of linear superposition, we obtain

\[
|\tilde{\mathcal{E}}_j| \leq C(\varepsilon_0 e^{\beta_j} + \sum_{0\leq k \leq j} e^{\beta(x_j-x_k)}[(\Psi^0 \mathcal{E})_k] + \sup_{j-1 \leq m \leq j} |h_m|[(\Psi^1 \mathcal{E})_k]|)
+ \sum_{0\leq k \leq j} e^{\beta(x_j-x_k)} \tau_k h_k)
+ C(\varepsilon_0 e^{\beta_j} + \sum_{j \leq k \leq j} e^{\beta_j(x_j-x_k)}[(\Psi^0 \mathcal{E})_k] + \sup_{j-1 \leq m \leq j} |h_m|[(\Psi^1 \mathcal{E})_k]|)
+ \sum_{j \leq k \leq j} e^{\beta_j(x_j-x_k)} \tau_k h_k),
\]

\( \Psi := (\Psi^0, \Psi^1, \Psi^2) \), where \( \beta < 0 \) is greater than the largest real part of the eigenvalues of \( A_+ \) associated with the invariant subspace \( \tilde{\Sigma}_+ \) complementary to \( \Sigma_+ \), and \( -\tilde{\theta} < \bar{\nu} < \nu \leq 0 \) is less than the smallest real part of the eigenvalues of \( A_+ \) associated with \( \Sigma_+ \); see Assumption 3.4. From (3.31), we readily obtain by a discrete version of calculation (2.8) in the argument of Lemma 2.1 the a priori estimate

\[
\sup_{0 \leq j \leq J} |\mathcal{E}_j| \leq C(e^{-\beta M} + \sup_j |\tau_j| + \sup_j h_j \sup_{0 \leq j \leq J} |\mathcal{E}_j|),
\]

19
which for sup \( j \) \( |h_j| \) sufficiently small implies

\[
(3.33) \quad \sup_{0 \leq j \leq J} |\mathcal{E}_j| \leq C\left(e^{-\bar{\theta}M} + \sup_j |\tau_j| \right),
\]

hence, by boundedness of \( |P_j| \) and \( W_j - \bar{W}(x_j) := P_j \mathcal{E}_j \), the desired bound \((3.27)\). A similar estimate yields contractivity of \( T \) in the sup-norm, hence existence and uniqueness in the class \( \ell^\infty[0, J] \).

Finally, observing in case \( \Sigma_+ = C^n \) that only the \( \sum_{j \leq k \leq J} \) terms in \((3.31)\) appear, we may bound \( \sup_{j-1 \leq m \leq j} |h_m|(|\Psi_1 \mathcal{E})_k| \) using \( j \leq k \) by

\[
\sup_{j-1 \leq m \leq j} |h_m| \sup_{k-\ell-1 \leq m \leq k+\ell} e^{-\bar{\theta}x_m} |h_m| \leq \sup_{0 \leq j \leq J} \left(e^{-\bar{\theta}-\bar{\nu}x_j} |h_j| \right) \sup_{0 \leq j \leq J} \left(e^{-\bar{\theta}x_m} |h_m| \right)
\]

and similarly for \( |(\Psi \mathcal{E})_k| \), to obtain by the same argument

\[
\sup_{0 \leq j \leq J} |\mathcal{E}_j| \leq C\left(e^{-\bar{\theta}M} + \sup_j |\tau_j| + \sup_{0 \leq j \leq J} \left(e^{-\bar{\theta}-\bar{\nu}x_j} |h_j| \right) \sup_{0 \leq j \leq J} |\mathcal{E}_j| \right),
\]

in place of \((3.32)\), yielding existence, uniqueness, and convergence under the relaxed mesh condition \( \sup_j |h_j|e^{-\bar{\theta}-\bar{\nu}x_j} \ll 1 \).

Remark 3.10. By Remark 3.8, Theorem 3.9 applies in particular to the one-sided, or Dirichlet, case \( \Sigma_+ = C^n \), with the implicit backward Euler method of Example 3.5.

Remark 3.11. Since \( -\bar{\theta} - \bar{\nu} < 0 \) by Assumption 3.7 in the case \( \Sigma_+ = C^n \) of our main interest, the mesh and truncation error conditions

\[
\sup_j |h_j|e^{-\bar{\theta}-\bar{\nu}x_j} \ll 1 \quad \text{and} \quad \sup_j |\tau_j| \leq \sup_j |h_j| e^{-\bar{\theta}x_j} \ll 1
\]

required for convergence allow for arbitrarily large step size as \( x_j \to +\infty \), hence our result indeed applies to shooting-type schemes with adaptive step size such as are used in practice.

Remark 3.12. The method of proof in the argument of Theorem 3.4, based on Lemma 2.3, is designed to handle the general boundary-value case. For shooting methods, there is a much simpler proof on \([L, M]\) with \( 1 \ll L \leq M \), as in the proof of Lemma 2.1, just separating constant-coefficient and exponentially decaying parts of \( A \) in the original \( W \) equation and obtaining contraction as in the proof of Lemma 2.1 from largeness of \( L \). The extension to the full domain \([0, M]\) then follows similarly as in the continuous case by standard convergence results for the Cauchy problem on a bounded domain. The conjugation argument is used to obtain contraction for globally defined schemes on the full domain \([0, M]\).
3.5 Convergence of the centered exterior product method

The centered exterior product method as described in the introduction consists of solving

\[ W' = (A(x, \lambda) - \mu S) W, \]

from \( x = +\infty \) to \( x = 0 \) for \( W^+ := W_1^+ \wedge \cdots \wedge W_k^+ \), where the linear operator \( A \) is defined by the Leibnitz formula

\[ AW_1 \wedge \cdots \wedge W_k := (AW_1 \wedge W_2 \wedge \cdots \wedge W_k) + \cdots + (W_1 \wedge \cdots \wedge AW_k) \]

with eigenvectors and eigenvalues \( \mathcal{R} = r_1 \wedge \cdots \wedge r_k \), \( \mu = a_1 + \cdots + a_k \), where \( r_j \) and \( a_j \) are eigenvectors of \( A \), and \( \mu S \) is the sum of the eigenvalues of \( A_+ \) associated with the \( k \)-dimensional stable subspace \( S_+ \) of \( A_+ \) obtained as the wedge product of a basis of \( S_+ \); solving the symmetric equation from \( x = -\infty \) to \( x = 0 \) for a solution asymptotic at \( -\infty \) to an eigenvector \( \mathcal{R} S_+ \) of \( A_+ \) associated with the unstable subspace \( S_- \) of \( A_- \); then evaluating the Evans function following (1.6) as

\[ D(\lambda) := \langle (W^+ \wedge W^-) | x=0 \rangle, \]

\[ W^- := W_1 \wedge \cdots \wedge W_{n-k}^- \], where \( \langle \cdot \rangle \) denotes coordinatization in the standard Euclidean basis, i.e. \( \eta =: \langle \eta \rangle (e_1 \wedge \cdots \wedge e_n) \) for an \( n \)-form \( \eta \).

