

# On highly oscillatory problems arising in electronic engineering

Marissa Condon  
School of Electronic Engineering  
Dublin City University  
Dublin 9  
Ireland

Alfredo Deaño  
Department of Applied Mathematics and Theoretical Physics  
Centre for Mathematical Sciences  
University of Cambridge  
Wilberforce Rd, Cambridge CB3 0WA, UK

Arieh Iserles  
Department of Applied Mathematics and Theoretical Physics  
Centre for Mathematical Sciences  
University of Cambridge  
Wilberforce Rd, Cambridge CB3 0WA, UK

August 1, 2008

## Abstract

In this paper, we consider linear ordinary equations originating in electronic engineering, which exhibit exceedingly rapid oscillation. Moreover, the oscillation model is completely different from the familiar framework of asymptotic analysis of highly oscillatory integrals.

Using a Bessel-function identity, we expand the oscillator into asymptotic series, and this allows us to extend Filon-type approach to this setting. The outcome is a time-stepping method that guarantees high accuracy regardless of the rate of oscillation.

## 1 Introduction

The focus of our attention in this paper is the discretization of ordinary differential equations of the form

$$\mathbf{y}' = A\mathbf{y} + E(t)\mathbf{g}(t), \quad t \geq 0, \quad \mathbf{y}(0) = \mathbf{y}_0 \in \mathbb{R}^d, \quad (1.1)$$

where  $A$  is a  $d \times d$  matrix,  $\mathbf{g}$  is a  $d$ -vector of functions while  $E$  is a  $d \times d$  matrix function,  $E_{k,l}(t) = \chi_{k,l} e^{\tau_{k,l} \sin \omega_{k,l} t}$ ,  $k, l = 1, \dots, d$ . While we may assume that the eigenvalues of  $A$

are of moderate size, the terms of  $E$  are highly oscillatory, since we allow for  $\max \omega_{k,l} \gg 1$ . Moreover, it is perfectly possible for different frequencies  $\omega_{k,l}$  to differ in size by many orders of magnitude.

The equation (1.1) is a model of more complicated, in general nonlinear, differential equations originating in electronic engineering. High-frequency signals abound in Radio Frequency (RF) communication systems. This is a consequence of the need for modulation: the imposition of a lower-frequency information signal onto a high-frequency carrier. The goal is to enable antennae of a manageable size to be employed for audio transmission. Antennae of the order of several miles to several thousand miles would be required if modulation was not performed. In RF communication systems, signals in the MHz frequency range and higher are common. Furthermore, nonlinearities abound in RF transmission systems owing to the presence of solid-state amplifiers, mixers and so on (Jeruchim, Balaban & Shanmugan 2000).

Most RF systems involve a linear part and a nonlinear part with the linear part due to the presence of linear resistors, inductors and capacitors and the nonlinear part due to amplifiers, mixers or nonlinear and controlled resistors and capacitors. The equations (1.1) are a simplified model with many of the nonlinearities approximated by linear terms. The occurrence of the  $e^{\tau_{k,l} \sin \omega_{k,l} t}$  is due to the input of sine-waves to terminals of circuits with diodes or transistors.

The recent explosion of developments in the RF and telecommunications industry has put pressure on circuit designers for faster simulations, faster designs and faster product output and the existing Computer Aided Design (CAD) tools have struggled to keep pace. In addition, the growing complexity of the modulation formats is rendering the software tools unacceptably slow and consequently, unsatisfactory. There is therefore, an urgent need for a complete revamp and update of the fundamental numerical processes within these CAD packages taking into account the modern developments and formats.

Some recent work in this direction is that by e.g. (Roychowdhury 2001) and subsequent work by Pulch (2005) and Dautbegovic, Condon & Brennan (2005). However, much more work is required to generate algorithms that are well-suited and effective for the application areas in hand.

On the face of it, solving (1.1) is trivial, because we can write the solution of this linear system explicitly as variation of constants,

$$\mathbf{y}(t_{n+1}) = e^{hA} \mathbf{y}(t_n) + \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-\xi)A} E(\xi) \mathbf{g}(\xi) d\xi, \quad (1.2)$$

where  $t_{n+1} = t_n + h$ . This, however, is not a very helpful observation because of the presence of highly oscillatory terms inside the integral. Specifically, rewriting (1.2) component-wise, we have

$$y_k(t_{n+1}) = \sum_{i=1}^d F_{k,i}(h) y_i(t_n) + \sum_{i=1}^d \sum_{j=1}^d \chi_{i,j} \int_{t_n}^{t_{n+1}} F_{k,i}(t_{n+1} - \xi) e^{\tau_{i,j} \sin \omega_{i,j} \xi} g_j(\xi) d\xi$$

for  $k = 1, \dots, d$ , where  $F(t) = e^{tA}$ . While the computation of the matrix exponential is standard, the intrinsic difficulty is represented by practical computation of integrals of the form

$$\int_{t_n}^{t_{n+1}} F_{k,i}(t_{n+1} - \xi) e^{\tau_{i,j} \sin \omega_{i,j} \xi} g_j(\xi) d\xi$$

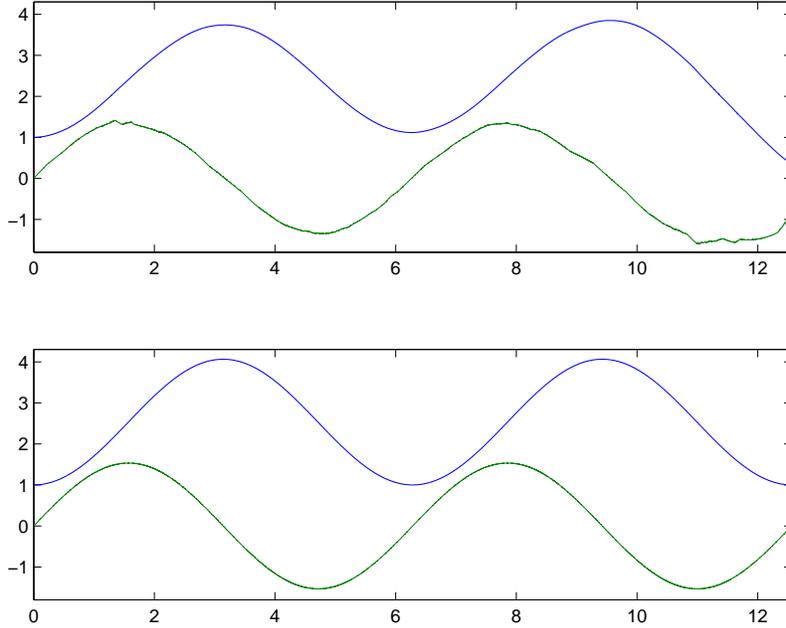


Figure 1.1: The numerical (top, with  $\text{RelTol} = 10^{-4}$ ) and true (bottom) solution of (1.4) in the first two periods.

$$= \frac{1}{2}h \int_{-1}^1 F_{k,i}(\frac{1}{2}h(1-x))g_j(t_n + \frac{1}{2}h(1+x))e^{\tau_{i,j} \sin \omega_{i,j} \left(t_n + \frac{1}{2}h(1+x)\right)} dx \quad (1.3)$$

for  $\omega_{i,j} \gg 1$ . Since classical numerical methods for non-oscillatory integrals, e.g. Gaussian quadrature, require the decomposition of the integration interval into  $\mathcal{O}(\omega)$  sub-panels (Davis & Rabinowitz 1984), and recalling that we have  $d^3$  such intervals in each step, they are completely unfit for purpose.

An alternative is provided by contemporary methods for highly oscillatory quadrature, an area that has undergone significant developments in the last few years. The problem, though, is that the integral (1.3) does not fit into the framework of traditional asymptotic theory for highly oscillatory integrals (Wong 2001): the latter is concerned with integrals of the form  $\int_{\Omega} f(\mathbf{x})e^{i\omega g(\mathbf{x})} d\mathbf{x}$ , where  $\omega \gg 1$  while neither  $f$  nor  $g$  are oscillatory. This is also the case with the methods for numerical calculation of highly oscillatory integrals that have been developed recently (Huybrechs & Vandewalle 2006, Iserles & Nørsett 2005, Olver 2006).

Yet another approach is to disregard the explicit formula (1.2) and use *exponential integrators* to solve the system (1.1). This is not very promising either. Most exponential integrators designed to cope with high oscillations do this in a Hamiltonian setting, which does not fit the paradigm of (1.1) (Grimm & Hochbruck 2006). Moreover, they are not designed to deal

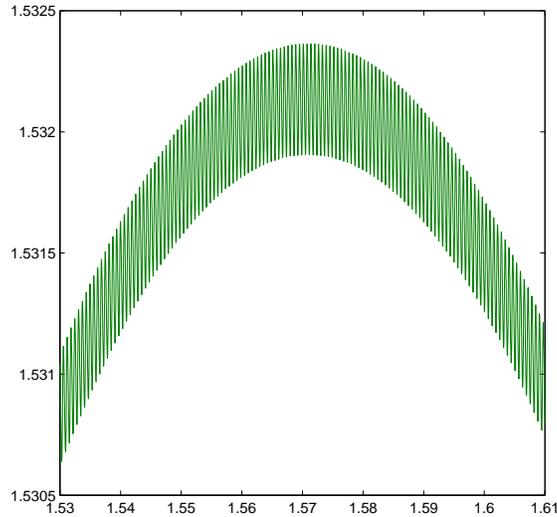


Figure 1.2: A close-up on  $y'$  within a narrow window, exhibiting rapid small-amplitude oscillations.

with the multiscale nature of (1.1) and with truly huge frequencies  $\omega_{i,j}$  therein. An exception to the Hamiltonian setting is provided in (Khanamirian 2008), but this does not advance us much since it takes us to the very same highly oscillatory quadrature methods which we have already deemed unsuitable in the last paragraph.

