

GMRES for oscillatory matrix-valued differential equations

Sheehan Olver

Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford, UK

`Sheehan.Olver@sjc.ox.ac.uk`

Abstract We investigate the use of Krylov subspace methods to solve linear, oscillatory ODEs. When we apply a Krylov subspace method to a properly formulated equation, we retain the asymptotic accuracy of the asymptotic expansion whilst converging to the exact solution. We will demonstrate the effectiveness of this method by computing Error and Mathieu functions.

Oxford University Computing Laboratory
Numerical Analysis Group
Wolfson Building
Parks Road
Oxford, England OX1 3QD

September, 2009

1. Introduction

Our aim is to compute the fundamental solution to the differential equation

$$Y'(t) = (B(t) + \omega A(t))Y(t), \quad t \in (a, b) \tag{1.1}$$

where A and B are $d \times d$ matrix-valued functions and ω is large. Applications for such equations include the computation of special functions (such as Airy, Bessel, Hypergeometric and Matthieu functions [1]) the time-independent Schrödinger equation and semi-discretizations of the linear time-dependent Schrödinger equation.

When the eigenvalues of A are imaginary, the solutions to (1.1) become more oscillatory as $\omega \rightarrow \infty$. Thus traditional time-stepping methods are inefficient for large ω . The accuracy of modified Magnus expansions [5] does not degenerate as ω increases when used to compute (1.1). On the other hand, the approach we construct actually improves with accuracy as ω increases, and at an arbitrarily high asymptotic order. Moreover, we solve the equation globally, allowing us to compute over unbounded domains and to higher accuracy than a time stepping approach can achieve.

The simplest form of (1.1) is when $d = 1$, in which case we obtain the solution exactly:

$$Y = \exp \int (B + \omega A) dt.$$

The nonhomogenous form of (1.1) is the Levin differential equation [6] (changing notation to emphasize that these are scalar functions)

$$\mathcal{L}y = y' + i\omega g'y = f. \tag{1.2}$$

A particular solution is

$$y = e^{-i\omega g} \int f e^{i\omega g} dt.$$

In other words, solving (1.2) allows us to compute oscillatory integrals:

$$\int_a^b f e^{i\omega g} dt = y(b)e^{i\omega g(b)} - y(a)e^{i\omega g(a)}.$$

Many methods have been developed in recent years for computing oscillatory integrals, with a recent review in [4]. One particular approach is to apply the GMRES method [10] directly to the differential equation [7, 8, 9], which we refer to as differential GMRES. By reformulating (1.2) as a shifted linear system

$$\mathcal{M}u = (\mathcal{M}_0 + i\omega)u = f, \tag{1.3}$$

where \mathcal{M}_0 represents a linear operator, we achieve an asymptotic error in residual of

$$\mathcal{O}(\omega^{-n-1}),$$

where n is the number of GMRES iterations. This is the same asymptotic order as an asymptotic expansion, however, differential GMRES actually converges for fixed ω , subject to a condition on the growth of f in the complex plane. In Section 2, we go into the details of this approach.

The goal, then, is to generalize this approach to the higher dimensional case of (1.1). This is accomplished by reforming the equation as a (matrix-valued) shifted linear system (1.3). By doing so, we obtain a method which also simultaneously achieves high asymptotic order and (based on numerical results) convergence.

Differential GMRES in its pure form requires taking derivatives (and integrals in the higher dimensional case) of the functions involved. In the general case, this is impractical. To avoid this, while we develop the framework in the infinite-dimensional setting, in practice we represent non-oscillatory functions as Chebyshev polynomials. This could be handled automatically and adaptive by the `chebfun` system [2], however, for concreteness and speed we use fixed order Chebyshev polynomials in our examples. More precisely, we represent functions by their values at Chebyshev–Lobatto points, and the FFT can be used to compute derivatives and anti-derivatives.