Equation (3.34) is of the form (3.1) considered in Section 3.1, so that we may apply our just-developed theory. Moreover, as \( \mu S_+ \) and \( \mu S_- \) respectively are the smallest real part and largest real part eigenvalues of \( A_+ \) and \( A_- \), we are in the more favorable one-sided, or “Dirichlet” case \( \Sigma_{\pm} = \mathbb{C}^N_{\pm} \) suitable for shooting methods, where \( N_+ := \binom{n}{k} \) and \( N_- := \binom{n-k}{n} \) are the dimensions of the systems for \( W^+ \) and \( W^- \). We may thus approximate the solution as described in Section 3.2 by solving a pair of finite difference schemes (3.12), on \([0, M]\) and \([0, -M]\), respectively, discretized as \( 0 = x_0 < x_1 < \cdots < x_J = M \) and \( 0 = x_0 > x_{-1} > \cdots > x_{-J} = -M \), with Dirichlet boundary conditions

\[ W^+_J = \mathcal{R} S_+, \quad W^-_{-J} = \mathcal{R} S_- \]

determining thereby a numerically approximated Evans function

\[ D^{M, h}(\lambda) := \langle (W^+_0 \wedge W^-_0) \rangle, \]

where \( h = (h_{-J}, \ldots, h_J) \) is the vector of mesh blocks used to discretize \([-M, M]\).

In the above discussion, we have implicitly assumed the consistent splitting hypothesis of [AGJ]: that the dimensions of the stable subspace of \( A_+ (\lambda) \) and the unstable subspace of \( A_- (\lambda) \) sum to full rank \( n \) on the subset of \( \lambda \) under consideration. By standard considerations [He1, GZ, Z1], the “region of consistent splitting” on which this holds typically includes the component of real \( +\infty \) in the complement of the essential spectrum of the associated linearized differential operator \( L \) of (1.1), whose point spectra the Evans function is designed

\[ \text{See Section 6.3.1 for numerical prescriptions of } \mathcal{R} S_{\pm}. \]
to determine. However, as pointed out in [GZ, ZH], it is sometimes useful to extend this region of investigation and study also eigenvalues embedded in the essential spectrum.

Following [GZ, ZH], we thus consider the problem in a slightly more general setting, substituting in place of consistent splitting the following gap assumption.

**Assumption 3.4.** On \( \Lambda \subset \mathbb{C} \), the spaces \( S_+ \) and \( S_- \) are invariant subspaces of \( A_+ \) and \( A_- \), analytic in \( \lambda \), with dimensions \( k \) and \( (n - k) \). Moreover, the spectral gaps \( \nu_+ \) and \( \nu_- \) defined as the maximum of the difference between the smallest real part of the eigenvalues of \( A_+ \) not associated with \( S_+ \) and the largest real part of those associated with \( S_+ \) and the maximum of the difference between the smallest real part of the eigenvalues of \( A_- \) associated with \( S_- \) and the largest real part of those not associated with \( S_- \), satisfy

\[
\nu_j > -\theta, \quad j = \pm,
\]

where \( \theta > 0 \) as in (3.2) is the exponential rate of convergence of \( A \) to \( A_\pm \) as \( x \to \pm\infty \).

Applying Theorem 3.9 in this context, we obtain the following convergence result.

**Corollary 3.13.** Under Assumptions 3.2, 3.3, and 3.4, for fixed \( 0 < \tilde{\theta} < \theta \), for \( M > 0 \) sufficiently large and \( \sup_j |h_j| e^{(-\tilde{\theta} - \tilde{\nu})x_j} \) sufficiently small, where \( -\tilde{\theta} < \tilde{\nu} \leq 0 \) is less than \( \min \nu_\pm \), there exist unique solutions \( W_\pm \) of (3.1), (3.37) determining an approximate Evans function \( D^{M,h}(\lambda) \) satisfying

\[
|D^{M,h} - D| \leq C(e^{-\tilde{\theta}M} + \sup_j |\tau_j|)
\]

uniformly on compact subsets of \( \Lambda \), where \( D \) is constructed following (3.36) from the solutions \( \hat{W}_\pm \) of (3.34) guaranteed by Lemma 3.1, \( \theta \) is as in (3.2), and \( \tau_j \) is truncation error.

**Proof.** Immediate, observing that Assumption 3.4 implies Assumption 3.1.

### 3.5.1 Mesh requirements, and computations in the essential spectrum

So long as we remain in the region of consistent splitting (see discussion above Assumption 3.4), i.e., away from the essential spectrum of the operator \( L \) whose point spectra we seek to study, we have \( \Re \sigma A \geq 0 \), with a simple eigenvalue at zero, from which we may obtain the stability property (3.17) of (ii) needed for the convergence proof, with value \( \tilde{\mu} = 0 \), even though the statement of (ii) is not strictly satisfied. Indeed, this situation holds for general difference schemes in case \( \Sigma_+ = \mathbb{C}^n \) whenever \( \Re \sigma A_+ \geq 0 \) and zero is a semisimple eigenvalue of \( A_+ \), yielding the same convergence results stated in Theorem 3.4 and Corollary 3.13. Moreover, the spectral gaps \( \nu_\pm \) of Assumption 3.4 are identically zero, and \( \tilde{\nu} \) (by semisimplicity) may be taken as zero as well.

Together with Remark 3.3, this shows that, for shooting methods based on an \( A \)-stable Cauchy solver (e.g., RK45), there is no requirement on the mesh size \( |h_j| \) beyond the requirement \( \sup_j |h_j| e^{(-\tilde{\theta} - \tilde{\nu})x_j} \leq \sup_j |h_j| e^{-\tilde{\theta}x_j} \) sufficiently small stated explicitly in the
Theorem (Corollary). This agrees with observations in [Br, BrZ, HuZ1, BHRZ, HLZ, CHNZ] in which centered exterior-product computations away from the essential spectrum are observed to require an extremely sparse mesh.

On the other hand, for \( \lambda \) inside the essential spectrum, one or more of the gaps \( \nu \pm \) becomes negative and \( \Re \sigma A \pm \) are no longer of one sign. As suggested by the simple computations of Remark 3.8, this might result in a much stricter requirement on \( |h_j| \) in order to satisfy condition (ii) (with \( \tilde{\mu} \) now strictly positive). Whether this is a real effect or just an artifact of our analysis is not clear, but this would be an interesting issue for further investigation.

A second interesting question, for general centered schemes, would be the relation between the dimension of the kernel of \( A \pm \) and the size of the coefficient \( C \) in convergence estimate (3.27); we conjecture that they are roughly proportional, information that could be useful in comparing schemes.