Finally, we can disregard the special structure of (1.1) and just use an all-purpose ODE solver, placing our trust in its error-control and variable-step strategies. Thus, we have solved the system

$$y'' + y = 2e^{\sin \omega t}, \quad t \geq 0, \quad y(0) = 1, \quad y'(0) = 0, \quad (1.4)$$

with the MATLAB routine `ode45`, employing different error tolerances and setting  $\omega = 10000$ . The solution of (1.4) is periodic with period  $2\pi$  and we have examined a numerical solution across two periods. We have set different values of the relative error tolerance parameter `RelTol`, setting in each case `AbsTol` =  $10^{-3} \times \text{RelTol}$ .

In Fig. 1.1 we present a numerical solution (admittedly, with the least relative error, `RelTol` =  $10^{-4}$ , yet tenfold smaller than the MATLAB default) of (1.4), comparing it with the exact solution. It is evident that the quality of the numerical solution deteriorates fairly rapidly. cursory examination of the exact solution might be misleading, since it appears to be a very ‘nice’ function, varying in a sedate manner. However, once we magnify the solution within a short window, as in Fig. 1.2, we note that it exhibits very rapid, small-amplitude oscillations. Such oscillations are bound to inhibit the step size in any standard error-control mechanism in all-purpose software and this, perhaps unsurprisingly, is reflected in Table 1. Another important observation is that the numerical (absolute) error falls substantially short of either relative or absolute error-tolerance parameters. This breakdown in error control has

Table 1: The performance of `ode45` in the interval  $[0, 4\pi]$  for different relative error tolerances for the system (1.4) with  $\omega = 10^4$ .

RelTol	number of steps	numerical error	
		in $y(4\pi)$	in $y'(4\pi)$
$10^{-4}$	61441	$-6.42_{-01}$	$-8.57_{-01}$
$10^{-5}$	123405	$9.61_{-04}$	$-2.19_{-02}$
$10^{-6}$	240645	$-1.01_{-04}$	$4.57_{-04}$
$10^{-7}$	377057	$1.94_{-06}$	$5.17_{-05}$

been already reported for other highly oscillatory ODE systems (Iserles 2002). Note that (1.4) is a toy problem, not just because we are interested in larger systems with many frequencies, but also because  $\omega = 10^4$  is a fairly small frequency within our framework. Realistic electronic circuits are likely to exhibit fast oscillations in the range of  $\approx 10^8$ . This, clearly, is beyond the scope of any standard ODE software.

The solution that we propose in this paper is to analyse the asymptotic behaviour of the integral (1.3), thereby creating the right tools for the extension of Filon-type quadrature (Iserles & Nørsett 2005) to this setting. This will lead not just to a practical algorithm for the calculation of (1.1) with arbitrarily large frequencies  $\omega_{i,j}$  (indeed, the higher the frequency, the better!) but will also serve us in future generalization of this equation to full nonlinear setting. Finally, asymptotic expansion and numerical computation of the highly oscillatory integral (1.3) and, in future publication, of its generalisations is of an independent mathematical interest.

## 2 The asymptotics of the ExpSin integral

Mindful of (1.3), we are concerned with the asymptotic behaviour of the integral

$$I[f] = \int_{-1}^1 f(x) e^{\tau \sin \omega(\alpha x + \beta)} dx, \quad (2.1)$$

where  $\alpha, \beta \in \mathbb{R}$ ,  $\tau \in \mathbb{C} \setminus \{0\}$  and  $\omega \gg 1$ . For a want of a better name, we call (2.1) the *ExpSin integral*.

Even the briefest examination of (2.1) highlights a crucial difference between the ExpSin integral and the ‘standard model’ of asymptotic theory of highly oscillatory integrals. Thus, suppose that we move the  $\omega$  to front of the sine function. It follows at once from the method of stationary phase (Wong 2001) that

$$\int_{-1}^1 f(x) e^{\tau \omega \sin(\alpha x + \beta)} dx = \mathcal{O}\left(\omega^{-\frac{1}{2}}\right), \quad \omega \gg 1,$$

provided that  $[-\beta + (m + \frac{1}{2})\pi]/\alpha \in [-1, 1]$  for some  $m \in \mathbb{Z}$ ,

$$\int_{-1}^1 f(x) e^{\tau \sin(\alpha x + \beta)} dx = \mathcal{O}\left(\omega^{-1}\right), \quad \omega \gg 1,$$

otherwise. On the other hand, for  $\tau \in \mathbb{R}$  and  $f(x) > 0$ ,  $x \in [-1, 1]$ , it follows at once that

$$0 < 2e^{-1} \min_{x \in [-1, 1]} f(x) \leq \mathbf{I}[f] \leq 2e \max_{x \in [-1, 1]} f(x)$$

and the integral is bounded away from zero uniformly in  $\omega \in \mathbb{R}$ . This is demonstrated in Fig. 2.1.

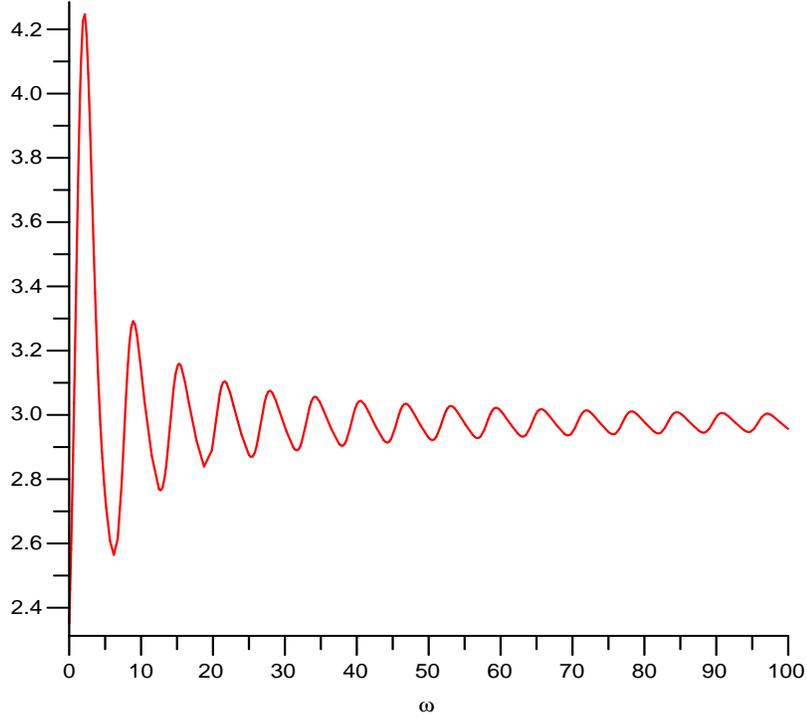


Figure 2.1: The integral  $\mathbf{I}[e^x]$  for  $\alpha = 1$ ,  $\beta = 0$  and  $0 \leq \omega \leq 100$ .

The key step toward the analysis of the ExpSin integral is the identity

$$e^{\tau \sin \theta} = I_0(\tau) + 2 \sum_{k=0}^{\infty} (-1)^k I_{2k+1}(\tau) \sin(2k+1)\theta + 2 \sum_{k=1}^{\infty} (-1)^k I_{2k}(\tau) \sin 2k\theta, \quad (2.2)$$

where  $I_k$  is the  $k$ th *modified Bessel function* (Abramowitz & Stegun 1964, p. 376, formula (9.6.35)). Letting  $\theta = \omega(\alpha x + \beta)$  in (2.1), we thus obtain

$$\begin{aligned} \mathbf{I}[f] &= I_0(\tau) \int_{-1}^1 f(x) dx + 2 \sum_{k=0}^{\infty} (-1)^k I_{2k+1}(\tau) \int_{-1}^1 f(x) \sin((2k+1)\omega(\alpha x + \beta)) dx \\ &\quad + 2 \sum_{k=1}^{\infty} (-1)^k I_{2k}(\tau) \int_{-1}^1 f(x) \cos(2k\omega(\alpha x + \beta)) dx. \end{aligned} \quad (2.3)$$

We thus express  $\mathbf{I}[f]$  as an infinite sum of integrals, all of which (except for the first) are themselves highly oscillatory. Before we expand these integrals in turn, it is useful to comment further about this sum. Since all oscillatory integrals are  $o(1)$  for  $\omega \gg 1$ , we deduce that

$$\lim_{\omega \rightarrow \infty} \mathbf{I}[f] = I_0(\tau) \int_{-1}^1 f(x) dx.$$

Moreover, we can deduce at once from (Abramowitz & Stegun 1964, p. 365, formula (9.3.1)) that

$$I_k(\tau) \sim \frac{1}{2\pi k} \left( \frac{e\tau}{2k} \right)^k, \quad k \gg 1.$$

Since the highly oscillatory integrals are small (as we will see soon, they are  $\mathcal{O}(\omega^{-1})$ ), we conclude that the infinite series converge very rapidly, at a spectral speed.

