Centre de Recherches Mathématiques
CRM Proceedings and Lecture Notes

On the numerical analysis of rapid oscillation

Arieh Iserles

Abstract. This paper surveys recent advances in the allied challenges of discretizing highly oscillatory ordinary differential equations and computing numerical quadrature of highly oscillatory integrals, and attempts to sketch the mathematical foundations of a general approach to these issues. Having described the Magnus, Cayley and Neumann methods for highly oscillatory ordinary differential equations and analysed the Filon quadrature of highly oscillatory integrals, its overall conclusion is that, once numerical methods are properly crafted, their accuracy increases in the presence of high oscillation.

1. Introduction

It is an article of faith in computational mathematics and scientific computing that highly oscillatory problems, ubiquitous in applications, are among the most intractable phenomena in numerical modelling. The contention of this paper is that this point of view is misleading, indeed woefully wrong. It is not just that properly crafted methods can discretize rapidly oscillating differential and integral equations perfectly well: the truth of the matter is that *such methods are likely to perform significantly better the higher the oscillation.* High oscillation is an ally of a numerical analyst!

In order to understand how to design efficient methods in the presence of high oscillation, we need first to understand why standard numerical methods perform so poorly in this setting. The reason is indeed quite elementary: the (local) error of a standard p-order time-stepping method can be expressed in the form $h^{p+1}\mathrm{D}(\boldsymbol{y},t) + O(h^{p+2})$, where D is either a $(p+1)$st derivative of the highly oscillating solution $\boldsymbol{y}$ or a $(p+1)$st-order elementary differential, composed of derivatives. However, each differentiation of a highly oscillatory function roughly multiplies locally the amplitude by the frequency. Therefore, the higher the oscillation and larger the frequency, the worse the local error. Moreover, the standard remedy to large error, increasing the order of the method, is a non-starter since, the larger the order, the greater the damage wrought by high oscillation.

The *Airy equation*

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + ty = 0, \quad t \geq 0, \qquad y(0) = 1,\ \frac{\mathrm{d}y(0)}{\mathrm{d}t} = 0, \tag{1.1}$$

1991 *Mathematics Subject Classification.* Primary 65L05, 41A55; Secondary 65D30.
Key words and phrases. High oscillation, Quadrature, Ordinary differential equations.

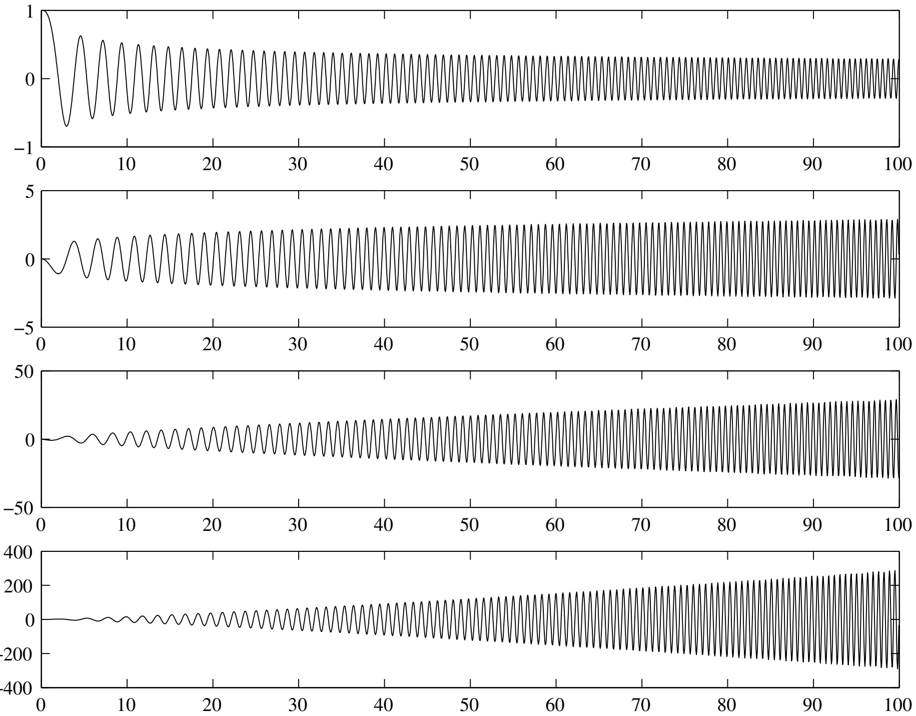

FIGURE 1. The solution of the Airy equation (1.1) and its first, second and third derivatives.

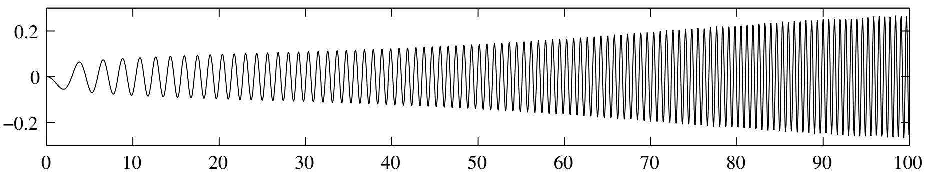

FIGURE 2. The error in solving (1.1) with a three-stage explicit Runge–Kutta method and time-step $h = \frac{1}{20}$.

is an excellent point of departure for our exposition. Figure 1 displays the exact solution y and its first three derivatives: it can be easily proved by the WKB method, that

$$y(t) \sim t^{-1/4} \sin(\tfrac{2}{3}t^{3/2}), \quad \frac{\mathrm{d}y(t)}{\mathrm{d}t} \sim t^{1/4} \cos(\tfrac{2}{3}t^{3/2}), \quad \frac{\mathrm{d}^2 y(t)}{\mathrm{d}t^2} \sim -t^{3/4} \sin(\tfrac{2}{3}t^{3/2}), \quad t \gg 1$$

and so on [**Ise02a**]. It is evident how fast the amplitude grows, once the function is differentiated. The consequences are for all to see in Figure 2, displaying the absolute error in the solution of (1.1) (converted into a first-order system) by the explicit third-order Runge–Kutta method with the tableau

$$\begin{array}{c|ccc} 0 & & & \\ \frac{1}{2} & \frac{1}{2} & & \\ 1 & -1 & 2 & \\ \hline & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{array}$$

[**HNW93**]. The principal error term consists of elementary differentials, among them the fourth derivative of y, and it is the latter that dominates error accumulation. This phenomenon has been analysed qualitatively in [**Ise02a**] using WKB analysis. For example, for the Airy equation (1.1) and *any* classical numerical method (multistep or Runge–Kutta scheme, a truncated Taylor expansion...) of an *even* order p the global error is of the form $ct^{p/2+5/4}\chi(t)$, $t \gg 1$, where χ is a highly oscillating function in the interval $[-1,1]$ and $c \neq 0$ is method-dependent.