4 Stability of continuous orthogonalization

We next address stability of the continuous orthogonalization method. Consider a solution \( \Omega \) of the continuous orthogonalization system (1.8)(i) restricted to the Stiefel manifold \( S = \{ \Omega : \Omega^* \Omega = I_k \} \), such that \( \bar{\Omega} \to \Omega_+ \) as \( x \to +\infty \). Evidently, the columns of \( \Omega_+ \) are an orthonormal basis for an invariant subspace \( \Sigma_+ \) of \( A_+ = \lim_{x \to +\infty} A(x) \). Of particular interest is the case arising in Evans function computations that \( \Sigma_+ \) is the stable or (at certain boundary points) a neutrally-stable subspace of \( A_+ \). Linearizing (1.8)(i) about \( \bar{\Omega} \), we obtain the linearized system

\[
\Omega' = (I - \bar{\Omega} \bar{\Omega}^*) A \Omega + \bar{\Omega} \bar{\Omega}^* A \bar{\Omega} - \bar{\Omega} \bar{\Omega}^* A \bar{\Omega},
\]

for which the limiting constant-coefficient equation as \( x \to +\infty \) is

\[
(4.1) \quad \Omega' = L \Omega := (I - \Omega_+ \bar{\Omega}_+^*) A \Omega - \Omega_+ \bar{\Omega}_+^* A \Omega - \Omega_+ \bar{\Omega}_+^* A \Omega_+.
\]

We wish to assess the backward stability of (4.1) with respect to general perturbations, both along the tangent manifold of the Stiefel manifold \( S \) and in transverse directions.

**Remark 4.1.** Properly speaking, (4.1) is linear in the pair of variables \( (\Omega, \Omega^*) \) and not \( \Omega \) alone, since matrix adjoint is not a linear operation. Along the tangent manifold, however, it is linear as we shall see in \( \Omega \) alone.

4.1 Asymptotic stability in tangential directions

The tangent manifold to \( S \) at \( \Omega_+ \) consists of directions \( \Omega \) such that \( D(\Omega^* \Omega)_{\Omega_+} \Omega = 0 \), or

\[
(4.2) \quad \Omega_+^* \Omega + \Omega_+^* \Omega_+ = 0,
\]

\[\text{In any case, this would presumably be confined to shooting methods.}\]
that is, for which $\Omega^*_+,\Omega$ is skew-symmetric. As the Stiefel manifold is invariant under (1.8)(i), this is evidently invariant under (4.1), as direct computation readily verifies. It is spanned by the direct sum of eigenvectors $\Omega^*_+, K, K$ skew, in the kernel of $L$ and eigenvectors

$$\begin{equation}
\Omega_{jk} := (I - \Omega^*_+ \Omega^*_+)r_j \sigma_k^*
\end{equation}$$

in the kernel of $\Omega^*_+$, where $r_j$ run through the eigenvectors of $A$ transverse to $\Sigma^+_+$ and $\sigma_k \in \mathbb{C}^k$ are left eigenvectors of $\alpha$ defined by $\alpha \Omega^*_+ := A^+ \Omega^*_+$. The former correspond to rotations within the same invariant subspace, the latter to perturbations outside $\Sigma^+_+$.

Under (4.2), we readily find that (4.1) simplifies to

$$\begin{equation}
\Omega' = L\Omega := (I - \Omega^*_+ \Omega^*_+^*)(A\Omega - \Omega \alpha),
\end{equation}$$

which, for eigenvectors (4.3) yields

$$\begin{equation}
L\Omega_{jk} = \Omega_{jk}(a_j - a_k),
\end{equation}$$

where $a_j$ and $a_k$ are the eigenvalue associated with $r_j$ and $\sigma_k$. Thus, the eigenvalues along the Stiefel manifold are zero ($k(k+1)/2$-fold) and $a_j - a_k$ ($(n-k) \times k$ in total), where $a_k$ belong to $\Sigma^+_+$ and $a_j$ to the complementary invariant subspace of $A^+$. In particular, when $\Sigma^+_+$ is the stable subspace of $A^+$, the Stiefel eigenvalues of $L$ have either zero or positive real part, and so (4.1) restricted to the tangent space, as claimed in the introduction, is (neutrally) stable in backward $x$, at least for the limiting system as $x \to +\infty$.

**Remark 4.2.** Reduced equation (4.4) is linear in $\Omega$ alone, since it does not involve $\Omega^*$.

### 4.2 Asymptotic stability in transverse directions

Off the Stiefel manifold, we face the difficulty pointed out in Remark 4.1 that $L$ strictly speaking is a linear function of $\Omega$ and $\Omega^*$, so that (4.1) actually represents a pair of coupled equations, complicating calculations. However, we can sidestep much of this difficulty by noting that the remaining modes not already treated are of form $\Omega^+\beta$, where $\beta \in \mathbb{C}^{k \times k}$ by a brief calculation satisfies

$$\begin{equation}
\beta' = -(\beta + \beta^*)\alpha,
\end{equation}$$

$A^+\Omega^+_+ =: \Omega^+_+\alpha$. Introducing $R := R\beta := (1/2)(\beta + \beta^*)$, we thus have the linear system

$$\begin{equation}
R' = -R\alpha - \alpha^* R,
\end{equation}$$

$$\begin{equation}
\beta' = -2R\alpha,
\end{equation}$$

which has block-triangular form

$$\begin{equation}
\begin{pmatrix}
R \\
\beta
\end{pmatrix}' = 
\begin{pmatrix}
\mathcal{M} & 0 \\
\mathcal{N} & 0
\end{pmatrix}
\begin{pmatrix}
R \\
\beta
\end{pmatrix},
\end{equation}$$

24
where \( MR := -\mathcal{R} \alpha - \alpha^* \mathcal{R} \) and \( NR := -2\mathcal{R} \alpha \).

This has the \( k^2 \)-dimensional kernel \( \mathcal{R} = 0 \) already identified in Section 4.1, and \( k^2 \)

eigenvectors

\[
(4.8) \quad \frac{\mathcal{R}_{jk}}{NR_{jk}/\mu_{jk}}, \quad \mathcal{R}_{jk} := l_j l_k^*,
\]

with eigenvalues \( \mu_{jk} := -(a_j^* + a_k) \), where \( l_j \) are left eigenvalues of \( \alpha \) and \( a_j \) the associated

eigenvalues, which are exactly the eigenvalues of \( A_+ \) restricted to \( \Sigma_+ \).

In particular, when \( \Sigma_+ \) is the stable subspace of \( A_+ \), the transverse eigenvalues of \( \mathcal{L} \) have

either zero or positive real part, and so the Stiefel manifold as indicated in the introduction, is (neutrally) stable in backward \( x \), at least for the limiting system as \( x \to +\infty \).

**Remark 4.3.** Alternatively, we may follow the simpler argument of the introduction to reach the same conclusion without finding explicitly the eigenmodes of the system.

### 4.3 Convergence of the polar coordinate method

Consider now the polar coordinate method, consisting of approximation of (1.8) on \([-M, 0]\) and \([0, M]\) by forward (resp. backward) Cauchy solvers, under Dirichlet boundary (i.e., Cauchy) conditions

\[
(4.9) \quad W^+_j = (\Omega_+, \log \tilde{\gamma}_+), \quad W_j^- = (\Omega_+, \log \tilde{\gamma}_+),
\]

where \( W^+_j \) are the numerical approximations of \((\Omega^+, \log \tilde{\gamma}_+)(x_j)\).