Let

$$C_{\sigma,\rho}[f] = \int_{-1}^1 f(x) \cos(\sigma x + \rho) dx, \quad S_{\sigma,\rho}[f] = \int_{-1}^1 f(x) \sin(\sigma x + \rho) dx,$$

therefore (2.3) becomes

$$\begin{aligned} \mathbf{I}[f] &= I_0(\tau) \int_{-1}^1 f(x) dx + 2 \sum_{k=0}^{\infty} (-1)^k I_{2k+1}(\tau) S_{(2k+1)\omega\alpha, (2k+1)\omega\beta}[f] \\ &\quad + 2 \sum_{k=1}^{\infty} (-1)^k I_{2k}(\tau) C_{2k\omega\alpha, 2k\omega\beta}[f]. \end{aligned} \quad (2.4)$$

Let us assume that  $f \in C^\infty[-1, 1]$ . It is fairly straightforward, although laborious, to expand  $C_{\sigma,\rho}[f]$  and  $S_{\sigma,\rho}[f]$  asymptotically in inverse powers of  $\sigma \neq 0$ . The obvious route, letting  $C_{\sigma,\rho}[f] + iS_{\sigma,\rho}[f] = \int_{-1}^1 f(x) e^{i(\sigma x + \rho)} dx$  and using an explicit expansion from (Iserles & Nørsett 2005), is probably less transparent than direct expansion. Integrating  $S_{\sigma,\rho}$ ,  $\sigma \neq 0$ , twice by parts we obtain

$$\begin{aligned} S_{\sigma,\rho}[f] &= -\frac{1}{\sigma} \int_{-1}^1 f(x) \frac{d}{dx} \cos(\sigma x + \rho) dx \\ &= -\frac{1}{\sigma} [f(1) \cos(\sigma + \rho) - f(-1) \cos(\sigma - \rho)] + \frac{1}{\sigma} \int_{-1}^1 f'(x) \cos(\sigma x + \rho) dx \\ &= -\frac{1}{\sigma} [f(1) \cos(\sigma + \rho) - f(-1) \cos(\sigma - \rho)] + \frac{1}{\sigma^2} \int_{-1}^1 f'(x) \frac{d}{dx} \sin(\sigma x + \rho) dx \\ &= -\frac{1}{\sigma} [f(1) \cos(\sigma + \rho) - f(-1) \cos(\sigma - \rho)] \\ &\quad + \frac{1}{\sigma^2} [f'(1) \sin(\sigma + \rho) + f'(-1) \sin(\sigma - \rho)] - \frac{1}{\sigma^2} S_{\sigma,\rho}[f'']. \end{aligned}$$

Iterating this expression yields the asymptotic expansion of  $S_{\sigma,\rho}[f]$  in inverse powers of  $\sigma$ ,

$$S_{\sigma,\rho}[f] \sim - \sum_{m=0}^{\infty} \frac{(-1)^m}{\sigma^{2m+1}} [f^{(2m)}(1) \cos(\sigma + \rho) - f^{(2m)}(-1) \cos(\sigma - \rho)] \quad (2.5)$$

$$+ \sum_{m=0}^{\infty} \frac{(-1)^m}{\sigma^{2m+2}} [f^{(2m+1)}(1) \sin(\sigma + \rho) + f^{(2m+1)}(-1) \sin(\sigma - \rho)], \quad \sigma \gg 1.$$

Likewise, using (2.5), we have

$$\begin{aligned} C_{\sigma,\rho}[f] &= \frac{1}{\sigma} \int_{-1}^1 f(x) \frac{d}{dx} \sin(\sigma x + \rho) dx \\ &= \frac{1}{\sigma} [f(1) \sin(\sigma + \rho) + f(-1) \sin(\sigma - \rho)] - \frac{1}{\sigma} S_{\sigma,\rho}[f'] \\ &\sim \sum_{m=0}^{\infty} \frac{(-1)^m}{\sigma^{2m+1}} [f^{(2m)}(1) \sin(\sigma + \rho) + f^{(2m)}(-1) \sin(\sigma - \rho)] \\ &\quad + \sum_{m=0}^{\infty} \frac{(-1)^m}{\sigma^{2m+2}} [f^{(2m+1)}(1) \cos(\sigma + \rho) - f^{(2m+1)}(-1) \cos(\sigma - \rho)], \quad \sigma \gg 1. \end{aligned} \tag{2.6}$$

Substituting (2.5) and (2.6) into (2.4) results in

$$\begin{aligned} \mathbf{I}[f] &\sim I_0(\tau) \int_{-1}^1 f(x) dx \\ &\quad + 2 \sum_{k=0}^{\infty} (-1)^k I_{2k+1}(\tau) \left\{ - \sum_{m=0}^{\infty} \frac{(-1)^m}{[(2k+1)\omega\alpha]^{2m+1}} [f^{(2m)}(1) \cos((2k+1)\omega(\alpha+\beta)) \right. \\ &\quad \quad \left. - f^{(2m)}(-1) \cos((2k+1)\omega(\alpha-\beta))] \right. \\ &\quad \left. + \sum_{m=0}^{\infty} \frac{(-1)^m}{[(2k+1)\omega\alpha]^{2m+2}} [f^{(2m+1)}(1) \sin((2k+1)\omega(\alpha+\beta)) \right. \\ &\quad \quad \left. + f^{(2m+1)}(-1) \sin((2k+1)\omega(\alpha-\beta))] \right\} \\ &\quad + 2 \sum_{k=1}^{\infty} (-1)^k I_{2k}(\tau) \left\{ \sum_{m=0}^{\infty} \frac{(-1)^m}{(2k\omega\alpha)^{2m+1}} [f^{(2m)}(1) \sin(2k\omega(\alpha+\beta)) \right. \\ &\quad \quad \left. + f^{(2m)}(-1) \sin(2k\omega(\alpha-\beta))] \right. \\ &\quad \left. + \sum_{m=0}^{\infty} \frac{(-1)^m}{(2k\omega\alpha)^{2m+2}} [f^{(2m+1)}(1) \cos(2k\omega(\alpha+\beta)) \right. \\ &\quad \quad \left. - f^{(2m+1)}(-1) \cos(2k\omega(\alpha-\beta))] \right\} \\ &= I_0(\tau) \int_{-1}^1 f(x) dx \\ &\quad + 2 \sum_{m=0}^{\infty} \frac{(-1)^m}{(\alpha\omega)^{2m+1}} \left[ -f^{(2m)}(1) \sum_{k=0}^{\infty} \frac{(-1)^k I_{2k+1}(\tau)}{(2k+1)^{2m+1}} \cos((2k+1)\omega(\alpha+\beta)) \right. \\ &\quad \left. + f^{(2m)}(1) \sum_{k=1}^{\infty} \frac{(-1)^k I_{2k}(\tau)}{(2k)^{2m+1}} \sin(2k\omega(\alpha+\beta)) \right] \end{aligned}$$

$$\begin{aligned}
& + f^{(2m)}(-1) \sum_{k=0}^{\infty} \frac{(-1)^k I_{2k+1}(\tau)}{(2k+1)^{2m+1}} \cos((2k+1)\omega(\alpha-\beta)) \\
& + f^{(2m)}(-1) \sum_{k=1}^{\infty} \frac{(-1)^k I_{2k}(\tau)}{(2k)^{2m+1}} \sin(2k\omega(\alpha-\beta)) \Bigg] \\
& + 2 \sum_{m=0}^{\infty} \frac{(-1)^m}{(\alpha\omega)^{2m+2}} \left[ f^{(2m+1)}(1) \sum_{k=0}^{\infty} \frac{(-1)^k I_{2k+1}(\tau)}{(2k+1)^{2m+2}} \sin((2k+1)\omega(\alpha+\beta)) \right. \\
& + f^{(2m+1)}(1) \sum_{k=1}^{\infty} \frac{(-1)^k I_{2k}(\tau)}{(2k)^{2m+2}} \cos(2k\omega(\alpha+\beta)) \\
& + f^{(2m+1)}(-1) \sum_{k=0}^{\infty} \frac{(-1)^k I_{2k+1}(\tau)}{(2k+1)^{2m+2}} \sin((2k+1)\omega(\alpha-\beta)) \\
& \left. - f^{(2m+1)}(-1) \sum_{k=1}^{\infty} \frac{(-1)^k I_{2k}(\tau)}{(2k)^{2m+2}} \cos(2k\omega(\alpha-\beta)) \right].
\end{aligned}$$

Let

$$\begin{aligned}
\Theta_m^{[1]}(\psi, \tau) &= 2 \sum_{k=0}^{\infty} \frac{(-1)^k I_{2k+1}(\tau)}{(2k+1)^{2m+1}} \cos((2k+1)\psi) \\
\Theta_m^{[2]}(\psi, \tau) &= 2 \sum_{k=1}^{\infty} \frac{(-1)^k I_{2k}(\tau)}{(2k)^{2m+1}} \sin(2k\psi), \\
\Phi_m^{[1]}(\psi, \tau) &= 2 \sum_{k=0}^{\infty} \frac{(-1)^k I_{2k+1}(\tau)}{(2k+1)^{2m+2}} \sin((2k+1)\psi), \\
\Phi_m^{[2]}(\psi, \tau) &= 2 \sum_{k=1}^{\infty} \frac{(-1)^k I_{2k}(\tau)}{(2k)^{2m+2}} \cos(2k\psi).
\end{aligned}$$

Note that the four functions are analytic in  $\psi, \tau$  for all  $m \in \mathbb{Z}_+$  and their convergence is assured. They are periodic in  $\psi$  of period  $2\pi$  for  $\Theta^{[1]}$  and  $\Phi^{[1]}$ , of period  $\pi$  otherwise.