2. Oscillatory integrals

GMRES [10] is an iterative algorithm originally developed for solving finite-dimensional linear systems

$$A\mathbf{v} = \mathbf{b} \quad \text{for} \quad A \in \mathbb{C}^{d \times d} \quad \text{and} \quad \mathbf{b} \in \mathbb{C}^d.$$

The Krylov subspace is defined as

$$\mathcal{K}_n[A, \mathbf{b}] = \text{span} \{ \mathbf{b}, \dots, A^{n-1}\mathbf{b} \}.$$

GMRES finds an element $\mathbf{v} \in \mathcal{K}_n[A, \mathbf{b}]$ such that the norm

$$\|A\mathbf{v} - \mathbf{b}\|_2$$

is minimized. This is accomplished through Arnoldi iteration [11], which finds an orthonormal matrix $Q = (q_1, \dots, q_n)$ whose columns span $\mathcal{K}_n[A, \mathbf{b}]$ and an upper Hessenberg (only zeros below the first subdiagonal) matrix $H \in \mathbb{C}^{(n+1) \times n}$ such that

$$AQ = QH \quad \text{and} \quad q_1 = \frac{\mathbf{b}}{\|\mathbf{b}\|_2}.$$

Then $\mathbf{c} \in \mathbb{C}^n$ is chosen to minimize the norm

$$\|H\mathbf{c} - \|\mathbf{b}\| \mathbf{e}_1\|_2.$$

The GMRES approximation is now $\mathbf{v} = Q\mathbf{c}$.

In [7, 8], this was generalized for unbounded, infinite-dimensional operators such as the differentiation operator \mathcal{D} . In exactly the same manner as the finite-dimensional case, given the linear operator \mathcal{L} , function f and a semi-inner product $\langle \cdot, \cdot \rangle$, Arnoldi iteration determines a row-vector $\mathbf{q} = (q_1, \dots, q_n)$ whose columns are functions which span the Krylov subspace $\mathcal{K}_n[\mathcal{L}, f]$ such that (where we use the convention $\mathcal{L}\mathbf{q} = (\mathcal{L}q_1, \dots, \mathcal{L}q_n)$)

$$\mathcal{L}\mathbf{q} = \mathbf{q}H \quad \text{and} \quad q_1 = \frac{f}{\|f\|}.$$

Similarly, differential GMRES finds a function $v \in \mathcal{K}_n[\mathcal{L}, f]$ that minimizes the seminorm

$$\|\mathcal{L}v - f\|.$$

This is accomplished by finding $\mathbf{c} \in \mathbb{C}^n$ that minimizes the finite-dimensional norm

$$\|H\mathbf{c} - \|\mathbf{b}\| \mathbf{e}_1\|_2,$$

thence $v = \mathbf{q}\mathbf{c}$.

Now consider the case of Arnoldi iteration applied to a shifted linear system of the form (1.3). We will denote the Hessenberg matrix produced by Arnoldi iteration for a particular value of ω as H_ω . A fact known from the finite-dimensional case which is also true in the infinite-dimensional case is that the orthonormal basis Q is independent of ω and

$$H_\omega = H_0 + i\omega I_{n,n+1}. \quad [3]$$

In other words, we only need to compute the Arnoldi iteration for one choice of ω to determine the GMRES approximation for all choices of ω . Furthermore, GMRES satisfies the property that the error in residual is

$$\|\mathcal{M}v - f\| = \mathcal{O}(\omega^{-n}). \quad [8]$$

In our particular case we wish to solve the Levin differential equation (1.2). But \mathcal{L} is not in the form of a shifted linear operator. If we assume that g' does not vanish, we can trivially put it into the required form:

$$\mathcal{M} = \mathcal{L} \frac{1}{g'} = \mathcal{D} \frac{1}{g'} + i\omega.$$

We thus apply differential GMRES to \mathcal{M} , f and a suitable inner product to obtain v . Then $y = \frac{v}{g'}$ and hence we approximate the oscillatory integral by

$$\frac{v(b)}{g'(b)} e^{i\omega g(a)} - \frac{v(a)}{g'(a)} e^{i\omega g(a)}.$$

since we represent functions by their values, we use the standard dot product on the sample vector at Chebyshev–Lobatto points.