**Corollary 4.4.** Under Assumptions 3.2, 3.3, and 3.4, for fixed \( 0 < \tilde{\theta} < \theta \), for \( M > 0 \) sufficiently large and \( \sup_j |h_j| e^{(\tilde{\theta} - \tilde{\nu})x_j} \) sufficiently small, where \( -\tilde{\theta} < \tilde{\nu} \leq 0 \) is less than \( \min \nu_\pm \), there exist unique solutions \( W^\pm \) of the discretization of (1.8) with boundary conditions (4.9) determining an approximate Evans function \( D^{M,h}(\lambda) \) satisfying

\[
(4.10) \quad |D^{M,h} - D| \leq C(e^{-\tilde{\theta}M} + \sup_j |\tau_j|)
\]

uniformly on compact subsets of \( \Lambda \), where \( D \) is constructed following (3.36) from the solutions \( W^\pm \) of (3.34) guaranteed by Lemma 3.1, \( \theta \) is as in (3.2), and \( \tau_j \) is truncation error.

**Proof.** Equivalently, we must establish the analog of Theorem 3.4. This follows in routine fashion by decomposing the error equation for the numerical difference scheme into its linear and nonlinear parts and treating the nonlinear part along with the conjugation error from transformation into \( Z \)-coordinates together as source terms in a fixed-point equation in a combination of the discrete linear argument of Corollary 3.13 and the continuous argument of Lemma 2.1 to obtain a contraction in the weighted \( \ell^\infty(Z^+) \) space defined by norm \( \|W\| := \sup_j |W_j e^{\tilde{\theta}x_j}| \), similarly as in the standard proof of the Stable Manifold Theorem.

We omit the straightforward but tedious details, except to mention one subtle point that will recur in later applications. Namely, the righthand side of (1.8)(i), considered
as a function on the complex-valued matrix $\Omega$, is not $C^1$. For, it involves the operation of matrix adjoint, which in turn involves complex conjugation, a non-analytic function on complex arguments. Thus, we cannot immediately apply the above-described argument to (1.8) as written as a complex-valued ODE, but must instead first decompose it into real and imaginary parts, or, as we prefer to do, consider the doubled system

\begin{align*}
\Omega' &= (I - \Omega\tilde{\Omega})A\Omega, \\
\tilde{\Omega}' &= \tilde{\Omega}A^*(I - \Omega\tilde{\Omega})
\end{align*}

in the pair of variables $(\Omega, \tilde{\Omega})$, with $\tilde{\Omega} := \Omega^*$. With this (purely internal) change, the argument goes through as described to yield the claimed result.

5 Boundary-value algorithms

Finally, on a more speculative note, we develop further some ideas of [S, HuZ1] regarding implementation of boundary-valued based Evans solvers for use in extremely large-scale systems, in the light of our new results.

5.1 Sandstede’s method

We first describe (perhaps an imperfect translation of) Sandstede’s original idea based on established projective boundary-value methods [Bel]. This consists, loosely speaking, of numerically solving the original, uncentered Evans system (2.1) on $[0, M]$ with mixed projective boundary conditions

\begin{equation}
\Pi_+ W^m_j = 0, \quad \Pi_0 W^j_0 = \alpha_m, \quad m = 1, \ldots, k,
\end{equation}

where $\Pi_+$ is the rank-$(n - k)$ unstable eigenprojection of $A_+$, $\Pi_0$ is a randomly chosen rank-$k$ projection, and $\alpha_m, m = 1, \ldots, k$ are $k$ random phase conditions determining a basis of $k$ independent solutions $W^m$. Here, as usual, $W^m_j$ denotes the numerical approximation of $W^m(x_j)$. For generic choices of $\Pi_0, \alpha$, this will give a numerically well-conditioned problem, so the procedure is to randomly select candidate values, then change these if the method does not converge after appropriate time.

The solution of the boundary-value scheme for a given parameter value is then obtained by Newton iteration combined with continuation/path-following. Once the method is running, initial guesses for $\Pi_0, \alpha$, and the solution itself may be chosen strategically based on the solution for nearby parameters, to improve conditioning/speed of convergence.

The disadvantages of this method are two: (i) decay at plus infinity means we don’t have control of the asymptotic behavior of solutions at $+\infty$, making it difficult to impose the desirable property of analyticity in $\lambda$; indeed, we prescribe solutions by phase conditions at $x = 0$, where we have no direct knowledge of the link to asymptotic behavior. (ii) (related) Note that applying a standard numerical difference scheme to the doubled system (4.11) yields an algorithm identical to what would be obtained by applying it to the original equation (1.8)(i).
prescription of random phase conditions at the origin is logarithmically complicated, requiring additional error control/programming beyond just solution of the Evans ODE.

Advantages of the method are the existence of a well-developed theory of convergence/error estimation for methods of this form, and a hoped-for dimensional advantage of iterative methods vs. shooting in the treatment of large systems.

5.2 A polar coordinate-based method

Following a suggestion of [HuZ1], we propose an alternative boundary-value scheme based on the polar coordinate method, using a Newton-based iterative boundary-value solver\(^8\) to approximate the solutions \((\Omega^+, \tilde{\Omega}^+, \tilde{\gamma}^+)\) and \((\Omega^-, \tilde{\Omega}^-, \tilde{\gamma}^-)\) of the doubled equations

\[
\Omega' = (I - \Omega \tilde{\Omega}) A_c \Omega, \quad \tilde{\Omega}' = \tilde{\Omega} A_c^* (I - \Omega \tilde{\Omega}), \quad \log \tilde{\gamma}' = \text{Trace}(\tilde{\Omega} A_c \tilde{\Omega}) - \text{Trace}(\tilde{\Omega} A_c \tilde{\Omega})_\pm
\]

on \([0, \pm M]\) and \([-M, 0]\) with Dirichlet boundary conditions

\[
(\Omega, \tilde{\Omega}, \tilde{\gamma})(\pm M) = (\Omega_\pm, \tilde{\Omega}_\pm, \tilde{\gamma}_\pm),
\]

starting with the exact solution \((\Omega, \tilde{\Omega}, \tilde{\gamma}) \equiv (\Omega_\pm, \tilde{\Omega}_\pm, \tilde{\gamma}_\pm)\) at \(c = 0\) and continuing via the homotopy

\[
A_c := A_\pm + c(A - A_\pm), \quad c \in [0, 1]
\]

to the desired solution of the full problem \(A_c = A\) at \(c = 1\).

This approach eliminates disadvantages (i)-(ii) of the standard approach and appears straightforward to code. Moreover, Corollary 4.4 gives a rigorous convergence result, indicating at least theoretical feasibility. What remains to be seen is whether it is practically useful on the scale of interest, and how its performance compares with standard uncentered schemes as described in Section 5.1. We hope to address these questions in future work.

Remark 5.1. As noted already in the proof of Corollary 4.4, the use of doubled coordinates is necessary in order that the ODE be \(C^1\) as a function of its arguments, since the matrix adjoint operation, since not analytic as a complex-valued function, is not \(C^1\). First-order smoothness is needed to apply Newton iteration, of which the first step is linearization.

5.3 A conjugation-based method

A possible drawback of the polar coordinate in numerically sensitive situations is its nonlinearity. An alternative, still more speculative, linear solution would be to use a Newton-based iterative solver to approximate on \([0, M]\) and \([-M, 0]\) solutions \(P^+\) and \(P^-\) of the conjugation equations

\[
P' = AP := A_\pm P - PA
\]

of (2.13), using projective boundary conditions

\[
\Pi_\pm (P_\pm - I)(\pm M) = 0, \quad \Pi_0 P_\pm = \alpha
\]

\(^8\)For example, MATLAB’s BVP5P.
starting with initial guess $P^\pm \equiv I$ and using a similar homotopy

$$A_c := A_\pm + c(A - A_\pm), \quad c \in [0,1]$$

from $A$ to its constant-coefficient limits $A_\pm$, defining an Evans approximant simply as

$$D^{h,M}(\lambda) := \det(P^+R^+,P^-R^-)|_{x=0},$$

where $R^\pm$ are matrices whose columns are bases of the stable (unstable) subspace of $A_\pm$.