In Fig. 2.2 we display the functions  $\Theta_0^{[i]}$  and  $\Phi_0^{[i]}$  for  $i = 1, 2$ . Note that the differences between  $\Theta_m^{[i]}$  for  $m \geq 1$  and  $\Theta_0^{[i]}$  (likewise, between  $\Phi_m^{[i]}$  and  $\Phi_0^{[i]}$ ) are very small, thus this figure is typical of all  $m$ s.

The four functions are infinite series. Yet, the speed of their convergence is so rapid that it is enough to restrict the range of summation to  $k \leq 6$  to attain machine accuracy.

Using  $\Theta_m^{[i]}$  and  $\Phi_m^{[i]}$  we can write conveniently the asymptotic expansion of the ExpSin integral  $I[f]$ .

**Lemma 1** *Let  $\alpha\omega \gg 1$ . Then*

$$\begin{aligned}
I[f] &\sim I_0(\tau) \int_{-1}^1 f(x) dx \\
&+ \sum_{m=0}^{\infty} \frac{(-1)^m}{(\alpha\omega)^{2m+1}} \left\{ f^{(2m)}(1) [\Theta_m^{[2]}(\omega(\alpha+\beta)) - \Theta_m^{[1]}(\omega(\alpha+\beta))] \right.
\end{aligned} \tag{2.7}$$

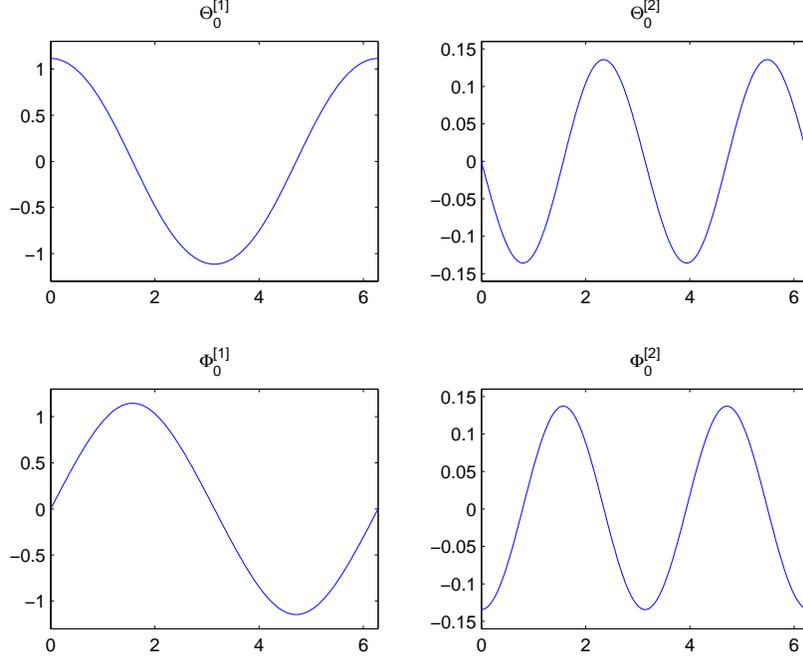


Figure 2.2: The functions  $\Theta_0^{[i]}(\psi, 1)$  and  $\Phi_0^{[i]}(\psi, 1)$  for  $i = 1, 2$  and  $0 \leq \psi \leq 2\pi$ .

$$\begin{aligned}
& + f^{(2m)}(-1)[\Theta_m^{[2]}(\omega(\alpha - \beta)) + \Theta_m^{[1]}(\omega(\alpha - \beta))] \Big\} \\
& + \sum_{m=0}^{\infty} \frac{(-1)^m}{(\alpha\omega)^{2m+2}} \left\{ f^{(2m+1)}(1)[\Phi_m^{[2]}(\omega(\alpha + \beta)) + \Phi_m^{[1]}(\omega(\alpha + \beta))] \right. \\
& \quad \left. - f^{(2m+1)}(-1)[\Phi_m^{[2]}(\omega(\alpha - \beta)) - \Phi_m^{[1]}(\omega(\alpha - \beta))] \right\}.
\end{aligned}$$

An immediate application of the expansion (2.7) is to the numerical calculation of  $I[f]$ . Truncating the series results for  $s \in \mathbb{N}$  in the *asymptotic method*

$$\begin{aligned}
I[f] & \approx \mathbf{A}_s[f] = I_0(\tau) \int_{-1}^1 f(x) dx \\
& + \sum_{m=0}^{\lfloor (s-1)/2 \rfloor} \frac{(-1)^m}{(\alpha\omega)^{2m+1}} \left\{ f^{(2m)}(1)[\Theta_m^{[2]}(\omega(\alpha + \beta)) - \Theta_m^{[1]}(\omega(\alpha + \beta))] \right. \\
& \quad \left. + f^{(2m)}(-1)[\Theta_m^{[2]}(\omega(\alpha - \beta)) + \Theta_m^{[1]}(\omega(\alpha - \beta))] \right\}
\end{aligned} \tag{2.8}$$

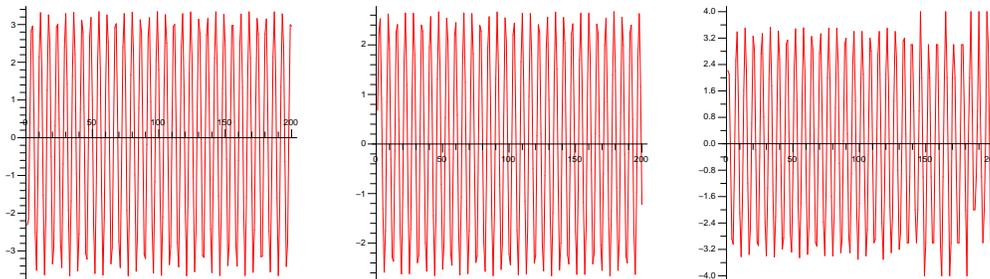


Figure 2.3: Scaled error  $\omega^{s+1}|\mathbf{A}_s[e^x] - \mathbf{I}[e^x]|$  (with  $\alpha = 1, \beta = 0$ ) for  $s = 1, 2, 3$  for  $\omega \in [0, 200]$ .

Table 2: Absolute errors  $|\mathbf{A}_s[e^x] - \mathbf{I}[e^x]|$  for  $s = 1, 2, 3$ .

$s$	$\omega = 10$	$\omega = 50$	$\omega = 100$	$\omega = 200$
1	$2.14_{-02}$	$3.96_{-04}$	$1.81_{-04}$	$7.39_{-05}$
2	$1.92_{-03}$	$2.02_{-05}$	$2.22_{-06}$	$1.53_{-07}$
3	$2.11_{-04}$	$1.44_{-07}$	$1.76_{-08}$	$1.89_{-09}$

$$+ \sum_{m=0}^{\lfloor s/2 \rfloor - 1} \frac{(-1)^m}{(\alpha\omega)^{2m+2}} \left\{ f^{(2m+1)}(1) [\Phi_m^{[2]}(\omega(\alpha + \beta)) + \Phi_m^{[1]}(\omega(\alpha + \beta))] \right. \\ \left. - f^{(2m+1)}(-1) [\Phi_m^{[2]}(\omega(\alpha - \beta)) - \Phi_m^{[1]}(\omega(\alpha - \beta))] \right\}.$$

and it is trivial to verify that

$$\mathbf{A}_s[f] = \mathbf{I}[f] + \mathcal{O}((\alpha\omega)^{-s-1}), \quad |\alpha\omega| \gg 1.$$

In a way of an example, we have used (2.8) to compute the integral  $\mathbf{I}[e^x]$  from Fig. 2.1. The results are displayed in Fig. 2.3 and they confirm the theoretical expectations on asymptotic behaviour. Indeed, they exceed it, because apparently the asymptotic regime sets already for very small  $|\alpha\omega|$ , rather than only for large frequencies. However, the figure does not exhibit transparently the actual absolute error and for this we refer to Table 2. It is clear that, while for large  $|\alpha\omega|$  we need relatively modest values of  $s$ , moderate frequencies call for large  $s$  and the method becomes expensive: this is only to be expected, because of the asymptotic nature of the method (2.8).

There are two obvious problems associated with the asymptotic method (2.8). Firstly, we said nothing about the numerical evaluation of the leading integral. This, however, can be accomplished easily using classical quadrature, since it is non-oscillatory. Secondly, as demonstrated in Table 2, the formula is useful only for sufficiently large value of  $|\alpha\omega|$ . Although, in our experience, asymptotic behaviour sets surprisingly rapidly, this is an undoubted shortcoming. In principle, not all the  $\omega_{i,j}$ s in (1.1) need be large and we do not wish to employ different quadrature rules for different  $\omega_{i,j}$ s, something that unduly complicates things.

Fortunately, a major lesson of recent advances in numerical quadrature of highly oscillatory integrals is that the main role of asymptotic formulæ like (2.7) is as a gateway to Filon-type (Iserles & Nørsett 2005) and Levin-type (Olver 2006) techniques. An important advantage of these methods is that they segue seamlessly into classical quadrature for small  $\omega$ , hence are uniformly effective throughout the entire range of frequencies. They are also typically significantly more accurate than the asymptotic method (2.8).