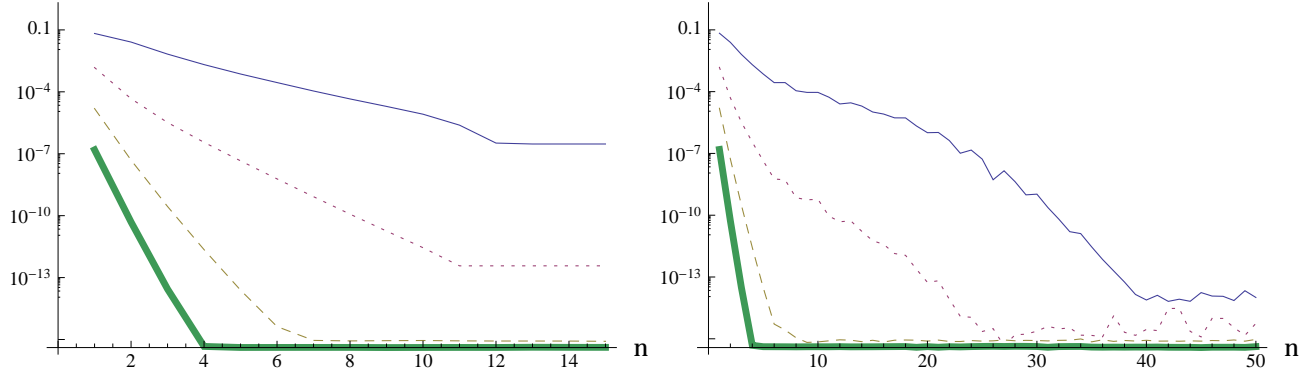


Figure 1: The error in computing $\frac{\sqrt{\pi}\operatorname{erfc}\sqrt{-i\omega}}{2\sqrt{-i\omega}}$ using differential GMRES with 15 (left figure) and 50 (right figure) mapped Chebyshev–Lobatto points, for $\omega = 1$ (plain), 10 (dotted), 100 (dashed) and 1000 (thick).

If the integral does contain a stationary point, i.e., g' vanishes in $[a, b]$, modifying the operator to take the form of a shifted linear operator is more complicated, and detailed in [9].

Consider the integral

$$\int_1^\infty e^{i\omega t^2} dt = \frac{\sqrt{\pi}\operatorname{erfc}\sqrt{-i\omega}}{2\sqrt{-i\omega}} \quad \text{for } \omega > 0.$$

Since the interval is unbounded, we represent functions by rational Chebyshev series — i.e., in terms of the basis $T_k(\frac{t-2}{t})$ — or, more precisely, by the values they take at the mapped Chebyshev–Lobatto points. In Figure 1, we compute the absolute error of our approximation for several choices of ω . As can be seen, the rate of convergence as $n \rightarrow \infty$ increases with the frequency, and the number of mapped Chebyshev–Lobatto points required to achieve machine precision accuracy decreases with the frequency.

3. Oscillatory differential equations

We will use the notation \exp to denote the matrix exponential, though we only apply it to diagonal matrices, where

$$\exp \operatorname{diag}(d_1, \dots, d_n) = \operatorname{diag}(\exp d_1, \dots, \exp d_n).$$

We will use the indefinite integral notation to denote

$$\int A(t) dt = \int_a^t A(t) dt.$$

We now consider the solution of (1.1). Motivated by the preceding section, our goal is to transform this equation into a nonhomogeneous shifted linear system. We will only consider

the case where A is diagonalizable with distinct eigenvalues. Thus assume that there exists a matrix-valued function V that is nonsingular for all $t \in [a, b]$ and diagonal matrix-valued function $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_d)$ where $\lambda_i \neq \lambda_j$ for any i and j , so that

$$AV = V\Lambda. \quad (3.1)$$

The requirement that V is smooth and nonsingular prevents application of this expansion to coalescing eigenvalues. This is similar to the case of stationary points for oscillatory integrals, and likewise outside the scope of this discussion. Note also that V is not determined uniquely, however, our approach works for any choice of V .

We apply the transformation $Y = VW$ to obtain

$$\begin{aligned} (VW)' &= (B + \omega A)W \Leftrightarrow \\ V'W + VW' &= (BV + \omega V\Lambda)W \Leftrightarrow \\ W' &= (H + \omega\Lambda)Y \quad \text{for} \quad H = V^{-1}BV - V^{-1}V'. \end{aligned}$$

We now apply the transformation

$$W = (I + U)e^{\int (\text{diag } H + \omega\Lambda) dt}.$$

Therefore

$$U' + (I + U)(\text{diag } H + \omega\Lambda) = (H + \omega\Lambda)(I + U).$$

We can rephrase this as

$$\mathcal{L}U = F \quad \text{for} \quad \mathcal{L}U = U' + U\text{diag } H - HU + \omega[U, \Lambda],$$

where $[U, \Lambda]$ is the standard matrix commutator $[U, \Lambda] = U\Lambda - \Lambda U$ and $F = H - \text{diag } H$.