Again, we have a rigorous convergence result, this time in the form of Theorem 3.3, but it is not clear whether the scheme is practically useful, or if so how its performance compares to that of the previously mentioned schemes. Moreover, besides sharing difficulty (ii) of the standard method, it has the additional difficulty that the dimension of $\Pi_+$ may change with different $\lambda$, necessitating still further modifications to the boundary conditions.

6 Postscript: initialization of eigenbases at infinity

For completeness, we describe, following [BrZ, HSZ, HuZ1, Z2], a simple and effective method for computing analytically-chosen initializing eigenbases at plus and minus spatial infinity. Combined with the integration methods described in the rest of the paper, this gives a basic working Evans solver that performs quite well in practice.\[10]

6.1 Kato’s ODE

Denote by $\Pi_+$ and $\Pi_-$ the eigenprojections of $A_+$ onto its stable subspace and $A_-$ onto its unstable subspace, with $A_\pm$ defined as in (1.2). Assume as in the introduction that the dimensions of the stable and unstable subspaces are constants $k$ and $n - k$ on the desired region of investigation $\lambda \in \Lambda$ and sum to $n$ (the “consistent splitting condition” of [AGJ]). By standard matrix perturbation theory, $\Pi_\pm$ are analytic in $\lambda$ for $\lambda \in \Lambda$ [K]. Introduce the complex ODE

$$R' = \Pi'R, \quad R(\lambda_*) = R_*,$$

where $'$ denotes $d/d\lambda$, $\lambda_* \in \Lambda$ is fixed, $\Pi = \Pi_\pm$, and $R = R_\pm$ with $R_+$ and $R_-$ $n \times k$ and $n \times (n - k)$ complex matrices, and $R_*$ is full rank and satisfies $\Pi(\lambda_*)R_* = R_*$: that is, its columns are a basis for the stable (resp. unstable) subspace of $A_+$ (resp. $A_-$).

Lemma 6.1 ([K, Z2]). There exists a global analytic solution $R$ of (6.1) on $\Lambda$ such that (i) $\text{rank}R \equiv \text{rank}R^*$, (ii) $\Pi R \equiv R$, and (iii) $\Pi R' \equiv 0$.

Proof. As a linear ODE with analytic coefficients, (6.1) possesses an analytic solution in a neighborhood of $\lambda_*$, that may be extended globally along any curve, whence, by the principle

\[9\] Described further in Section 6

\[10\] Optimized versions may be found in the STABLAB package developed by J. Humpherys.
of analytic continuation, it possesses a global analytic solution on any simply connected
domain containing \( \lambda \). Property (i) follows likewise by the fact that \( R \) satisfies a linear
ODE. Differentiating the identity \( \Pi^2 = \Pi \) following \( K \) yields \( \Pi \Pi' = \Pi' \Pi = \Pi' \), whence,
multiplying on the right by \( \Pi \), we find the key property

\[
(6.2) \quad \Pi \Pi' = 0. 
\]

From (6.2), we obtain

\[
(\Pi R - R)' = (\Pi'R + \Pi R' - R') = \Pi'R + (\Pi - I)\Pi'R = \Pi\Pi'R,
\]

which, by \( \Pi \Pi' \Pi = 0 \) and \( \Pi^2 = \Pi \) gives

\[
(\Pi R - R)' = -\Pi\Pi'(\Pi R - R), \quad (\Pi R - R)(\lambda) = 0,
\]

from which (ii) follows by uniqueness of solutions of linear ODE. Expanding \( \Pi R' = \Pi \Pi'R \)
and using \( \Pi R = R \) and \( \Pi \Pi' \Pi = 0 \), we obtain \( \Pi R' = \Pi \Pi'R \Pi = 0 \), verifying (iii).

Remark 6.2. Property (iii) indicates that the Kato basis is an optimal choice in the sense
that it involves minimal variation in \( R \). It is also useful as a direct characterization of the
Kato basis independent of (6.1); see [HSZ, BDG] or Example 6.6 below.

6.2 Numerical implementation

Choose a set of mesh points \( \lambda_j, j = 0, \ldots, J \) along a path \( \Gamma \subset \Lambda \) and denote by \( \Pi_j := \Pi(\lambda_j) \)
and \( R_j \) the approximation of \( R(\lambda_j) \). Typically, \( \lambda_0 = \lambda_J \), i.e., \( \Gamma \) is a closed contour.

6.2.1 Computing \( \Pi_j \)

Given a matrix \( A \), one may efficiently (\( \sim 32n^3 \) operations; see [GvL, SB]) compute by
“ordered” Schur decomposition\(^{11}\), i.e., Schur decomposition \( A = QUQ^{-1} \), \( Q \) orthogonal
and \( U \) upper triangular, for which also the diagonal entries of \( U \) are ordered in increasing
real part, an orthonormal basis

\[
\tilde{R}_u := (Q_{k+1}, \ldots, Q_n),
\]

of its unstable subspace, where \( Q_{k+1}, \ldots, Q_n \) are the last \( n - k \) columns of \( Q \), \( n - k \) the
dimension of the unstable subspace. Performing the same procedure for \( -A \), \( A^* \), and
\( -A^* \) we obtain orthonormal bases \( \tilde{R}_u, \tilde{L}_u, \tilde{L}_s \) also for the stable subspace of \( A \) and the
unstable and stable subspaces of \( A^* \), from which we may compute the stable and unstable
eigenprojections in straightforward and numerically well-conditioned manner via

\[
(6.3) \quad \Pi_u := \tilde{R}_u (\tilde{L}_u^* \tilde{R}_u)^{-1} \tilde{L}_u^*, \quad \Pi_s := \tilde{R}_s (\tilde{L}_s^* \tilde{R}_s)^{-1} \tilde{L}_s^*.
\]

Applying this to matrices \( A_j^\pm := A_j(\lambda_j) \), we obtain the projectors \( \Pi_j^\pm := \Pi_j(\lambda_j) \). Hereafter,
we consider \( \Pi_j^\pm \) as known quantities.

\(^{11}\)Supported, for example, in MATLAB and LAPACK.
Remark 6.3. It is tempting to instead simply call an eigenvalue–eigenvector solver and express \( \Pi s = \sum r_j l_j \), where \( r_j, l_j \) are left and right eigenvalues associated with stable eigenvectors. However, this becomes ill-conditioned near points where stable eigenvalues collide, as frequently happens for cases resulting from other than simple scalar equations.