### 3 A Filon-type method

An alternative to the asymptotic method (2.8) is a *Filon-type method*. Thus, let  $\nu \geq 2$ , nodes  $-1 = c_1 < c_2 < \dots < c_\nu = 1$  and multiplicities  $m_1, m_2, \dots, m_\nu \in \mathbb{N}$ . We interpolate the function  $f$  in a Hermite sense at the nodes  $\mathbf{c}$  by a polynomial  $p$  of degree  $\sum_{i=1}^\nu m_i - 1$ ,

$$p^{(j)}(c_k) = f^{(j)}(c_k), \quad j = 0, \dots, m_k - 1, \quad k = 1, 2, \dots, \nu. \quad (3.1)$$

The *Filon-type method* for the highly oscillatory integral (2.1) is defined as

$$\mathbf{F}[f] = \int_{-1}^1 p(x) e^{\tau \sin \omega(\alpha x + \beta)} dx. \quad (3.2)$$

**Theorem 2** Let  $s = \min\{m_1, m_\nu\}$ . Then for every  $f \in C^\infty[-1, 1]$

$$\mathbf{F}[f] - \mathbf{I}[f] \sim I_0(\tau) \mathbf{E}[f] + \mathcal{O}(\omega^{-s-1}), \quad \omega \gg 1, \quad (3.3)$$

where  $\mathbf{E}[f] = \int_{-1}^1 [p(x) - f(x)] dx$ .

*Proof* We use the method of proof from (Iserles & Nørsett 2005). Since both  $\mathbf{F}$  and  $\mathbf{I}$  are linear operators,  $\mathbf{F}[f] - \mathbf{I}[f] = \mathbf{I}[p - f]$  and the theorem follows at once from letting  $p - f$  in (2.8) and noting that the interpolation conditions (3.1) annihilate asymptotic terms for  $m = 0, 1, \dots, \lfloor (s-1)/2 \rfloor$  in the first sum in (2.7) and  $m = 0, 1, \dots, \lfloor s/2 \rfloor - 1$  in the second.  $\square$

Note that the internal nodes  $c_2, \dots, c_{\nu-1}$  have no influence upon the asymptotic order of the error. However, they have three important functions. Firstly, good choice of such points minimizes the non-oscillatory quadrature error  $\mathbf{E}[f]$ , one of the two components of the quadrature error in (3.3). Secondly, intuitively speaking, the method (3.2) is nothing but the asymptotic quadrature  $\mathbf{A}_s$ , applied to the interpolation error  $p - f$  rather than to the original function  $f$ . Thus, the smaller we make the interpolation error, the better. Thirdly, unlike (2.8), the Filon-type method is relevant throughout the range of frequencies  $\omega \in \mathbb{R}$ . In particular, when  $|\omega|$  is small then  $\mathbf{F}[f] = \mathbf{E}[f] + \mathcal{O}(\omega)$ , the reason being that  $\mathbf{I}[f] = \int_{-1}^1 f(x) dx + \mathcal{O}(\omega)$  and  $\mathbf{F}[f] = \mathbf{I}[p] = \int_{-1}^1 p(x) dx + \mathcal{O}(\omega) = \mathbf{Q}[f] + \mathcal{O}(\omega)$ . Thus, rendering  $|\mathbf{E}[f]|$  small is vital also in this regime.

#### 3.1 Implementation of the Filon-type method

The implementation of (3.2) is based on the premise that we can integrate (2.1) exactly once  $f$  is a polynomial. Thus, let

$$p(x) = \sum_{r=0}^q p_r x^r, \quad \text{where} \quad q = \sum_{i=1}^\nu m_i - 1.$$

Then

$$\mathbf{F}[f] = \sum_{r=0}^q p_r \int_{-1}^1 x^r e^{\tau \sin \omega(\alpha x + \beta)} dx = \sum_{r=0}^q p_r \mu_r(\omega). \quad (3.4)$$

The *moments*  $\mu_r$  can be calculated directly from the asymptotic expansion (2.7) since the latter terminates in that case,

$$\begin{aligned} \mu_r(\omega) &= \frac{1 + (-1)^r}{r + 1} I_0(\tau) \\ &+ \sum_{m=0}^{\lfloor r/2 \rfloor} \frac{(-1)^m}{(\alpha\omega)^{2m+1}} \frac{r!}{(r - 2m)!} \left\{ [\Theta_m^{[2]}(\omega(\alpha + \beta)) - \Theta_m^{[1]}(\omega(\alpha + \beta))] \right. \\ &\quad \left. + (-1)^r [\Theta_m^{[2]}(\omega(\alpha - \beta)) + \Theta_m^{[1]}(\omega(\alpha - \beta))] \right\} \\ &+ \sum_{m=0}^{\lfloor (r-1)/2 \rfloor} \frac{(-1)^m}{(\alpha\omega)^{2m+2}} \frac{r!}{(r - 2m - 1)!} \left\{ [\Phi_m^{[2]}(\omega(\alpha + \beta)) + \Phi_m^{[1]}(\omega(\alpha + \beta))] \right. \\ &\quad \left. + (-1)^r [\Phi_m^{[2]}(\omega(\alpha - \beta)) - \Phi_m^{[1]}(\omega(\alpha - \beta))] \right\}, \quad r \in \mathbb{Z}_+. \end{aligned}$$

Note that (3.4) is not a practical means to calculate  $\mathbf{F}[f]$ . Like in the case of non-oscillatory quadrature, it is advantageous to express  $p$  in terms of *cardinal polynomials*,

$$p(x) = \sum_{k=1}^{\nu} \sum_{j=0}^{m_k-1} \ell_{k,j}(x) f^{(j)}(c_k),$$

where each  $\ell_{k,j}$  is a polynomial of degree  $q$  such that

$$\ell_{k,j}^{(i)}(c_n) = \begin{cases} 1, & k = n, i = j, \\ 0, & \text{otherwise} \end{cases} \quad (3.5)$$

for  $i = 0, 1, \dots, m_n - 1, j = 0, 1, \dots, m_k - 1, k, n = 1, 2, \dots, \nu$ . Letting

$$b_{k,j} = \mathbf{I}[\ell_{k,j}] = \int_{-1}^1 \ell_{k,j}(x) e^{\tau \sin(\omega(\alpha x + \beta))} dx, \quad j = 0, 1, \dots, m_k - 1, \quad k = 1, 2, \dots, \nu$$

(which we can do once-for-all in terms of the moments  $\mu_r$ ) we obtain

$$\mathbf{F}[f] = \sum_{k=1}^{\nu} \sum_{j=0}^{m_k-1} b_{k,j} f^{(j)}(c_k), \quad (3.6)$$

a form reminiscent of classical quadrature (Davis & Rabinowitz 1984).

To illustrate our construction by few simple examples, let us assume (mostly to render notation more transparent) that  $\alpha = 1$  and  $\beta = 0$ , whereby

$$\begin{aligned} \mu_0(\omega) &= 2I_0(\tau) + \frac{2}{\omega} \Theta_0^{[2]}(\omega), \\ \mu_1(\omega) &= -\frac{2}{\omega} \Theta_0^{[1]}(\omega) + \frac{2}{\omega^2} \Phi_0^{[1]}(\omega), \end{aligned}$$

$$\begin{aligned}\mu_2(\omega) &= \frac{2}{3}I_0(\tau) + \frac{2}{\omega}\Theta_0^{[2]}(\omega) + \frac{4}{\omega^2}\Phi_0^{[2]}(\omega) - \frac{4}{\omega^3}\Theta_1^{[2]}(\omega), \\ \mu_3(\omega) &= -\frac{2}{\omega}\Theta_0^{[1]}(\omega) + \frac{6}{\omega^2}\Phi_0^{[1]}(\omega) + \frac{12}{\omega^3}\Theta_1^{[1]}(\omega) - \frac{12}{\omega^4}\Phi_1^{[1]}(\omega), \\ \mu_4(\omega) &= \frac{2}{5}I_0(\tau) + \frac{2}{\omega}\Theta_0^{[2]}(\omega) + \frac{8}{\omega^2}\Phi_0^{[2]}(\omega) - \frac{24}{\omega^3}\Theta_1^{[2]}(\omega) - \frac{48}{\omega^4}\Phi_1^{[2]}(\omega) + \frac{48}{\omega^5}\Theta_2^{[2]}(\omega).\end{aligned}$$

(Note that we have suppressed the dependence of  $\Phi_m^{[i]}$  and  $\Theta_m^{[i]}$  on  $\tau$ .)