Needless to say, the numerical solution in Figure 2 is useless and the standard 'panic button' is of little use: decreasing the step-size further or using variable-step strategy rapidly leads to unacceptably small steps [**Ise02b**].

A key to the development of *effective* numerical methods for highly oscillatory problems is to stand the reasoning of the last paragraph on its head: *if differentiating a rapidly oscillating function is so bad, since the amplitude grows, surely integrating such a function must be good.* Thus, we need somehow to reformulate the underlying differential equation so that its numerical solution entails integration of known highly oscillatory functions.

We commence from a highly oscillatory ordinary differential (ODE) system

$$\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = \boldsymbol{f}(t,\boldsymbol{y}), t \geq t_0, \qquad \boldsymbol{y}(t_0) = \boldsymbol{y}_0. \tag{1.2}$$

To time-step from t_n to $t_{n+1} = t_n + h$, it is convenient to rewrite (1.2) in the form

$$\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = A(t)\boldsymbol{y} + \boldsymbol{g}(t,\boldsymbol{y}), \quad t \geq t_n, \qquad \boldsymbol{y}(t_n) = \boldsymbol{y}_n,$$

where $\boldsymbol{g}$ is, in some sense, small. A good choice is to let A be the Jacobian matrix or its approximation.

Although it is possible to construct counterexamples, in most cases of interest the matrix A is stable (no eigenvalues have positive real parts) and it has some large eigenvalues in $\mathrm{i}\mathbb{R}$. In other words, the solution of the 'frozen' system

$$\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = A(\tilde{t})\boldsymbol{y}, \quad t \geq t_n, \qquad \boldsymbol{y}(t_n) = \boldsymbol{y}_n,$$

where $\tilde{t} \geq t_n$, is highly oscillatory. To time-step (1.2) from t_n to $t_{n+1} = t_n + h$, we choose $\alpha \geq 0$, set $\tilde{A} = A(t_n + \alpha h)$, represent the solution in the form

$$\boldsymbol{y}(t_n + \tau) = \mathrm{e}^{\tau\tilde{A}}\boldsymbol{x}(\tau), \qquad \tau \geq 0 \tag{1.3}$$

and let $\boldsymbol{y}_{n+1} = \mathrm{e}^{h\tilde{A}}\boldsymbol{x}_1$, where $\boldsymbol{x}_1$ is our approximation of $\boldsymbol{x}(h)$. For nonlinear equations it makes sense to choose $\alpha = 0$ but if the system (1.2) is linear and the Jacobian matrix is independent of $\boldsymbol{y}$, a superior strategy is to exploit time symmetry and set $\alpha = \frac{1}{2}$.

Following [**Ise02a**], we solve for the function $\boldsymbol{x}$ in (1.3), observing that it obeys the modified differential equation

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\tau} = B(\tau)\boldsymbol{x} + \mathrm{e}^{-\tau\tilde{A}}\boldsymbol{g}(t_n + \tau, \mathrm{e}^{\tau\tilde{A}}\boldsymbol{x}), \quad \tau \geq 0, \qquad \boldsymbol{x}(0) = \boldsymbol{y}_n. \tag{1.4}$$

where

$$B(\tau) = \mathrm{e}^{-\tau\tilde{A}}[A(t_n+\tau) - \tilde{A}]\mathrm{e}^{\tau\tilde{A}}$$

At a first glance, the ODE system (1.4) is nothing but a means of scrambling (1.2) into a more complicated form, a familiar mathematical vice. Yet, it has one redeeming virtue in comparison with the original system: the matrix $B(\tau) = \mathrm{e}^{-\tau\tilde{A}}[A(t+\tau) - \tilde{A}]\mathrm{e}^{\tau\tilde{A}}$ is itself highly oscillatory as a function of τ. This can be exploited by numerical methods that are based upon the integration of the vector field. Such methods are described in Section 2.

The careful reader would have noticed that converting the task in hand to the integration of highly oscillating functions is all very well, but in last analysis such integration is likely to be replaced by quadrature. This is again a numerical problem often foiled by high oscillation. Indeed, the situation is considerably worse than most practitioners would imagine. Suppose that we approximate the integral

$$I_h[f] = \int_0^h f(x)\mathrm{e}^{\mathrm{i}\omega x}\,\mathrm{d}x = h\int_0^1 f(hx)\mathrm{e}^{\mathrm{i}h\omega x}\,\mathrm{d}x, \qquad \omega > 0, \tag{1.5}$$

by a Gauss–Christoffel quadrature [**DR80, Gau68**]. Thus, we choose distinct quadrature nodes $c_1, c_2, \ldots, c_\nu \in [0,1]$, replace the integrand by its polynomial interpolation at the above points and integrate exactly. The outcome is the quadrature formula

$$Q_h^{\mathrm{GC}}[f] = h\sum_{l=1}^{\nu} b_l f(c_l h)\mathrm{e}^{\mathrm{i}c_l h\omega}. \tag{1.6}$$

How good is (1.6)? Everybody is familiar with the standard analysis of quadrature methods: by design, the method is at least of order ν and this can be improved, provided that the collocation polynomial satisfies favourable orthogonality conditions. This, however, makes sense *only* for suitably small *natural frequency* $h\omega$, when (1.5) is not oscillatory at all. Otherwise, a trivial argument can be used to argue that (1.6) is complete nonsense. For any $f \in \mathrm{L}_1[0,1]$ we know from the Riemann–Lebesgue Lemma that, for fixed $h > 0$, $\lim_{h\omega\to\infty} I_h[f] = 0$. (Actually, we know more: if f is smooth then $I_h[f] \sim O(\omega^{-1})$, $h\omega \to \infty$ – this is a consequence of the van der Corput Lemma and will be used in Section 3.) On the other hand, for large $h\omega$, quadrature (1.6) consists of sampling a virtually-random highly oscillating function $f(hx)\mathrm{e}^{\mathrm{i}h\omega x}$ at ν fixed points in $[0,1]$ and forming a linear combination. Seen from this perspective, it is evident that a Gauss–Christoffel quadrature carries the error of $O(h)$ for $h\omega \gg 1$: the procedure of using a random number generator instead, scaling its output by h, is equally useless and much cheaper. If you wish to rescue Gauss–Christoffel at all costs, the only remedy is to split (1.5) into sufficiently many sub-intervals, each with small natural frequency, and apply (1.6) separately in each sub-interval.