6.2.2 First-order integration scheme

Approximating \( \Pi'(\lambda_j) \) to first order by the finite difference \( (\Pi_{j+1} - \Pi_j) / (\lambda_{j+1} - \lambda_j) \) and substituting this into a first-order Euler scheme gives

\[
R_{j+1} = R_j + (\lambda_{j+1} - \lambda_j) \frac{\Pi_{j+1} - \Pi_j}{\lambda_{j+1} - \lambda_j} R_j,
\]

or \( R_{j+1} = R_j + \Pi_{j+1} R_j - \Pi_j R_j \), yielding by the property \( \Pi_j R_j = R_j \) (preserved exactly by the scheme) the simple greedy algorithm

\[
(6.4) \quad R_{j+1} = \Pi_{j+1} R_j.
\]

It is a remarkable fact [Z2] (a consequence of Lemma 6.1) that, up to numerical error, evolution of (6.4) about a closed loop \( \lambda_0 = \lambda_J \) yields the original value \( R_J = R_0 \).

6.2.3 Second-order scheme

To obtain a second-order discretization of (6.1), we approximate \( R_{j+1} - R_j \approx \Delta \lambda_j \Pi'_{j+1/2} R_{j+1/2} / 2 \), good to second order, where \( \Delta \lambda_j := \lambda_{j+1} - \lambda_j \). Noting that \( R_{j+1/2} \approx \Pi_{j+1/2} R_j \) to second order, by (6.4), and approximating \( \Pi_{j+1/2} \approx \frac{1}{2} (\Pi_{j+1} + \Pi_j) \), also good to second order, and \( \Pi'_{j+1/2} \approx (\Pi_{j+1} - \Pi_j) / \Delta \lambda_j \), we obtain, combining and rearranging,

\[
R_{j+1} = R_j + \frac{1}{2} (\Pi_{j+1} - \Pi_j)(\Pi_{j+1} + \Pi_j) R_j.
\]

Stabilizing by following with a projection \( \Pi_{j+1} \), we obtain after some rearrangement the reduced second-order explicit scheme

\[
(6.5) \quad R_{j+1} = \Pi_{j+1} [I + \frac{1}{2} \Pi_j (I - \Pi_{j+1})] R_j.
\]

This is the version we recommend for serious computations. For individual numerical experiments the simpler greedy algorithm (6.4) will often suffice (see discussion, [Z2]).

Remark 6.4. Arbitrarily higher-order schemes may be obtained by Richardson extrapolation starting from scheme (6.4) or (6.5); see [Z2]. In practice, this does not seem useful.

6.3 Initialization of Evans function ODE

Finally, we describe the conversion of analytic bases \( R_\pm(\lambda) \) into initial data for the centered exterior product or polar coordinate method.
6.3.1 Centered exterior product scheme

Denote $R^+ = (R^+_1, \ldots, R^+_k)$ and $R^- = (R^-_1, \ldots, R^-_{n-k})$. Then, the initializing wedge products $\mathcal{R}_{S^\pm}$ of Section 3.5 are given simply by

\begin{equation}
\mathcal{R}_{S^+} := R^+_1 \wedge \cdots \wedge R^+_k \quad \text{and} \quad \mathcal{R}_{S^-} := R^-_1 \wedge \cdots \wedge R^-_{n-k}.
\end{equation}

6.3.2 Polar coordinate scheme

For each $\lambda \in \Lambda$, we may efficiently compute matrices $\Omega_{\pm}(\lambda)$ whose columns form orthonormal bases for $S_{\pm}$, by the same ordered Schur decomposition used in the computation of $\Pi_{\pm}$. This need not even be continuous with respect to $\lambda$. Equating

$$\Omega_{+}\tilde{\alpha}_{+}(\lambda) = R_{+}(\lambda), \quad \Omega_{-}\tilde{\alpha}_{-}(\lambda) = R_{-}(\lambda),$$

for some $\tilde{\alpha}_{\pm}$, we obtain

$$\tilde{\alpha}_{+}(\lambda) = \Omega^*_+ R_{+}(\lambda), \quad \tilde{\alpha}_{-}(\lambda) = \Omega^*_- R_{-}(\lambda),$$

and therefore the exterior product of the columns of $R_{\pm}$ is equal to the exterior product of the columns of $\Omega_{\pm}$ times

\begin{equation}
\tilde{\gamma}_{\pm}(\lambda) := \det(\Omega^*_R)(\lambda)).
\end{equation}

Thus, we may initialize the polar coordinate ODE (1.8) with

\begin{equation}
\Omega = \Omega_{\pm}, \quad \tilde{\gamma} = \det(\Omega^*_R).\end{equation}

**Remark 6.5.** The wedge products represented by polar coordinates $(\tilde{\gamma}, \Omega)_{\pm}(\lambda)$, with $\tilde{\gamma}_{\pm}$ defined as in (6.7), are the same products $\mathcal{R}_{S^\pm}$ defined in (6.6). In particular, they are analytic with respect to $\lambda$, though coordinates $\tilde{\gamma}$ and $\Omega$ in general are not.

6.4 Error control

As the integration of Kato's ODE is carried out on a bounded closed curve, standard error estimates apply and convergence is essentially automatic, and we shall not discuss it. In applications, we are often interested in determining the winding number of $D$ about such a curve. For this purpose, following [Br, BrZ], we introduce a simple a posteriori “Rouche” check limiting the step-size in $\lambda$ by the requirement that the relative change in $D(\lambda)$ be less than a conservative 0.1. (By Rouche’s Theorem, relative error less than one is sufficient to obtain the correct winding number.) In contrast to integration in $x$ of the Evans system, integration in $\lambda$ of the Kato system is a one-time cost, so not a rate-determining factor in the performance of the overall code. However, the computation time is sensitive (proportional) to the number of mesh points in $\lambda$, so this should be held down as much as possible.
6.5 Finer points: two exceptional cases

We conclude by pointing out two commonly occurring cases that can give trouble if not expected, and describe some practical resolutions.

6.5.1 Behavior near the origin

For the linearized operators $L$ arising in the study of stability of traveling-waves of certain systems such as viscous conservation laws or Cahn–Hilliard and nonlinear Schrödinger equations, the point $\lambda = 0$ is embedded in the essential spectrum of $L$. In computing a winding number around some bounded portion of the set $\{\Re\lambda \geq 0\}$ of possible unstable eigenvalues, we must pass through or near this value, at which the spectral gap (see discussion above Assumption 3.4) between stable and unstable subspaces of $A_{\pm}$ goes to zero. In this case, the eigenprojections $\Pi_{\pm}$ lose their characterization as stable (resp. unstable) eigenprojections of $A_{\pm}$, so must be computed in a different way than the ordered Schur decomposition described above. Worse, they may lose analyticity, possessing a branch singularity, at $\lambda = 0$.

To avoid the former problem, we typically just compute near but not at $\lambda = 0$. However, this leads to occasional uncertainty/bad results near the origin and should probably be improved in the analytic case by instead computing $\Pi_{\pm}$ at points within or on the essential spectrum boundary of $L$ by analytic extrapolation from values at points outside. This is an important practical area for further algorithm development. The latter problem, concerning behavior near a branch singularity, is discussed in the next subsection.

6.5.2 Behavior near a branch singularity

For certain problems, especially those involving additional parameters, e.g. multi-d [HLyZ2] or families of one-dimensional waves that pass through characteristic values [BHZ], there may appear for certain parameters branch points in the eigenvalues of $A_{\pm}$ as a function of $\lambda$, at which $\Pi_{\pm}$ therefore blow up [K]. This requires some adjustment in order to restore good behavior.