We commence from  $\nu = 2$ ,  $\mathbf{c} = [-1, 1]$  and  $\mathbf{m} = [1, 1]$ , whereby  $p(x) = \frac{1}{2}(1-x)f(-1) + \frac{1}{2}(1+x)f(1)$ . This results in the method

$$\mathbf{F}[f] = \frac{1}{2}[\mu_0(\omega) - \mu_1(\omega)]f(-1) + \frac{1}{2}[\mu_0(\omega) + \mu_1(\omega)]f(1).$$

Next, we consider  $\mathbf{c} = [-1, 1]$  and  $\mathbf{m} = [2, 2]$ , whereby

$$\begin{aligned}p(x) &= \frac{1}{4}(1+x)(2+x-x^2)f(1) + \frac{1}{4}(1-x)(2-x-x^2)f(-1) \\ &\quad - \frac{1}{4}(1-x)(1+x)^2f'(1) + \frac{1}{4}(1-x)^2(1+x)f'(-1)\end{aligned}$$

and

$$\begin{aligned}\mathbf{F}[f] &= \frac{1}{4}(2\mu_0 + 3\mu_1 - \mu_3)f(1) + \frac{1}{4}(2\mu_0 - 3\mu_1 + \mu_3)f(-1) \\ &\quad - \frac{1}{4}(\mu_0 + \mu_1 - \mu_2 - \mu_3)f'(1) + \frac{1}{4}(\mu_0 - \mu_1 - \mu_2 + \mu_3)f'(-1).\end{aligned}$$

As our final example, we let  $\nu = 3$ ,  $\mathbf{c} = [-1, 0, 1]$  and  $\mathbf{m} = [2, 1, 2]$ . We now have

$$\begin{aligned}p(x) &= \frac{1}{4}x(1+x)^2(3-2x)f(1) + (1-x^2)^2f(0) - \frac{1}{4}x(1-x)^2(3+2x)f(-1) \\ &\quad - x(1-x)(1+x)^2f'(1) - x(1-x)^2(1+x)f'(-1),\end{aligned}$$

therefore

$$\begin{aligned}\mathbf{F}[f] &= \frac{1}{4}(3\mu_1 + 4\mu_2 - \mu_3 - 2\mu_4)f(1) + (\mu_0 - 2\mu_2 + \mu_4)f(0) \\ &\quad + \frac{1}{4}(-3\mu_1 + 4\mu_2 + \mu_3 - 2\mu_4)f(-1) + \frac{1}{4}(-\mu_1 - \mu_2 + \mu_3 + \mu_4)f'(1) \\ &\quad + \frac{1}{4}(-\mu_1 + \mu_2 + \mu_3 - \mu_4)f'(-1).\end{aligned}$$

### 3.2 Hermite–Birkhoff quadrature

Wishing to minimise the non-oscillatory error  $\mathbf{E}[f]$ , we have the freedom of choosing nodes and weights, subject to  $c_1 = -1$ ,  $c_\nu = 1$  and  $s = \min\{m_1, m_\nu\}$ , a procedure that has been already considered in (Iserles & Nørsett 2006). Let

$$\tilde{b}_{k,j} = \int_{-1}^1 \ell_{k,j}(x) dx, \quad j = 0, 1, \dots, m_k - 1, \quad k = 1, 2, \dots, \nu,$$

where the  $\ell_{k,j}$ s were defined in (3.5), and

$$\mathbf{Q}[f] = \sum_{k=1}^{\nu} \sum_{j=0}^{m_k-1} \tilde{b}_{k,j} f^{(j)}(c_k), \quad (3.7)$$

we have  $\mathbf{E}[f] = \mathbf{Q}[f] - \int_{-1}^1 f(x) dx$ . Therefore  $\mathbf{Q}$  is a *Hermite–Birkhoff quadrature* (Micchelli & Rivlin 1973) for the computation of a non-oscillatory integral.

**Theorem 3** Let  $m_1 = m_\nu = s$  and  $m_2 = m_3 = \dots = m_{\nu-1} \equiv 1$ . The quadrature (3.7) is of maximal order  $2\nu + 2s - 4$  (in other words, is exact for all polynomials of degree  $2\nu + 2s - 5$  when  $c_2, c_3, \dots, c_{\nu-1}$  are the zeros of the Jacobi polynomial  $P_{\nu-2}^{(s,s)}$ ).

*Proof* A straightforward generalisation of the familiar proof on the order of Gauss–Christoffel quadrature (Davis & Rabinowitz 1984). Let

$$\mathbb{P}_{\nu+2s-2}[x] \ni u(x) = (1-x^2)^s P_{\nu-2}^{(s,s)}(x),$$

where  $\mathbb{P}_n[x]$  is the set of  $n$ th-degree polynomials. Given any  $w \in \mathbb{P}_{2\nu+2s-5}[x]$ , it follows by the Euclidean algorithm that there exist  $p \in \mathbb{P}_{\nu-3}[x]$  and  $q \in \mathbb{P}_{\nu+2s-3}[x]$  such that  $w = pu + q$ . Recalling that  $P_{\nu-2}^{(s,s)}$  is orthogonal in  $(-1, 1)$  with respect to the weight function  $(1-x^2)^s$  (Abramowitz & Stegun 1964), we have

$$\int_{-1}^1 p(x)u(x) dx = \int_{-1}^1 p(x)P_{\nu-2}^{(s,s)}(x)(1-x^2)^s dx = 0,$$

while  $\mathbf{Q}[pu] = 0$  because  $u(c_k) = 0$ ,  $k = 1, 2, \dots, \nu$  and  $u^{(j)}(\pm 1) = 0$ ,  $j = 0, 1, \dots, s-1$ . Therefore

$$\mathbf{E}[w] = \mathbf{Q}[w] - \int_{-1}^1 w(x) dx = \mathbf{Q}[q] - \int_{-1}^1 q(x) dx.$$

The right-hand side vanishes because the weights are interpolatory. This is standard argument in classical quadrature and follows in a Hermite–Birkhoff setting by counting  $\nu + 2s$  degrees of freedom and observing that the underlying linear system (ensuring that  $\mathbf{Q}$  is at least of order  $\nu + 2s$ ) is nonsingular, being a limiting case of Lagrangian interpolation with  $\nu + 2s$  nodes. We deduce that  $\mathbf{E}[w] = 0$  for every  $w \in \mathbb{P}_{2\nu+2s-5}[x]$ , hence order  $2\nu + 2s - 4$ .

It remains to prove that no other choice of internal nodes  $c_2, c_3, \dots, c_{\nu-1}$  may increase the order. To this end it is sufficient to single out one polynomial  $v \in \mathbb{P}_{2\nu+2s-4}[x]$  such that  $\mathbf{E}[v] \neq 0$  for any choice of internal nodes. We thus choose

$$v(x) = (1-x^2)^s \prod_{k=2}^{\nu-1} (x-c_k)^2,$$

where  $c_2, c_3, \dots, c_{\nu-1} \in (-1, 1)$  are arbitrary. Trivially,  $\int_{-1}^1 v(x) dx > 0$ . On the other hand,  $v(c_k) = 0$ ,  $k = 1, 2, \dots, \nu$  and  $v^{(j)}(\pm 1) = 0$ ,  $j = 0, 1, \dots, s-1$ , imply that  $\mathbf{Q}[v] = 0$ . It thus follows that  $\mathbf{E}[v] < 0$  and the maximal order is indeed  $2(\nu + s - 2)$ .  $\square$

To flesh out numbers, herewith few explicit quadratures (3.7):

$$\begin{aligned} \nu = 2, s = 2 : \quad \mathbf{Q}[f] &= f(1) + f(-1) - \frac{1}{3}[f'(1) - f'(-1)], \\ \nu = 2, s = 3 : \quad \mathbf{Q}[f] &= f(1) + f(-1) - \frac{2}{3}[f'(1) - f'(-1)] + \frac{1}{15}[f''(1) + f''(-1)], \\ \nu = 3, s = 2 : \quad \mathbf{Q}[f] &= \frac{1}{15}[7f(1) + 16f(0) + 7f(-1)] - \frac{1}{15}[f'(1) - f'(-1)], \\ \nu = 4, s = 2 : \quad \mathbf{Q}[f] &= \frac{1}{135}[37f(1) + 98f(\frac{\sqrt{7}}{7}) + 98f(-\frac{\sqrt{7}}{7}) + 37f(-1)] \end{aligned}$$

Table 3: Absolute errors in approximating  $\int_{-1}^1 e^x dx$  by  $\mathbf{Q}[e^x]$  with  $s = m_1 = m_\nu = 2$  (top),  $s = m_1 = m_\nu = 3$  (bottom),  $m_2 = m_3 = \dots = m_{\nu-1} = 1$  and  $\nu = 2, \dots, 7$ .

$\nu$	2	3	4	5	6	7
$s = 2$	$4.77_{-02}$	$2.21_{-04}$	$7.42_{-07}$	$1.74_{-09}$	$2.93_{-12}$	$3.71_{-15}$
$s = 3$	$1.34_{-03}$	$2.61_{-06}$	$4.65_{-09}$	$6.61_{-12}$	$7.43_{-15}$	$6.77_{-18}$

$$\begin{aligned}
& -\frac{1}{45}[f'(1) - f'(-1)], \\
\nu = 3, s = 3 : \quad \mathbf{Q}[f] &= \frac{1}{35}[19f(1) + 32f(0) + 19f(-1)] - \frac{4}{35}[f'(1) - f'(-1)] \\
& + \frac{1}{105}[f''(1) + f''(-1)], \\
\nu = 4, s = 3 : \quad \mathbf{Q}[f] &= \frac{1}{1120}[391f(1) + 729f(\frac{1}{3}) + 729f(-\frac{1}{3}) + 391f(-1)] \\
& - \frac{13}{280}[f'(1) - f'(-1)] + \frac{1}{420}[f''(1) + f''(-1)].
\end{aligned}$$

The order, in each case, is  $2(\nu + s - 2)$ .

Of course, there is nothing to prevent us from using higher multiplicities with internal nodes, except that we might lose the attractive feature of Theorem 3, reminiscent of the Gauss–Christoffel quadrature, namely that maximal order exceeds by  $\nu$  the number of degrees of freedom. Thus, for example, choosing  $\nu = 3$ ,  $\mathbf{c} = [-1, 0, 1]$  and  $m_k \equiv 2$ , the coefficient of  $f'(0)$  is nil and we recover the sixth-order formula with  $\nu = 3$ ,  $s = 1$ , above. On the other hand, once we let  $\mathbf{m} = [2, 3, 2]$ , we obtain

$$\mathbf{Q}[f] = \frac{1}{35}[11f(1) + 48f(0) + 11f(-1)] - \frac{1}{35}[f'(1) - f'(-1)] + \frac{8}{105}f''(0),$$

of order 8.