The situation is typical of naive numerical approach to high oscillation: familiar and trusted methods are useless and the remedy appears to be to purge high oscillation by zooming into sufficiently small subproblems. This is *the exact opposite* of the approach advocated in this paper, to use algorithms that thrive on rapid oscillation. Thus, in Section 3 we present an analysis, further elaborated in [**Ise04**] and [**Ise03**], of the *Filon method,* a superior alternative to (1.6) whose performance improves in the presence of high oscillation.

2. Highly oscillatory ODE solvers

Our point of departure is the modified ODE system (1.4) but, for ease of exposition, we commence by assuming $\boldsymbol{g} \equiv \mathbf{0}$, namely that (1.2) is a linear system. In Subsection 2.4 we will comment briefly how to extend the methods to a fully nonlinear setting.

2.1. The Magnus method.

Given the matrix linear ODE system

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = B(t)Y, \quad t \geq t_0, \qquad Y(t_0) = Y_0, \tag{2.1}$$

we represent its solution in the form $Y(t) = \mathrm{e}^{\Omega(t)}Y_0$ and treat Ω as a new unknown. It is possible to show that

$$\Omega' = B - \tfrac{1}{2}[\Omega, B] + \tfrac{1}{12}[\Omega,[\Omega,B]] - \tfrac{1}{720}[\Omega,[\Omega,[\Omega,[\Omega,B]]]] + \cdots, \quad t \geq 0, \tag{2.2}$$

where $\Omega(0) = O$. A crucial feature of (2.2) renders it of an interest in geometric numerical integration. Let $\mathcal{G}$ be a finite-dimensional Lie group and $\mathfrak{g}$ the corresponding Lie algebra. If $Y_0 \in \mathcal{G}$ and $B(t) \in \mathfrak{g}$ for $t \geq t_0$ then $Y(t)$ evolves in $\mathcal{G}$ and it is often important to retain this structural feature of (2.1) under discretization. However, $\mathcal{G}$ is, in general, a nonlinear construct and the scope for retention of its structure by a direct application of numerical solvers to (2.1) is limited. On the other hand, $\mathfrak{g}$ is a *linear* space and, as long as we solve (2.2) by employing just linear-space operations and commutators, we are assured of respecting its structure. Hence the interest in using (2.2) as a platform for the development of *Lie-group methods,* an endeavour described at length in [**IMKNZ00**].

A powerful technique to solve (2.2) has been introduced by Wilhelm Magnus [**Mag54**] and analysed in detail from the numerical standpoint by Iserles and Nørsett [**IN99**]. Thus, the function Ω can be expanded in the form

$$\begin{aligned}\Omega(t) \;=\; & \int_0^t B(x)\,\mathrm{d}x - \tfrac{1}{2}\int_0^t\int_0^{x_1}[B(x_2),B(x_1)]\,\mathrm{d}x_2\,\mathrm{d}x_1 \\ & + \tfrac{1}{4}\int_0^t\int_0^{x_1}\int_0^{x_2}[[B(x_3),B(x_2)],B(x_1)]\,\mathrm{d}x_3\,\mathrm{d}x_2\,\mathrm{d}x_1 \\ & + \tfrac{1}{12}\int_0^t\int_0^{x_1}\int_0^{x_1}[B(x_3),[B(x_2),B(x_1)]]\,\mathrm{d}x_3\,\mathrm{d}x_2\,\mathrm{d}x_1 + \cdots .\end{aligned} \tag{2.3}$$

Truncating (2.3) results in a high-quality numerical method. We observe, following the discussion in Section 1, that the algorithm consists of (multivariate) integration, hence is highly suitable for rapidly oscillating ODE systems, once we apply it to the linear system

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = B(\tau)\boldsymbol{x}, \;\; \tau \geq 0, \quad \boldsymbol{x}(0) = \boldsymbol{y}_n, \quad \text{where} \quad B(t) = \mathrm{e}^{-\tau\tilde{A}}[A(t_n+\tau) - \tilde{A}]\mathrm{e}^{\tau\tilde{A}}.$$

In particular, we obtain a seventh-order method for the highly oscillatory ODE (1.2) (with $\boldsymbol{g} \equiv \mathbf{0}$) by letting $\boldsymbol{y}_{n+1} = \mathrm{e}^{h\tilde{A}}\mathrm{e}^{\tilde{\Omega}(h)}\boldsymbol{y}_n$, where

$$\tilde{\Omega}(h) = \int_0^h B(x)\,\mathrm{d}x - \tfrac{1}{2}\int_0^h\int_0^{x_1}[B(x_2),B(x_1)]\,\mathrm{d}x_2\,\mathrm{d}x_1. \tag{2.4}$$

2.2. The Cayley method. A major shortcoming of the Magnus method is that it requires the calculation of a matrix exponential, a task that is prohibitively expensive when the number of variables is large. In our context, we need to compute the exponential and its inverse once per step, to transform the variables in line with (1.3), while using the Magnus expansion means further computation of another exponential, that of $\tilde{\Omega}(h)$. Although a matrix exponential can be suitably approximated and two exponentials can be conveniently amalgamated by using the BCH formula [**IMKNZ00**], arguably a superior alternative is to dispose of the exponential altogether. In place of the transformation (1.3) we may thus use the *Cayley transformation*

$$\boldsymbol{y}(t_n+\tau)=(I-\tfrac{1}{2}\tau\tilde{A})^{-1}(I+\tfrac{1}{2}\tau\tilde{A})\boldsymbol{x}(\tau), \qquad \tau\geq 0.$$

This replaces (1.4) with the ODE system

$$\frac{\mathrm{d}x}{\mathrm{d}\tau}=B(\tau)\boldsymbol{x}+R^{-1}(\tau\tilde{A})\boldsymbol{g}(t_n+\tau,R(\tau\tilde{A})\boldsymbol{x}), \quad \tau\geq 0, \qquad \boldsymbol{x}(0)=\boldsymbol{y}_0,$$

where

$$B(\tau)=(I+\tfrac{1}{2}\tau\tilde{A})^{-1}(I-\tfrac{1}{2}\tau\tilde{A})[A(t_n+\tau)-\tilde{A}-\tfrac{1}{4}\tau^2A(t_n+\tau)\tilde{A}^2](I-\tfrac{1}{2}\tilde{A})^{-2},$$

while $R(z)=(1+\frac{1}{2}z)/(1-\frac{1}{2}z)$.