**Example 6.6.** A model for this situation is the eigenvalue equation for a scalar convected heat equation $\lambda u + \eta u' = u''$ with convection coefficient $\eta$ passing through zero. The coefficient matrix for the associated first-order system is

\[
A := \begin{pmatrix} 0 & 1 \\ \lambda & \eta \end{pmatrix}.
\]

Then, the stable eigenvector of $A$ determined by Kato’s ODE (6.1) is

\[
R(\eta, \lambda) := \frac{(\eta^2/4 + 1)^{1/4}}{(\eta^2/4 + \lambda)^{1/4}} \begin{pmatrix} 1, -\eta/2 - \sqrt{\eta^2/4 + \lambda} \end{pmatrix}^T,
\]

which, apart from the divergent factor $\frac{(\eta^2/4+1)^{1/4}}{(\eta^2/4+\lambda)^{1/4}}$, is a smooth function of $\sqrt{\eta^2/4 + \lambda}$.

\[\text{\textsuperscript{12}}\text{This would also allow computations within the essential spectrum, up to now not systematically carried out (though see [Br] for some preliminary results in this direction).}\]
Proof. By straightforward computation, \( \mu_{\pm}(\lambda) := \mp(\eta/2+\sqrt{\eta^2/4+\lambda}) \) and \( V_{\pm} := (1, \mu_{\pm}(\lambda))^T \) are eigenvalues and eigenvectors of the matrix \( A \) of (6.9) in Example 6.6. The associated Kato eigenvectors \( V_{\pm} \) are determined uniquely, up to a constant factor independent of \( \lambda \), by the property that there exist corresponding left eigenvectors \( \tilde{V}_{\pm} \) such that

\[
(\tilde{V} \cdot V)^\pm \equiv \text{constant}, \quad (\tilde{V} \cdot \dot{V})^\pm \equiv 0,
\]

where “\( \cdot \)” denotes \( d/d\lambda \); see Lemma 6.1(iii).

Computing dual eigenvectors \( \tilde{V}_{\pm} = (\lambda + \mu^2)^{-1}(\lambda, \mu_{\pm}) \) satisfying \( (\tilde{V} \cdot V)^\pm \equiv 1 \), and setting \( V_{\pm} = c_{\pm} V_{\pm}^\pm, \tilde{V}_{\pm} = V_{\pm} / c_{\pm} \), we find after a brief calculation that (6.11) is equivalent to the complex ODE

\[
\dot{c}_{\pm} = -\left(\tilde{V} \cdot \dot{V}\right)^\pm c_{\pm} = -\left(\frac{\dot{\mu}}{2\mu - \eta}\right) c_{\pm},
\]

which may be solved by exponentiation, yielding the general solution \( c_{\pm}(\lambda) = C (\eta^2/4 + \lambda)^{-1/4} \). Initializing without loss of generality at \( c_{\pm}(1) = 1 \), we obtain (6.10).

Remark 6.7. It is straightforward to generalize by the same method the computation of Example 6.6 to branch singularities of general order \( s \).

The computation of Example 6.6 indicates that the Kato basis blows up at \( \lambda = 0 \) as \( (\eta^2 + 4\lambda)^{-1/4} \) as \( \eta \) crosses the characteristic value \( \eta = 0 \), hence does not give a choice that is continuous across the entire range of parameters. However, the same example shows that there is a different choice \( (1, -\eta/2 - \sqrt{\eta^2/4+\lambda})^T \) that is continuous, possessing only a square-root singularity. We can effectively exchange one for another, by rescaling the Kato basis by factor \( (\eta^2 + 4\lambda)^{1/4} \). See [BHZ] for examples/further details.

This issue is mainly important in problems with parameters, or possessing branch singularities, but can also arise for a problem without parameter or singularity that happens to lie near a related problem with branch singularities. In such a case the “invisible” branch singularity could serve as an organizing center directing the Kato flow without the user being aware of it. Thus, it is important to be alert to this possibility.

References

[ACR] U.M. Ascher, H. Chin, and S. Reich, Stabilization of DAEs and invariant manifolds, Numer. Math. 67 (1994) 131–149.

[AGJ] J. Alexander, R. Gardner, and C.K.R.T. Jones, A topological invariant arising in the analysis of traveling waves. J. Reine Angew. Math. 410 (1990) 167–212.

[AS] J. C. Alexander and R. Sachs. Linear instability of solitary waves of a Boussinesq-type equation: a computer assisted computation, Nonlinear World 2 (1995) 471–507.

[AlB] L. Allen and T. J. Bridges, Numerical exterior algebra and the compound matrix method, Numer. Math. 92 (2002) 197–232.
B. Barker, J. Humpherys, and K. Zumbrun, *One-dimensional stability of parallel shock layers in isentropic magnetohydrodynamics*, in preparation.

B. Barker, J. Humpherys, K. Rudd, and K. Zumbrun, *Stability of viscous shocks in isentropic gas dynamics*, Comm. Math. Phys. 281 (2008), no. 1, 231–249.

W.-J. Beyn, *The numerical computation of connecting orbits in dynamical systems*, IMA J. Numer. Analysis 9 (1990) 379–405.

W.-J. Beyn, *Zur Stabilität von Differenzenverfahren für Systeme linearer gewöhnlicher Randwerttaufgaben*, Numer. Math. 29 (1978) 209–226.

T.J. Bridges, *The Orr-Sommerfeld equation on a manifold*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci., 455 (1999) 3019–3040.

T.J. Bridges, G. Derks, and G. Gottwald, *Stability and instability of solitary waves of the fifth-order KdV equation: a numerical framework*, Phys. D, 172(1-4):190–216, 2002.

T. J. Bridges and S. Reich, *Computing Lyapunov exponents on a Stiefel manifold*, Phys. D 156 (2001) 219–238.

L.Q. Brin, *Numerical testing of the stability of viscous shock waves*, Math. Comp., 70(235):1071–1088, 2001.

L.Q. Brin and K. Zumbrun, *Analytically varying eigenvectors and the stability of viscous shock waves*, Mat. Contemp., 22:19–32, 2002, Seventh Workshop on Partial Differential Equations, Part I (Rio de Janeiro, 2001).

W.A. Coppel, *Stability and asymptotic behavior of differential equations*, D.C. Heath and Co., Boston, MA (1965) viii+166 pp.

N. Costanzino, J. Humpherys, T. Nguyen, and K. Zumbrun, *Spectral stability of noncharacteristic boundary layers of isentropic Navier–Stokes equations*, to appear, Arch. for Rat. Mech. Anal.

A. Davey, *An automatic orthonormalization method for solving stiff boundary value problems*, J. Comput. Phys. 51 (1983) 343–356.

J. W. Demmel, L. Dieci and M. J. Friedman. *Computing connecting orbits via an improved algorithm for continuing invariant subspaces*. SIAM J. Sci. Comput. 22 (2000) 81–94.

G. Dahlquist, *Convergence and stability in the numerical integration of ordinary differential equations*, Math. Scand., V. 4 (1956) 33–53.