In Table 3 we display errors  $\mathbf{E}[f]$  committed by Hermite–Birkhoff methods consistent with the conditions of Theorem 3, with  $s = 2$  and increasing values of  $\nu$ , applied to the function  $f(x) = e^x$ . The decrease in error is consistent with Theorem 3.

### 3.3 Numerical examples for Filon-type methods

According to (3.3), the error of Filon-type methods has two components. The asymptotic component decays with increasing  $\omega$  but  $I_0(\tau)\mathbf{E}[f]$  is independent of  $\omega$ . Thus, unlike in the case of Filon-type methods for ‘classical’ highly oscillatory integrals (Iserles & Nørsett 2005) and in variance with the asymptotic method  $\mathbf{A}[f]$ , the error does not tend to zero for  $\omega \rightarrow \infty$ . This is demonstrated in Fig. 3.4, where we display the absolute error  $\mathbf{F}[f] - \mathbf{I}[f]$  for the Filon-type method with  $\mathbf{c} = [-1, 0, 1]$ ,  $\mathbf{m} = [2, 1, 2]$  and the function  $f(x) = e^x$ . For  $\omega \gg 1$  the error asymptotes to  $\approx -2.79_{04} = I_0(1)\mathbf{E}[e^x]$  (cf. Table 3,  $\nu = 3$ , for  $\mathbf{E}[e^x]$ ).

In Fig. 3.5 we display the asymptotic error component  $\mathbf{F}[f] - \mathbf{I}[f] - I_0(\tau)\mathbf{E}[f]$ , scaled by  $\omega^3$ , for two different Filon-type methods of an asymptotic order two. In both cases, consistently with Theorem 2, the scaled error asymptotes to a constant.

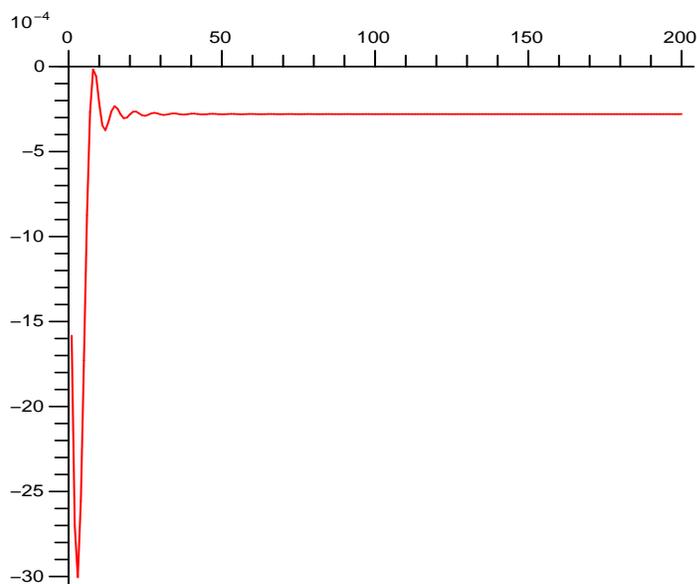


Figure 3.4: The error  $F[e^x] - I[e^x]$  for  $c = [-1, 0, 1]$ ,  $m = [2, 1, 2]$  and  $\omega \in [0, 200]$ .

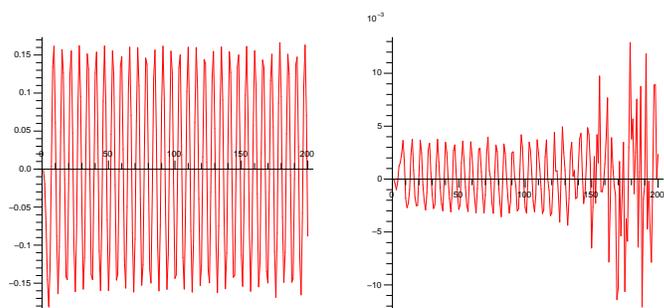


Figure 3.5: Scaled errors  $\omega^3 |F[e^x] - I[e^x] - I_0(1)E[e^x]|$  for  $c = [-1, 0, 1]$ ,  $m = [2, 1, 2]$  (left),  $c = [-1, -\frac{\sqrt{7}}{7}, \frac{\sqrt{7}}{7}, 1]$ ,  $m = [2, 1, 1, 2]$  (right) and  $\omega \in [0, 200]$ .

It is instructive to compare absolute errors at different values of  $\omega$  for asymptotic and Filon-type methods. In an important aspect, this comparison is heavily weighed against Filon-type methods, because the asymptotic method (2.8) assumes that  $\int_{-1}^1 f(x) dx$  is calculated exactly: in practice we need to replace the integral by quadrature. Nonetheless, and even bearing in mind that  $E[f]$  makes up the major share of error in (3.2), Filon-type methods are

Table 4: Absolute errors  $|\mathbf{F}[e^x] - \mathbf{I}[e^x]|$  for different Filon-type methods.

Methods	$\omega = 10$	$\omega = 50$	$\omega = 100$	$\omega = 200$
$\mathbf{c} = [-1, 0, 1], \mathbf{m} = [2, 1, 2]$	$2.18_{-04}$	$2.80_{-04}$	$2.79_{-04}$	$2.79_{-04}$
$\mathbf{c} = [-1, -\frac{\sqrt{7}}{7}, \frac{\sqrt{7}}{7}, 1], \mathbf{m} = [2, 1, 1, 2]$	$2.75_{-06}$	$9.63_{-07}$	$9.43_{-07}$	$9.40_{-07}$
$\mathbf{c} = [-1, 0, 1], \mathbf{m} = [3, 1, 3]$	$9.22_{-07}$	$3.31_{-06}$	$3.31_{-06}$	$3.31_{-06}$
$\mathbf{c} = [-1, -\frac{1}{3}, \frac{1}{3}, 1], \mathbf{m} = [3, 1, 1, 3]$	$7.97_{-09}$	$5.88_{-09}$	$5.88_{-09}$	$5.88_{-09}$
$\mathbf{c} = [-1, -\frac{\sqrt{33}}{11}, 0, \frac{\sqrt{33}}{11}, 1], \mathbf{m} = [3, 1, 1, 1, 3]$	$9.83_{-09}$	$1.40_{-11}$	$7.66_{-12}$	$8.28_{-12}$
$\mathbf{c} = [-1, -\frac{\sqrt{65}}{11}, 0, \frac{\sqrt{65}}{11}, 1], \mathbf{m} = [3, 1, 3, 1, 3]$	$1.18_{-10}$	$1.09_{-13}$	$9.16_{-15}$	$1.21_{-14}$

evident superior.

The error for asymptotic methods is displayed in Table 2 and, predictably, it starts unacceptably high but becomes increasingly small, tending to zero for  $\omega \gg 1$ . Not so for Filon-type methods, exhibited in Table 4. The *uniform error* of (3.2) is considerably smaller, because the performance for small and moderate  $\omega$ s is considerably better. On the other hand, the error for  $\omega \gg 1$  does not tend to zero, as we have already repeatedly observed. Overall, it is clear that Filon-type methods significantly decrease the error at the cost of few extra function evaluations, even when the integral in  $\mathbf{A}_s[f]$  is computed exactly.

We note in passing that the fixed error component  $\mathbf{E}[f]$  assumes significantly smaller importance in the setting of ordinary differential equations and the solution of the integral (1.2). In that case  $\mathbf{E}[f]$  is scaled by  $h^q$ , where  $h = t_{n+1} - t_n$  is the length of the integration interval and  $q$  is the order of the Hermite–Birkhoff quadrature. In this setting Filon-type methods are likely to outperform asymptotic methods by a large margin, since the latter are largely insensitive to the length of integration integral.

## 4 Numerical examples

We bring the equation (1.2) into a form appropriate for the application of Filon-type methods in the interval  $[-1, 1]$ ,

$$\mathbf{y}(t_{n+1}) = e^{hA}\mathbf{y}(t_n) + \frac{h}{2} \int_{-1}^1 e^{\frac{1}{2}h(1-x)A} \mathbf{E}(h(n + \frac{1}{2} + \frac{1}{2}x)) \mathbf{g}(h(n + \frac{1}{2} + \frac{1}{2}x)) dx. \quad (4.1)$$

Our time-stepping routine is obtained by replacing integrals with appropriate Filon-type methods.

In the specific context of equation (1.4), the time-stepping formula (4.1), combined with a Filon-type solver, becomes

$$\begin{aligned} y_{n+1,1} &= y_{n,1} \cos h + y_{n,2} \sin h + h\mathbf{F}[\sin(\frac{1}{2}h(1-x))], \\ y_{n+1,2} &= -y_{n,1} \sin h + y_{n,2} \cos h + h\mathbf{F}[\cos(\frac{1}{2}h(1-x))], \end{aligned}$$

where the Filon-type methods are applied with  $\alpha = \frac{1}{2}, \beta = n + \frac{1}{2}$  and  $\omega$  replaced by  $h\omega$ .