The above procedure might seem advantageous, since forming $B(\tau)$ requires the solution of two linear algebraic systems, a task which is easier than the calculation of the exponential and allows greater latitude in the exploitation of possible sparsity and structure in the matrix $\tilde{A}$. Unfortunately, the cubic term in the square brackets means that, unlike the matrix that we have encountered in the previous subsection, B need not be small. This has been confirmed by unpublished numerical experiments and motivates us to abandon this approach,

The exponential transformation (1.3) is here to stay, at least for the time being. However, we can follow the reasoning in [**DR02**] and avoid the computation of the second exponential, associated with the Magnus method. The matrix $\tilde{A}$ has often a great deal of structure that renders it amenable to an affordable calculation (or approximation) of its exponential: one example will feature in the next subsection and several publications have been devoted to effective approximation of the exponential in this setting [**IMKNZ00, IZ02, MKQZ01**]. On the other hand, the matrix $\tilde{\Omega}(h)$ of the last subsection is likely to have neither serendipitous structure nor sparsity, hence the calculation of its exponential is usually a more formidable task. The following technique dispenses with this calculation altogether.

Taking a leaf from [**Ise01**], we solve (1.4) by a Cayley transformation: letting $\boldsymbol{x}(\tau)=R(\Omega(\tau))\boldsymbol{y}_n$, we observe that the matrix function Ω obeys the ODE

$$\Omega'=A-\tfrac{1}{2}[\Omega,A]-\tfrac{1}{4}\Omega A\Omega, \quad t\geq 0, \qquad \Omega(0)=O. \tag{2.5}$$

(We have assumed for simplicity, and consistently with the framework of Subsection 2.1, that $\boldsymbol{g}\equiv\boldsymbol{0}$.) The solution of (2.5) can be expanded similarly to that of the equation (2.2),

$$\begin{aligned}\Omega(h) \;=\; & \int_0^h B(x)\,\mathrm{d}x - \tfrac{1}{2}\int_0^h\int_0^{x_1}[B(x_2),B(x_1)]\,\mathrm{d}x_2\,\mathrm{d}x_1 \\ & + \tfrac{1}{4}\int_0^h\int_0^{x_1}\int_0^{x_2}[[B(x_3),B(x_2)],B(x_1)]\,\mathrm{d}x_3\,\mathrm{d}x_2\,\mathrm{d}x_1 \\ & - \tfrac{1}{4}\int_0^h\int_0^{x_1}B(x_2)\,\mathrm{d}x_2 B(x_1)\int_0^{x_1}B(x_2)\,\mathrm{d}x_2\,\mathrm{d}x_1 + \cdots .\end{aligned}$$

Observe that the first two terms above coincide with the truncated Magnus expansion (2.4). Since for all intents and purposes this is all we need, we conclude that there is little *practical* difference between Magnus and Cayley: the only real distinction is that, in place of calculating $\exp\tilde{\Omega}(h)$, we compute $R(\tilde{\Omega}(h))$.

Note further that the Cayley transformation is a Lie-group method only for some Lie groups [**Ise01**]. Yet, this is of little relevance in the context of highly oscillatory problems. The important feature of the method is that it reduces time-stepping to the computation of integrals of highly oscillatory functions.

2.3. The Neumann method. The Magnus method is highly suitable for small systems, while the scope of the Cayley method is wider. However, none can be advocated for the very large systems originating in a semi-discretization of evolutionary partial differential equations. This subsection introduces a technique, based on the Neumann expansion, as applied to the linear *Schrödinger equation*

$$\mathrm{i}\frac{\mathrm{d}u}{\mathrm{d}t} = \nabla u + \lambda(t,\boldsymbol{x})u, \qquad t \geq 0, \quad \boldsymbol{x} \in \Theta,$$

with appropriate initial and boundary conditions. We assume here that $\lambda(t,\boldsymbol{x}) \gg 1$ for $t \gg 1$, rendering the solution highly oscillatory. For simplicity, we assume just two spatial dimensions, but the method can be easily generalized.

Discretizing the Laplacian with the standard *five-point stencil,* we obtain the linear ODE system

$$\mathrm{i}\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = [\tfrac{1}{(\Delta x)^2}S + \Lambda(t)]\boldsymbol{u}, \qquad t \geq 0,$$

where Λ is a diagonal matrix. More importantly, the 'discrete Laplacian' S is a partitioned Toeplitz, symmetric, tridiagonal (TST) matrix, made out of smaller TST matrices. Its eigenvector matrix Q is known and, more pertinently, a product of the form $Q\boldsymbol{v}$ can be formed rapidly using the *fast Fourier transform (FFT)* [**Ise96**].

To time-step from t_n to $t_{n+1} = t_n + \Delta t$, we first rewrite the ODE in the form

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = [A(t) + \tilde{\Lambda}(t)]\boldsymbol{u}, \qquad t \geq 0,$$

where

$$A(t) = -\tfrac{\mathrm{i}}{(\Delta x)^2}S - \mathrm{i}\hat{\lambda}(t)I, \quad \tilde{\Lambda}(t) = \mathrm{i}[\hat{\lambda}(t) - \Lambda(t)]$$

and $\hat{\lambda}(t) = \int_\Theta \lambda(t,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}/\mathrm{vol}\,\Theta$. Note that A shares the eigenvectors of S, while the diagonal matrix $\tilde{\Lambda}$ is small. Letting $\tilde{A} = A(t_n + \frac{1}{2}\Delta t)$, we transform variables

$$\boldsymbol{u}(t_n + \tau) = \mathrm{e}^{\tau\tilde{A}}\boldsymbol{v}(\tau), \qquad \tau \geq 0,$$

whereby

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}\tau} = B(\tau)\boldsymbol{v}, \quad \tau \geq 0, \qquad \boldsymbol{v}(0) = \boldsymbol{u}^n, \tag{2.6}$$

and

$$B(\tau) = \mathrm{e}^{-\tau\tilde{A}}[A(t_n+\tau) - \tilde{A} - \tilde{\Lambda}(t_n+\tau)]\mathrm{e}^{\tau\tilde{A}}, \qquad \tau \geq 0.$$

The computation of $\exp(\tau\tilde{A})$ can be accomplished with FFT, hence is affordable even for large systems. However, the matrix B is dense, with far less structure than, say, $\tilde{A}$, and both Magnus and Cayley transformations are prohibitively expensive. We propose instead to use the *Neumann expansion* [**Atk97**, p. 516], occasionally known in an ODE context as the *Peano expansion* [**Gan59**, p. 152]:

$$\boldsymbol{v}(\Delta t) = \sum_{m=0}^{\infty} \mathcal{N}_m(\Delta t)\boldsymbol{u}^n, \tag{2.7}$$

where $\mathcal{N}_0 \equiv I$ and

$$\begin{aligned} \mathcal{N}_m(t) &= \int_0^t B(x)\mathcal{N}_{m-1}(x)\mathrm{d}x \\ &= \int_0^t\int_0^{x_1}\cdots\int_0^{x_{m-1}} B(x_m)B(x_{m-1})\cdots B(x_1)\,\mathrm{d}x_m\,\mathrm{d}x_{m-1}\cdots\mathrm{d}x_1, \quad m \in \mathbb{N}. \end{aligned}$$