G. Dahlquist, *Stability and error bounds in the numerical integration of ordinary differential equations*, Trans. of the Royal Inst. Of Tchn., Stockholm, Sweden, Nr. 130 (1959) 87 pp.

L. Dieci and T. Eirola. *Applications of smooth orthogonal factorizations of matrices*. In Numerical methods for bifurcation problems and large-scale dynamical systems (Minneapolis, MN, 1997). Springer, IMA Vol. Math. Appl. 119 (2000) 141–162.

L. Dieci and T. Eirola. *On smooth decompositions of matrices*. SIAM J. Matrix Anal. Appl. 20 (1999) 800–819.

L. Dieci and M. J. Friedman. *Continuation of invariant subspaces*. Numer. Lin. Alg. Appl. 8 (2001) 317–327.
[Dr] L. O. Drury, *Numerical solution of Orr-Sommerfeld-type equations*, J. Comput. Phys. 37 (1980) 133–139.

[Er] J.J. Erpenbeck. *Stability of steady-state equilibrium detonations*, Physics of Fluids 5 (1962) 604–614.

[GJ1] R. Gardner and C.K.R.T. Jones, *A stability index for steady state solutions of boundary value problems for parabolic systems*, J. Diff. Eqs. 91 (1991), no. 2, 181–203.

[GJ2] R. Gardner and C.K.R.T. Jones, *Traveling waves of a perturbed diffusion equation arising in a phase field model*, Ind. Univ. Math. J. 38 (1989), no. 4, 1197–1222.

[GZ] R. Gardner and K. Zumbrun, *The gap lemma and geometric criteria instability of viscous shock profiles*, CPAM 51. 1998, 797-855.

[GLZ] F. Gesztesy, Y. Latushkin, and K. Zumbrun, *Derivatives of (Modified) Fredholm Determinants and Stability of Standing and Traveling Waves*, J. Math. Pures Appl. (9) 90 (2008), no. 2, 160–200.

[GB] F. Gilbert and G. Backus, *Propagator matrices in elastic wave and vibration problems*, Geophysics, 31 (1966) 326–332.

[GvL] G. H. Golub and C. F. Van Loan, *Matrix computations*, Johns Hopkins University Press, Baltimore (1996).

[HNW1] Hairer, Nosrett, and Wanner, *Solving ordinary differential equations. I. Nonstiff problems*, Second edition. Springer Series in Computational Mathematics, 8. Springer-Verlag, Berlin, 1993. xvi+528 pp. ISBN: 3-540-56670-8 65-02.

[HNW2] Hairer, Nosrett, and Wanner, *Solving ordinary differential equations. II. Stiff and differential-algebraic problems*, Second edition. Springer Series in Computational Mathematics, 14. Springer-Verlag, Berlin, 1996. xvi+614 pp. ISBN: 3-540-60452-9 65-02.

[He] D. Henry, *Geometric theory of semilinear parabolic equations*, Lecture Notes in Mathematics, Springer–Verlag, Berlin (1981), iv + 348 pp.

[HLZ] J. Humpherys, O. Lafitte, and K. Zumbrun, *Stability of isentropic viscous shock profiles in the high-mach number limit*, Preprint (2007).

[HLyZ1] J. Humpherys, G. Lyng, and K. Zumbrun, *Spectral stability of ideal-gas shock layers*, to appear, Archive for Rat. Mech. Anal.

[HLyZ2] J. Humpherys, G. Lyng, and K. Zumbrun, *Multidimensional spectral stability of large-amplitude Navier-Stokes shocks*, in preparation.

[HSZ] J. Humpherys, B. Sandstede, and K. Zumbrun, *Efficient computation of analytic bases in Evans function analysis of large systems*, Numer. Math. 103 (2006), no. 4, 631–642.

[HuZ1] J. Humpherys and K. Zumbrun, *An efficient shooting algorithm for evans function calculations in large systems*, Physica D, 220(2):116–126, 2006.

[HuZ2] J. Humpherys and K. Zumbrun, *Efficient numerical stability analysis of detonation waves in ZND*, in preparation.

[J] C.K.R.T. Jones, Discussion session, AIM workshop on Stability Criteria for Multi-Dimensional Waves and Patterns, May 2005.
[KS] T. Kapitula and B. Sandstede, *Stability of bright solitary-wave solutions to perturbed nonlinear Schrödinger equations*, Phys. D, 124 (1998) 58–103.

[K] T. Kato, *Perturbation theory for linear operators*. Springer–Verlag, Berlin Heidelberg (1985).

[Kr] H.O. Kreiss, *Difference approximations for boundary and eigenvalue problems for ordinary differential equations*, Math. Comp. 26 (1972) 605–624.

[LS] H.I. Lee and D.S. Stewart, *Calculation of linear detonation instability: one-dimensional instability of plane detonation*, J. Fluid Mech. 216 (1990) 103–132.

[LPSS] G. J. Lord, D. Peterhof, B. Sandstede, and A. Scheel, *Numerical computation of solitary waves in infinite cylindrical domains*, SIAM J. Numer. Anal. 37 (2000) 1420–1454.

[MeZ] G. Métivier and K. Zumbrun, *Large viscous boundary layers for noncharacteristic nonlinear hyperbolic problems*, Mem. Amer. Math. Soc. 175 (2005), no. 826, vi+107 pp.

[NR1] B. S. Ng and W. H. Reid, *An initial value method for eigenvalue problems using compound matrices*, J. Comput. Phys. 30 (1979) 125–136.

[NR2] B. S. Ng and W. H. Reid, *A numerical method for linear two-point boundary value problems using compound matrices*, J. Comput. Phys. 33 (1979) 70–85.

[NR3] B. S. Ng and W. H. Reid, *On the numerical solution of the Orr-Sommerfeld problem: asymptotic initial conditions for shooting methods*, J. Comput. Phys., 38 (1980) 275–293.

[NR4] B. S. Ng and W. H. Reid, *The compound matrix method for ordinary differential systems*, J. Comput. Phys. 58 (1985) 209–228.

[OZ] M. Oh and K. Zumbrun, *Stability of periodic solutions of viscous conservation laws with viscosity- 1. Analysis of the Evans function*, Arch. Ration. Mech. Anal. 166 (2003), no. 2, 99–166.

[PW] R. L. Pego and M. I. Weinstein. *Eigenvalues, and instabilities of solitary waves*, Philos. Trans. Roy. Soc. London Ser. A 340 (1992) 47–94.

[S] B. Sandstede, *Private Communication*, Little Compton meeting on Evans function techniques (1998).

[SB] J. Stoer and R. Bulirsch, *Introduction to numerical analysis*, Springer-Verlag, New York (2002).

[Z1] K. Zumbrun, *Stability of large-amplitude shock waves of compressible Navier-Stokes equations*, With an appendix by Helge Kristian Jenssen and Gregory Lyng. Handbook of mathematical fluid dynamics. Vol. III, 311–533, North-Holland, Amsterdam, (2004).

[Z2] K. Zumbrun, *A local greedy algorithm and higher-order extensions for global continuation of analytically varying subspaces*, To appear, Quarterly J. Appl. Math.

[ZH] K. Zumbrun and P. Howard, *Pointwise semigroup methods and stability of viscous shock waves*. Indiana Mathematics Journal V47 (1998), 741–871.