The pointwise error for two Filon-type methods is displayed in Fig. 4.1 for step-sizes  $h \in \{\frac{\pi}{100}, \frac{\pi}{200}, \frac{\pi}{400}, \frac{\pi}{800}\}$ . (We have used there the more precise Filon-type method with step-size  $h = \frac{\pi}{1600}$  as our ‘true’ solution.) The first Filon-type method uses only function values

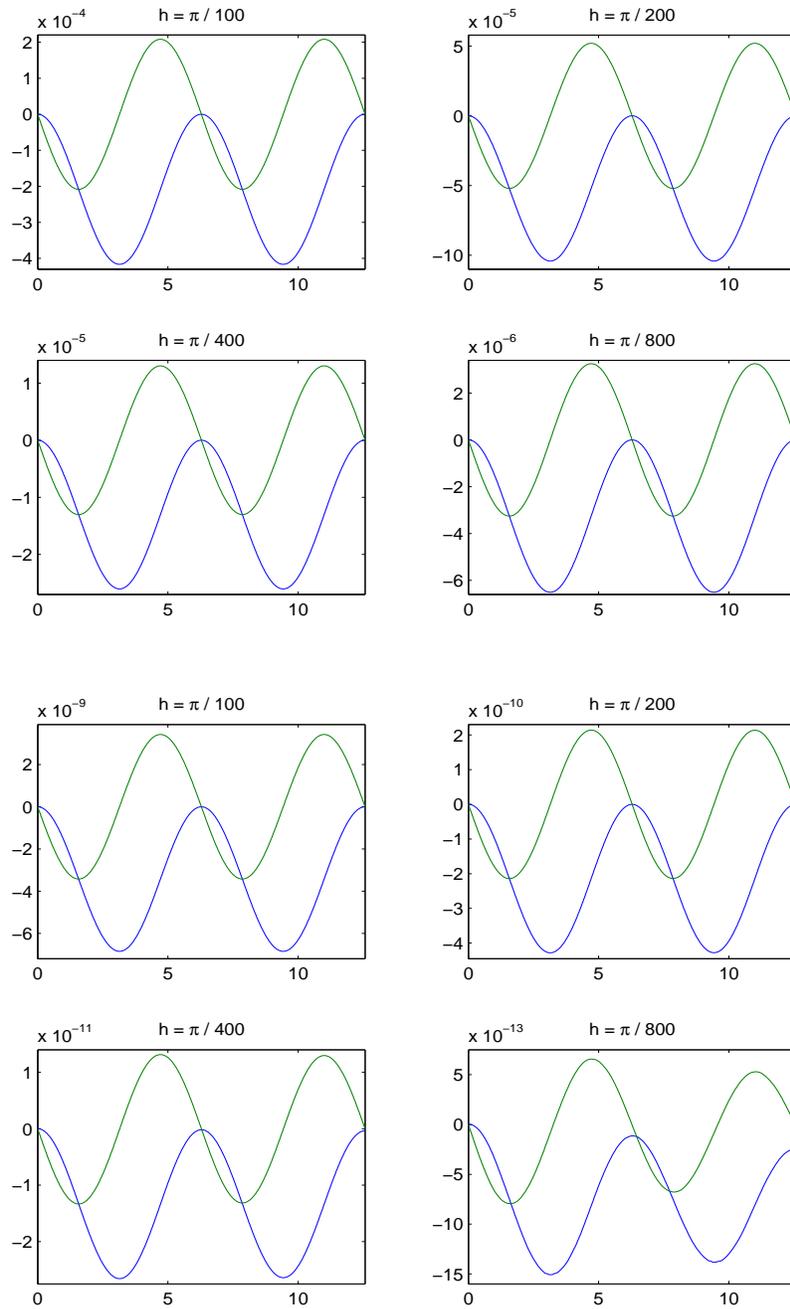


Figure 4.1: Pointwise absolute error for two Filon-type methods.

at the endpoints, the second uses both function values and derivatives there. (We did not use any internal points but note in passing that their incorporation would have further reduced the error.) A comparison with Table 1 is striking: at the cost of just 400 steps with the plain-vanilla Filon-type method (requiring just one new function evaluation per step!) we produce better accuracy than `ode45` with 240645 steps.

Note that the errors in Fig. 4.1 appear to be the same periodic function, scaled by suitable powers of  $h$  (except for the second method with  $h = \frac{\pi}{800}$ , but this is likely to be a machine-precision artefact). So is the exact solution (cf. bottom of Fig. 1.1) but these two functions are different. The reason for periodicity can tell us something about the properties of our method, hence it bears some elaboration. Recall that for large  $\omega$  the major source for the error is the classical quadrature error (scaled by a suitable Bessel function). Let  $\mathbf{e}_n = \mathbf{y}_n - \mathbf{y}(t_n)$  and consider just the error originating in classical quadrature. It readily follows that

$$\mathbf{e}_{n+1} = e^{hA} \mathbf{e}_n + h\mathbf{q}, \quad (4.2)$$

where the vector  $\mathbf{q}$  contains the contribution of classical quadrature error for the different components. For example, in the present case, for ‘plain-vanilla’ Filon we have

$$\mathbf{q} = I_0(1) \begin{bmatrix} h \sin h - 2 + 2 \cos h \\ h + h \cos h - 2 \sin h \end{bmatrix} = \begin{bmatrix} \mathcal{O}(h^4) \\ \mathcal{O}(h^3) \end{bmatrix}.$$

Bearing in mind that  $\mathbf{e}_0 = \mathbf{0}$ , the solution of (4.2) is

$$\mathbf{e}_n = h(I - e^{hA})^{-1}(I - e^{nhA})\mathbf{q},$$

which in the present case becomes

$$\mathbf{e}_n = \frac{h}{2(1 - \cos h)} \begin{bmatrix} 1 - \cos h & \sin h \\ -\sin h & 1 - \cos h \end{bmatrix} \begin{bmatrix} 1 - \cos nh & -\sin nh \\ \sin nh & 1 - \cos nh \end{bmatrix} \mathbf{q}.$$

Periodicity is clear, as is the fact that in Fig. 4.1 the error is a linear combination of just two harmonics,  $\sin nh$  and  $\cos nh$ .

Be it as may, it is crystal clear that even the most elementary Filon-type methods enjoy tremendous advantage in comparison to state-of-the-art general ODE software like `ode45` when applied with high frequencies. Another important advantage of Filon-type methods, which is not apparent from our comparison, is that both the error and the computational effort are roughly uniform in frequency, while classical ODE solvers deteriorate with increasing frequency. Note that we have implemented Filon-type methods in the most straightforward manner, with constant step size and without any error control. (cf. (Iserles & Nørsett 2004) for error control for Filon-type methods.) It is highly likely that more sophisticated implementation would have resulted in even more striking outcome.

## Acknowledgements

The work of Marissa Condon was supported by Science Foundation Ireland under Principal Investigator Grant No. 05/IN.1/I18, while Alfredo Deaño acknowledges financial support from the program of postdoctoral grants of the Spanish Ministry of Education and Science.

## References

- Abramowitz, M. & Stegun, I. A., eds (1964), *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, DC.
- Dautbegovic, E., Condon, M. & Brennan, C. (2005), ‘An efficient nonlinear circuit simulation technique’, *IEEE Trans. Microwave Theory & Techniques* **53**, 548–555.
- Davis, P. J. & Rabinowitz, P. (1984), *Methods of Numerical Integration*, 2nd edn, Academic Press, Orlando, FL.
- Grimm, V. & Hochbruck, M. (2006), ‘Error analysis of exponential integrators for oscillatory second-order differential equations’, *J. Phys. A: Math. Gen.* **39**, 5495–5507.
- Huybrechs, D. & Vandewalle, S. (2006), ‘On the evaluation of highly oscillatory integrals by analytic continuation’, *SIAM J. Num. Anal.* **44**, 1026–1048.
- Iserles, A. (2002), ‘Think globally, act locally: solving highly-oscillatory ordinary differential equations’, *Appld Num. Anal.* **43**, 145–160.
- Iserles, A. & Nørsett, S. P. (2004), ‘On quadrature methods for highly oscillatory integrals and their implementation’, *BIT* **44**, 755–772.
- Iserles, A. & Nørsett, S. P. (2005), ‘Efficient quadrature of highly oscillatory integrals using derivatives’, *Proc. Royal Soc. A* **461**, 1383–1399.
- Iserles, A. & Nørsett, S. P. (2006), From high oscillation to rapid approximation I: Modified Fourier expansions, Technical Report NA2006/05, DAMTP, University of Cambridge. To appear in *IMA J. Num. Anal.*.
- Jeruchim, M. C., Balaban, P. & Shanmugan, K. S. (2000), *Simulation of Communication Systems, Modeling, Methodology and Techniques*, 2nd edn, Kluwer Academic/Plenum Publishers, New York.
- Khanamirian, M. (2008), Quadrature methods for systems of highly oscillatory ODEs. part I, Technical Report NA2008/01, DAMTP, University of Cambridge.
- Micchelli, C. A. & Rivlin, T. J. (1973), ‘Quadrature formulæ and Hermite–Birkhoff interpolation’, *Adv. in Maths* **11**, 93–112.
- Olver, S. (2006), ‘Moment-free numerical integration of highly oscillatory functions’, *IMA J. Num. Anal.* **26**, 213–227.
- Pulch, R. (2005), ‘Multi-time scale differential equations for simulating frequency modulated signals’, *Appld Numer. Maths* **53**, 421–436.
- Roychowdhury, J. (2001), ‘Analysing circuits with widely separated time scales using numerical PDE methods’, *IEEE Trans. Circuits Sys. I, Fund. Theory & Applics* **48**, 578–594.
- Wong, R. (2001), *Asymptotic Approximations of Integrals*, SIAM, Philadelphia.