The expansion (2.7) possesses a number of useful features. Firstly, it consists of repeated integration of highly oscillatory functions, hence is suitable for rapidly oscillating systems. Secondly, it does not require the computation of exponentials. This, together with the low cost of forming the matrix B in (2.6), means that the expense of time-stepping is fairly moderate. Thirdly, the standard obstacle to the usefulness of Neumann series in, for example, the numerical solution of integral equations, but also for ODEs, is the convergence requirement $\|\boldsymbol{K}\| < 1$, where

$$\boldsymbol{K}[F] = \int_0^t B(x)F(x)\,\mathrm{d}x$$

[**AMO03**]. This, however, is not a problem in the present setting *precisely because we are integrating highly oscillatory functions.*

2.4. Nonlinear equations. Let $\mathrm{e}^{\Omega(\tau)}$ be the fundamental solution of (2.1). The solution of (1.4) can be written using *nonlinear variation of constants*

$$\boldsymbol{x}(\tau) = \mathrm{e}^{\Omega(\tau)}\boldsymbol{y}_n + \int_0^\tau \mathrm{e}^{\Omega(\tau-\xi)}\mathrm{e}^{\xi\tilde{A}}\boldsymbol{g}(t_n+\xi, \mathrm{e}^{\xi\tilde{A}}\boldsymbol{x}(\xi))\,\mathrm{d}\xi.$$

This is an implicit equation, but it can be solved in few *waveform relaxation* iterations,

$$\begin{aligned} \boldsymbol{x}^{[0]}(\tau) &\equiv \boldsymbol{y}_n, \\ \boldsymbol{x}^{[l+1]}(\tau) &= \mathrm{e}^{\Omega(\tau)}\boldsymbol{y}_n + \int_0^\tau \mathrm{e}^{\Omega(\tau-\xi)}\mathrm{e}^{\xi\tilde{A}}\boldsymbol{g}(t_n+\xi, \mathrm{e}^{\xi\tilde{A}}\boldsymbol{x}^{[l]}(\xi)\,\mathrm{d}\xi, \qquad l \in \mathbb{Z}_+. \end{aligned}$$

Typically, two or three iterations suffice and we observe that each requires integration with highly oscillating integrands.

This brief section represents work in progress, which is close in spirit to recent methods for stiff ordinary differential equations, pioneered by Kassam and Trefethen [**KT04**].

2.5. An example: the Airy equation. Converting the linear oscillator

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + g(t)y = 0, \quad t \geq 0, \qquad y(0) = y_0, \quad \frac{\mathrm{d}y(0)}{\mathrm{d}t} = y_0',$$

where we assume that $g(t) \gg 1$ for $t \gg 1$, into a two-variable system and subjecting it to the transformation (1.4) results in

$$B(\tau) = [g(t_n + \tau) - \omega^2] \begin{bmatrix} (2\omega)^{-1} \sin 2\omega\tau & \omega^{-2} \sin^2 \omega\tau \\ -\cos^2 \omega\tau & -(2\omega)^{-1} \sin 2\omega\tau \end{bmatrix},$$

where $\omega = \sqrt{g(t_n + \frac{1}{2}h)} \gg 1$.

We evaluate just the lowest-order methods, applied to the Airy oscillator (1.1). Thus, Magnus and Cayley, both with just a single integral, yield

$$\boldsymbol{y}_{n+1} = \mathrm{e}^{h\tilde{A}} \mathrm{e}^{\tilde{\Omega}(h)} \boldsymbol{y}_n \qquad \text{and} \qquad \boldsymbol{y}_{n+1} = \mathrm{e}^{h\tilde{A}} R(\tilde{\Omega}(h)) \boldsymbol{y}_n$$

respectively, where

$$\tilde{\Omega}(h) = (\sin h\omega_n - h\omega_n \cos h\omega_n) \begin{bmatrix} \dfrac{\cos h\omega_n}{8\omega_n^3} & \dfrac{\sin h\omega_n}{4\omega_n^4} \\ \dfrac{\sin h\omega_n}{4\omega_n^2} & -\dfrac{\cos h\omega_n}{8\omega_n^3} \end{bmatrix}, \qquad \omega_n = \sqrt{t_n + \tfrac{1}{2}h}.$$

Moreover, the Neumann expansion with a single term reduces to

$$\boldsymbol{y}_{n+1} = \mathrm{e}^{h\tilde{A}} [I + \tilde{\Omega}(h)] \boldsymbol{y}_n.$$

In the present case, one can think of the Cayley and Neumann expansions as approximating Magnus through the replacement of e^z by its [1/1] and [1/0] Padé approximations respectively.

Figure 3 displays the global error in solving the Airy equation (1.1) with the above Magnus method and the step size $h = \frac{1}{4}$. The contrast with Figure 2 could not have been greater and it represents a tangible demonstration of the advantage of the approach of this paper. It has been proved in [**Ise02a**] that the global error when (1.1) is solved with Magnus accumulates like $ct^{-1/4}\chi(t)$, where χ oscillates in $[-1, 1]$. This is confirmed in the bottom plot in Figure 3.

Both Cayley and Neumann produce results which are virtually indistinguishable from Figure 3. Specifically, the Cayley and Magnus solutions differ by just $\approx 10^{-10}$, whether Neumann departs from Magnus by at most $\approx 10^{-5}$.

3. Quadrature of highly oscillatory integrals

Discretising highly oscillatory ODEs by converting the problem into the integration of highly oscillatory functions is a powerful approach which can be easily spoilt by bad quadrature. Occasionally, like in Subsection 2.5, we can evaluate integrals in an explicit form, but this is an exception and, in general, we need to use quadrature. Having already observed in Section 1 that a Gauss–Christoffel quadrature (1.6) is unsuitable for our needs, we look elsewhere for numerical salvation. Fortunately, quality quadrature methods for highly oscillatory integrals are available in the literature, e.g. the Zamfirescu method [**Zam63**], the Levin method [**Lev96**] and, in particular, the Filon method [**Fil28, Fli60**]. With the sole exception of Levin's method, the analysis of these algorithms was lacking until recently [**Ise04, Ise03**]. We simply have had no idea how good such methods can be.