Our goal now is to premultiply this operator by an inverse to the commutator operator. Because of our choice of transformations, we have ensured that F has zeros along the diagonal. Thus we can utilize the following inverse to the commutator:

Definition 3.1 For $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ with distinct entries and M with zeros along the diagonal,

$$\mathcal{Q}M = \begin{pmatrix} 0 & \frac{m_{12}}{\lambda_2 - \lambda_1} & \cdots & \frac{m_{1n}}{\lambda_n - \lambda_1} \\ \frac{m_{21}}{\lambda_1 - \lambda_2} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{m_{(n-1)n}}{\lambda_n - \lambda_{n-1}} \\ \frac{m_{n1}}{\lambda_1 - \lambda_n} & \cdots & \frac{m_{n(n-1)}}{\lambda_{n-1} - \lambda_n} & 0 \end{pmatrix}.$$

From inspection, it is clear then that $[\mathcal{Q}F, \Lambda] = \mathcal{Q}[F]\Lambda - \Lambda\mathcal{Q}[F] = F$. However, the term

$$\mathcal{L}\mathcal{Q}F = (\mathcal{Q}F)' + (\mathcal{Q}F)\text{diag } H - H\mathcal{Q}F + \omega F$$

will not necessarily have zeros along the diagonal, Because of the term $H\mathcal{Q}F$. In other words, we cannot generate the Krylov subspace for $\mathcal{L}\mathcal{Q}$ and F . Fortunately, diagonal matrices lie in the kernel of the commutator. Hence we use the following, alternative commutator inverse:

$$[\cdot, \Lambda]^{-1}U = \mathcal{Q}U + \int \text{diag}(H\mathcal{Q}U) dt.$$

Then (using $D = \int \text{diag}(H\mathcal{Q}F) dt$)

$$\begin{aligned} \mathcal{M}F &= \mathcal{L}[\cdot, \Lambda]^{-1}F = (\mathcal{Q}F)' + \text{diag}(H\mathcal{Q}F) - H\mathcal{Q}F \\ &\quad + (\mathcal{Q}F)\text{diag}H + D\text{diag}H - HD + \omega F. \end{aligned}$$

Since $\mathcal{Q}F$ has zeros along the diagonal, the first and fourth terms also have zeros along the diagonal. The second term cancels the diagonal of the third term. Finally, since D is diagonal

$$D\text{diag}H - HD = (\text{diag}H - H)D$$

also has zeros on the diagonal. Thus \mathcal{M} successfully maps the set of infinitely differentiable matrix-valued functions with zeros along the diagonal to itself.

Without modification, we can now construct a differential GMRES method for \mathcal{M} and F , provided an appropriate semi-inner product is used. We will use the Frobenius inner product, where the dot product of two functions remains the dot product of their values at Chebyshev–Lobatto points. This returns a function G which satisfies

$$F \approx \mathcal{M}G = \mathcal{L}[\cdot, \Lambda]^{-1}G.$$

Therefore, $Y' \approx (B + \omega A)Y$ for

$$Y = V(I + [\cdot, \Lambda]^{-1}G)e^{\int (\text{diag}H + \omega\Lambda) dt}.$$

The fundamental solution is then

$$Y_{\mathbb{F}}(t) = Y(t)Y(0)^{-1}.$$

4. Example: Mathieu functions

Consider the system

$$\begin{aligned} Y' &= (B_{\alpha} + \omega A)Y, \\ B_{\alpha} &= \begin{pmatrix} 0 & 1 \\ \alpha & 0 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 0 & 1 \\ -\cosh 2t & 0 \end{pmatrix}. \end{aligned}$$

The eigenvalues of A are $\pm\sqrt{\cosh 2t}$, hence the solution is highly oscillatory. Indeed, as $t \rightarrow \infty$, it becomes exponentially more oscillatory. The fundamental solution to this equation