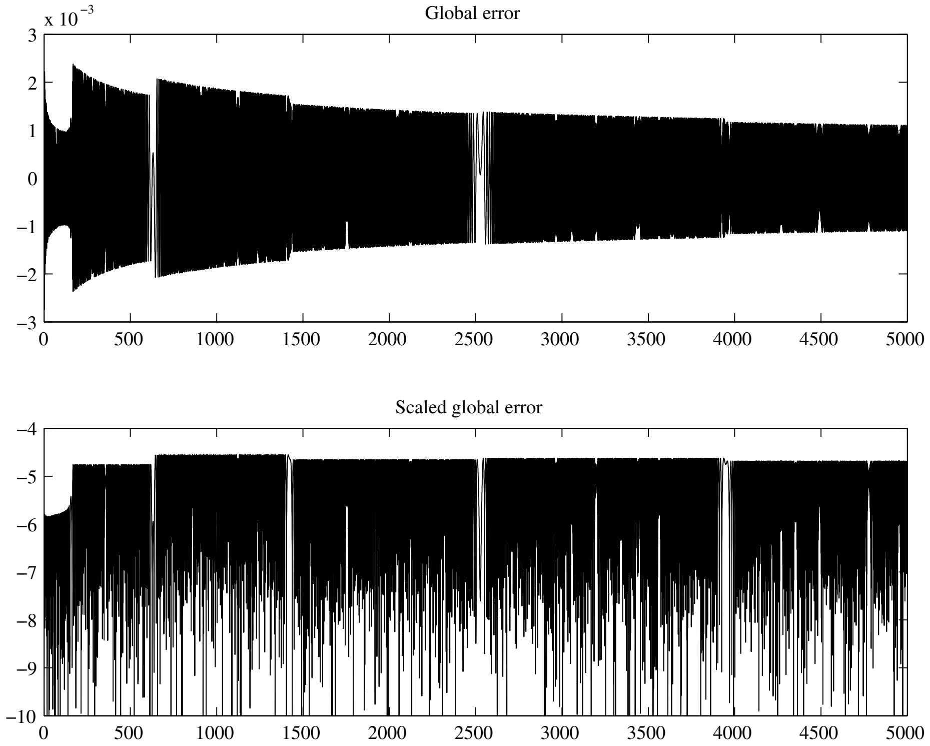

FIGURE 3. The error in solving (1.1) with the single-term Magnus method and $h = \frac{1}{4}$. The top plot displays the global error $|y_n - y(t_n)|$ and the bottom the scaled error $\log|y_n - y(t_n)| + \frac{1}{4}\log t$.

3.1. The Filon method. In the sequel we focus on the Filon method, since it is the most general and natural of quadrature methods for highly oscillatory integrals and also since it is the most comprehensively understood. It is, indeed, remarkable that a method introduced (by Louis Napoleon George Filon, [**Fil28**]) in 1928 and then either forgotten or often misunderstood turns out to be so effective.

The idea is really simple: to approximate the integral

$$I[f] = \int_0^1 f(x)K(\omega, x)\,\mathrm{d}x, \tag{3.1}$$

where K oscillates rapidly for $\omega \gg 1$, we choose quadrature nodes $c_1 < c_2 < \cdots < c_\nu$ in $[0, 1]$ and interpolate $f(x)$ (rather than interpolating $f(x)K(\omega, x)$ in Gauss–Christoffel) there by a polynomial of degree $\nu - 1$. Provided that we can evaluate the moments

$$\mu_m(\omega; K) = \int_0^1 x^m K(\omega, x)\,\mathrm{d}x, \qquad m \in \mathbb{Z}_+,$$

explicitly for $0 \le m \le \nu - 1$, substitution of the interpolating polynomial in place of f in (3.1) results in the *Filon quadrature*

$$Q^{\mathrm{F}}[f] = \sum_{l=1}^{\nu} b_l(\omega) f(c_l). \tag{3.2}$$

We note in passing that moments can be evaluated explicitly, e.g. with standard symbolic algebra software, for a wide range of functions K of practical interest, e.g. $\mathrm{e}^{\mathrm{i}\omega(x-d)^m}$ for $d \in [0,1]$ and $m \in \{1,2,3\}$, $\mathrm{e}^{\mathrm{i}\omega|x-d|}/|x-d|^\alpha$ for $d \in [0,1]$ and $\alpha < 1$, the Bessel oscillator $\mathrm{J}_\nu(\omega x)$ and the Airy oscillator $\mathrm{Ai}(-\omega x)$.

Numerical experiments with (3.2) are remarkable, while elementary and intuitive arguments (sadly missing from some early treatments of Filon's method, that should be better left in their obscurity) imply that the method is likely to be very effective. In particular, by design the method integrates *exactly* all polynomials f of degree $\le \nu - 1$ and, unlike $f(x)K(\omega,x)$, a smooth function f can be interpolated very well by such polynomials. Having said this, there is a need for more formal analysis, both to understand the method better and to develop suitable strategies for good choice of quadrature nodes.

3.2. Regular oscillators. We commence our analysis from the regular oscillator $K(\omega,x) = \mathrm{e}^{\mathrm{i}\omega x}$, i.e.

$$I[f] = \int_0^1 f(x)\mathrm{e}^{\mathrm{i}\omega x}\,\mathrm{d}x. \tag{3.3}$$

This is equivalent to setting $h = 1$ in (1.5): it is easy to incorporate an additional small parameter $h > 0$ into the theory but we opt for (3.3) for simplicity's sake.

It is interesting to consider the quadrature error $E[f] = Q^{\mathrm{F}}[f] - I[f]$ in three regimes: $0 < \omega \ll 1$, $\omega = O(1)$ and $\omega \gg 1$. The first corresponds to a non-oscillatory regime and it is of interest to check how much (if at all) Filon is worse than Gauss–Christoffel in this setting. The second regime is consistent with a persistent numerical 'folk wisdom', namely that for optimal results a function should be sampled a fixed number of times per period. Finally, the regime $\omega \gg 1$ addresses our main concern, highly oscillatory problems.

A comprehensive analysis in [**Ise04**] determines $E[f]$ in all three regimes for arbitrary smooth functions f.

(1) Suppose that the Gauss–Christoffel method for (3.3) with $\omega = 0$, using the same nodes, is of order $p \in \{\nu, \nu+1, \ldots, 2\nu\}$: this is equivalent to

$$\int_0^1 x^m \gamma(x)\,\mathrm{d}x = 0, \quad m = 0,1,\ldots,p-\nu-1,$$

where $\gamma(t) = \prod_{l=1}^{\nu}(t - c_l)$. Then so is the Filon quadrature (3.2) for $0 < \omega \ll 1$.

(2) If $\omega = O(1)$ then (3.2) recovers only polynomials of degree $\nu - 1$, in other words is of order ν. Thus, the received 'folk wisdom' is perversely wrong: this is the *worst* regime.

(3) If $\omega \gg 1$ then, for a general choice of quadrature nodes, the error is $E[f] = O(\omega^{-1})$ but, if $c_1 = 0$ and $c_\nu = 1$ then $E[f] = O(\omega^{-2})$. The method is still of order ν but, remarkably, for high frequencies it is not polynomial recovery but the oscillation itself that kills the error. The higher the oscillation, the better Filon's quadrature and this most welcome phenomenon can be greatly enhanced by choosing the endpoints as quadrature nodes.

3.3. Irregular oscillators and the geometric model. Let g be a smooth, real function and set $K(\omega, x) = \mathrm{e}^{\mathrm{i}\omega g(x)}$: in place of (3.3) we thus integrate

$$I[f] = \int_0^1 f(x)\mathrm{e}^{\mathrm{i}\omega g(x)}\,\mathrm{d}x. \tag{3.4}$$

In the sequel we focus on the case $\omega \gg 1$, since it is both the most relevant to our discussion and the most comprehensively understood. Our exposition follows [**Ise03**].

In order to examine the behaviour of $E[f]$ for $\omega \gg 1$, it is highly relevant to mention the main step used in [**Ise04**] to elicit the error asymptotics for $g(x) = x$, namely studying the asymptotic behaviour of the moments,

$$\mu_m(\omega; \mathrm{e}^{\mathrm{i}\omega x}) \sim \frac{1^m \mathrm{e}^{\mathrm{i}\omega} - 0^m}{\mathrm{i}\omega} + O(\omega^{-2}), \qquad m \in \mathbb{Z}_+, \quad \omega \gg 1. \tag{3.5}$$

In a manner made precise in [**Ise04**], choosing the nodes $c_1 = 0$ and $c_\nu = 1$ eliminates the influence of the 0^m and 1^m terms in (3.5) on the error, rendering $E[f] = O(\omega^{-2})$. With far greater generality, it has been established in [**Ise03**] that if the moments of K obey the *geometric model*

$$\mu_m(\omega; K) \sim \sum_{j=1}^{r} v_j(\omega)\omega^{-\alpha_j} d_j^m + O(\omega^{-\beta}), \qquad m \in \mathbb{Z}_+, \quad \omega \gg 1, \tag{3.6}$$

where $d_1, d_2, \ldots, d_r \in [0, 1]$, $v_1, v_2, \ldots, v_r \in \mathrm{L}_\infty[0, \infty)$ and

$$0 < \alpha_1, \alpha_2, \ldots, \alpha_r < \beta$$

then, using *all* the points $d_1, d_2, \ldots, d_r$ as quadrature nodes yields $E[f] \sim O(\omega^{-\beta})$. In a manner that can be made precise, choosing each d_j eliminates the influence of $v_j(\omega)\omega^{-\alpha_j}$ on the error.

We have already mentioned an example of K that is consistent with the geometric model (3.6), namely $K(\omega, x) = \mathrm{e}^{\mathrm{i}\omega x}$. In that case $r = 2$, $d_1 = 0$, $d_2 = 1$, $v_1(x) \equiv -1$, $v_2(x) = \mathrm{e}^{\mathrm{i}x}$, $\alpha_1 = \alpha_2 = 1$ and $\beta = 2$. [**Ise03**] establishes very broad conditions for the consistency of $K(\omega, x) = \mathrm{e}^{\mathrm{i}\omega g(x)}$ with the geometric model.

The simplest case is when $g' \neq 0$ in $[0, 1]$ and either g'' does not change sign there, whereby we let $s = 1$, or there exists $s \geq 3$ so that $g^{(s)} \neq 0$ in $[0, 1]$. (Such condition is always satisifed when g is a polynomial.) Then (3.6) is satisfied with $r = 2$, $d_1 = 0$, $d_2 = 1$, $\alpha_1 = \alpha_2 = 1/s$ and $\beta = 2/s$. The proof uses the van der Corput Lemma from harmonic analysis [**Ste93**].

The case when g' vanishes at distinct points in $(0, 1)$ is more complicated. The van der Corput lemma is combined with the method of stationary phase from perturbation theory, using bump functions. Firstly, suppose that g' vanishes at $\tilde{d} \in (0, 1)$ and nowhere else in $[0, 1]$. In other words, there exists an integer $\delta \geq 1$ such that

$$g'(\tilde{d}) = g''(\tilde{d}) = \cdots = g^{(\delta-1)}(\tilde{d}) = 0, \quad g^{(\delta)}(\tilde{d}) \neq 0.$$

The method of stationary phase can be used to prove that there exist functionals a_i such that

$$I[f] \sim \omega^{1/\delta} \sum_{i=1}^{\infty} a_i[f]\omega^{-i/\delta}, \qquad \omega \gg 1,$$

for every function f with a sufficiently small compact support in a neighbourhood of $\tilde{d}$ [**Ste93**].

Suppose now that g' vanishes at the $r-2$ distinct points $d_j \in (0,1)$ and that the lowest-order nonzero derivative at d_j is $g^{(\delta_j)}$. We use bump functions to 'stitch' compactly-supported functions together, adding contributions from the endpoints $d_1 = 0$ and $d_r = 1$. It then follows that g obeys the quadratic model with $\alpha_1 = \alpha_r = 1$ and $\alpha_j = 1/\delta_j$ for $j = 2, \ldots, r-1$. Also the value of β can be determined and it depends on the parity of the integers δ_j [**Ise03**].

The analysis may break down when g' vanishes at an endpoint, e.g. in the case $K(\omega, x) = \mathrm{e}^{\mathrm{i}\omega x^2}$. Otherwise it provides the optimal strategy for the choice of quadrature points in the Filon method: we need $\nu \geq r$ and

$$d_1, d_2, \ldots, d_r \in \{c_1, c_2, \ldots, c_\nu\}.$$

3.4. Multivariate integration. Higher-order methods from Section 2 require multivariate integration, a notoriously difficult computational task. Fortunately, the special form of Magnus and Cayley integrals renders them amenable to a very effective quadrature [**IMKNZ00**] and a future publication will extend this to the Neumann expansion.