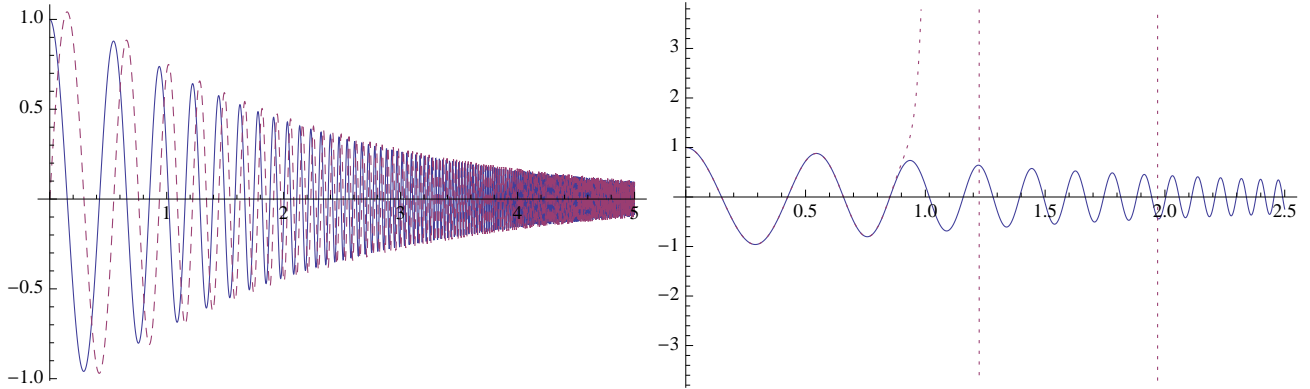


Figure 2: In the left graph, the first row of the differential GMRES approximation to the fundamental solution of the Mathieu equation with $\alpha = 0$ and $\omega = 10$. In the right graph, a comparison of the $(1, 1)$ -entry of the GMRES approximation (solid) to the real part of MATHEMATICA's built-in routine (dotted).

can be written in terms of Mathieu functions [1] as

$$\begin{pmatrix} C\left(\alpha(1+\omega), \frac{1}{2}\omega(1+\omega), it\right) & -S\left(\alpha(1+\omega), \frac{1}{2}\omega(1+\omega), it\right) \\ i\frac{C'(\alpha(1+\omega), \frac{1}{2}\omega(1+\omega), it)}{1+\omega} & -i\frac{S'(\alpha(1+\omega), \frac{1}{2}\omega(1+\omega), it)}{1+\omega} \end{pmatrix}.$$

We apply our approach to approximate this function with $\alpha = 0$ and $\omega = 10$ over the interval $(0, 5)$. As can be seen in Figure 2, our solution is equal to the built-in MATHEMATICA routine (after scaling to obtain the fundamental solution) for computing Mathieu functions when t is small. As t increases, the MATHEMATICA routine quickly explodes, whereas our method remains nicely behaved. Furthermore, the true solution must be real, as is our approximate solution, whilst the MATHEMATICA routine grows a nonzero imaginary component. Comparison with a numerical ODE solver with a very small step size reveals that our solution is indeed the correct one. We omit a graph of the convergence rate for different values of ω , which is similar in behaviour to Figure 1: the larger ω is, the faster the rate of convergence.

References

- [1] Abramowitz, M. and Stegun, I., *Handbook of Mathematical Functions*, National Bureau of Standards Appl. Math. Series, #55, U.S. Govt. Printing Office, Washington, D.C., 1970.
- [2] Battles, Z. and Trefethen, L.N., An extension of MATLAB to continuous functions and operators, *SIAM J. Sci. Comput.* **25** (2004), 1743–1770.

- [3] Datta, B.N. and Saad, Y., Arnoldi methods for large Sylvester-like observer matrix equations and an associated algorithm for partial pole assignment, *Linear Algebra Appl* **154–156** (1991), 225–244.
- [4] Huybrechs, D. and Olver, S. , Highly oscillatory quadrature , *in: Highly Oscillatory Problems: Computation, Theory and Applications*, ed. B. Engquist et al, Cambridge University Press, Cambridge, 2009.
- [5] Iserles, A. , On the method of Neumann series for highly oscillatory equations , *Bit Numerical Mathematics* **44** (2004), 473–488.
- [6] Levin, D., Procedures for computing one and two-dimensional integrals of functions with rapid irregular oscillations, *Math. Comp.* **38** (1982), 531–538.
- [7] Olver, S., GMRES for the differentiation operator, *SIAM J. Numer. Anal.*, to appear.
- [8] Olver, S., Shifted GMRES for oscillatory integrals, *Numer. Math.*, to appear.
- [9] Olver, S. , Fast, numerically stable computation of oscillatory integrals with stationary points , preprint, Technical Report NA-09/2, Computing Laboratory, Oxford University, 2009.
- [10] Saad, Y. and Schultz, M., GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems , *SIAM J. Sci. Stat. Comput.* **7** (1986), 856–869.
- [11] Trefethen, L. and Bau III, D., *Numerical Linear Algebra* , SIAM,1997.