The first stage is to reduce the expense in function evaluations. Thus, suppose that $c_1, c_2, \ldots, c_\nu \in [0,1]$ are given quadrature nodes and that we have evaluated $B(c_l h)$, $l = 1, 2, \ldots, \nu$. It is elementary that the first integral in (2.3) can be integrated with a Gauss–Christoffel quadrature of order p, say. In other words, there exist weights $b_l^{[1]}$, $l = 1, 2, \ldots, \nu$ so that

$$h \sum_{l=1}^{\nu} b_l^{[1]} B(c_l h) = \int_0^h B(x)\,\mathrm{d}x + O(h^{p+1}).$$

It was proved in [**IN99**] that there also exist weights $b_{i,j}^{[2]}$, $i, j = 1, 2, \ldots, \nu$, such that

$$h^2 \sum_{i=1}^{\nu} \sum_{j=1}^{\nu} b_{i,j}^{[2]} [B(c_i h), B(c_j h)] = \int_0^h \int_0^{x_1} [B(x_2), B(x_1)]\,\mathrm{d}x_2\,\mathrm{d}x_1 + O(h^p).$$

Similarly for *all* Magnus or Cayley integrals: with just ν evaluations of the matrix B we can evaluate the entire truncated expansion to order p.

The drawback of this approach is that, although the number of function evaluations is very small, this is not the case with the volume of linear algebra, since we need to evaluate a very large number of commutators. The number of commutators, however, can be reduced a very great deal by techniques pioneered by Munthe-Kaas and Owren and by Blanes, Casas and Ros [**IMKNZ00**].

All this is true for 'standard' Magnus or Cayley, where Lagrange points represent the obvious choice of quadrature nodes. It has been proved in [**Ise04**] that this approach remains equally valid for the Filon quadrature and highly oscillating integrands.

4. Conclusion

The present review of recent (and some new) results in the computation of highly oscillatory problems represents a snapshot of work in progress, rather than a mature and comprehensive theory. Yet, assembling all the above results does more than just presenting a number of disjoint algorithms, useful as they might be. General principles emerge.

Firstly, highly oscillatory differential equations should be solved by algorithms that integrate highly oscillating functions. There might be diverse ways of designing such algorithms: we have described three in Section 2 but more and better might well exist.

Secondly, quadrature of highly oscillatory functions is possible, indeed easy, but its optimal implementation calls for an understanding of the asymptotic behaviour of moments.

Thirdly, and most importantly, high oscillation is not an enemy of computation. Any discretization method that becomes less accurate (or more expensive) in the presence of rapid oscillation is, arguably, a *bad* method. And good methods do exist, although much is yet to be done in this area. The rule of a thumb is that high oscillation should assist precision and lower the cost of computation. The main lesson of this paper is that this is an attainable goal.

References

[AMO03] N. D. Aparicio, S. J. A. Malham, and M. Oliver, *Precomputation methods and the Evans function*, Tech. report, Heriot-Watt University, Edinburgh, 2003.

[Atk97] K. E. Atkinson, *The numerical solution of integral equations of the second kind*, Cambridge University Press, Cambridge, 1997.

[DR80] P. J. Davis and P. Rabinowitz, *Methods of numerical integration*, 2nd ed., Academic Press, Orlando, FL, 1980.

[DR02] F. Diele and S. Ragni, *The global error of Magnus methods based on the Cayley map for some oscillatory problems*, Tech. report, University of Bari, 2002.

[Fil28] L. N. G. Filon, *On a quadrature formula for trigonometric integrals*, Proc. Royal Soc. Edinburgh **49** (1928), 38–47.

[Fli60] E. A. Flinn, *A modification of Filon's method of numerical integration*, J ACM **7** (1960), 181–184.

[Gan59] F. R. Gantmacher, *Applications of the theory of matrices*, Wiley Interscience, New York, 1959.

[Gau68] W. Gautschi, *Construction of Gauss–Christoffel quadrature formulas*, Maths Comp. **22** (1968), 251–270.

[HNW93] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving ordinary differential equations I: Nonstiff problems*, 2nd revised ed., Springer Verlag, Berlin, 1993.

[IMKNZ00] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna, *Lie-group methods*, Acta Numerica **9** (2000), 215–365.

[IN99] A. Iserles and S. P. Nørsett, *On the solution of linear differential equations in Lie groups*, Phil. Trans Royal Society A **357** (1999), 983–1020.

[Ise96] A. Iserles, *A first course in the numerical analysis of differential equations*, Cambridge University Press, Cambridge, 1996.

[Ise01] ———, *On Cayley-transform methods for the discretization of Lie-group equations*, Found. Comp. Maths **1** (2001), 129–160.

[Ise02a] ———, *On the global error of discretization methods for highly-oscillatory ordinary differential equations*, BIT **42** (2002), 561–599.

[Ise02b] ———, *Think globally, act locally: Solving highly-oscillatory ordinary differential equations*, Appld Num. Anal. **43** (2002), 145–160.

[Ise03] ———, *On the numerical quadrature of highly-oscillating integrals II: Irregular oscillators*, Tech. Report NA2003/09, DAMTP, University of Cambridge, 2003.

[Ise04] ———, *On the numerical quadrature of highly-oscillating integrals I: Fourier transforms*, IMA J. Num. Anal. (2004), to appear.

[IZ02] A. Iserles and A. Zanna, *Efficient computation of the matrix exponential by generalized polar decompositions*, Tech. Report NA2002/09, DAMTP, University of Cambridge, 2002.

[KT04] A.-K. Kassam and L. N. Trefethen, *Fourth-order time stepping for stiff PDEs*, SIAM J. Sc. Comput. (2004), to appear.

[Lev96] D. Levin, *Fast integration of rapidly oscillatory functions*, J. Comput. Appl. Maths **67** (1996), 95–101.

[Mag54] W. Magnus, *On the exponential solution of differential equations for a linear operator*, Comm. Pure and Appl. Math **VII** (1954), 649–673.

[MKQZ01] H. Munthe-Kaas, R. G. W. Quispel, and A. Zanna, *Generalized polar decompositions on Lie groups with involutive automorphisms*, Found. Comp. Maths **1** (2001), 297–324.

[Ste93] E. Stein, *Harmonic analysis: Real-variable methods, orthogonality, and oscillatory integrals*, Princeton University Press, Princeton, NJ, 1993.

[Zam63] I. Zamfirescu, *An extension of Gauss's method for the calculation of improper integrals*, Acad. R.P. Romîne Stud. Cerc. Mat. **14** (1963), 615–631, (in Romanian).

Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, Wilberforce Rd, Cambridge CB3 0WA, United Kingdom

E-mail address: ai@damtp.cam.ac.uk