Internat. Math. Nachrichten
Nr. ?? (2013), 1–17

Long-term control of oscillations in differential equations

Ernst Hairer and Christian Lubich

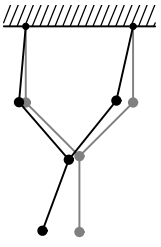
Université de Genève and Universität Tübingen

Ordinary differential equations arise everywhere in science – Newton's law in physics, N-body problems in astronomy and in molecular dynamics, engineering problems in robotics, population models in biology, and many more. Since their analytic solution can be obtained only in exceptional situations, one is usually restricted to numerical simulations and/or to qualitative investigations of the flow. This article reviews a recent technique - the modulated Fourier expansion - which permits to get insight into the long-term behaviour of numerical solutions of multi-value methods as well as of analytic solutions of highly oscillatory differential equations.

1 Two problems with high-frequency oscillations

We start with illustrating the main results of this survey by two typical situations. In the first example, high oscillations are due to the discretization, whereas in the second example oscillations are inherent in the differential equation.

Long-term integration of Hamiltonian systems with multistep methods. We consider the equations of motion for a coupled triple pendulum in the plane as shown in the small figure. Due to the closed loop it is simpler to write them in Cartesian rather than in minimal coordinates. We denote the coordinates of the four mass points by Q_1, Q_2, Q_3, Q_4, so that (assuming mass and gravity constant equal to one) the potential energy U is the sum of the vertical components, $U = e_2^{\mathsf{T}}(Q_1 + Q_2 + Q_3 + Q_4)$. The coordinates have to satisfy five constraints: The distance between the upper mass points and their suspension point is fixed, say equal to one, and the mass points are connected by massless rods of length one as shown in the small figure (initial position is



ISSN 0020-7926

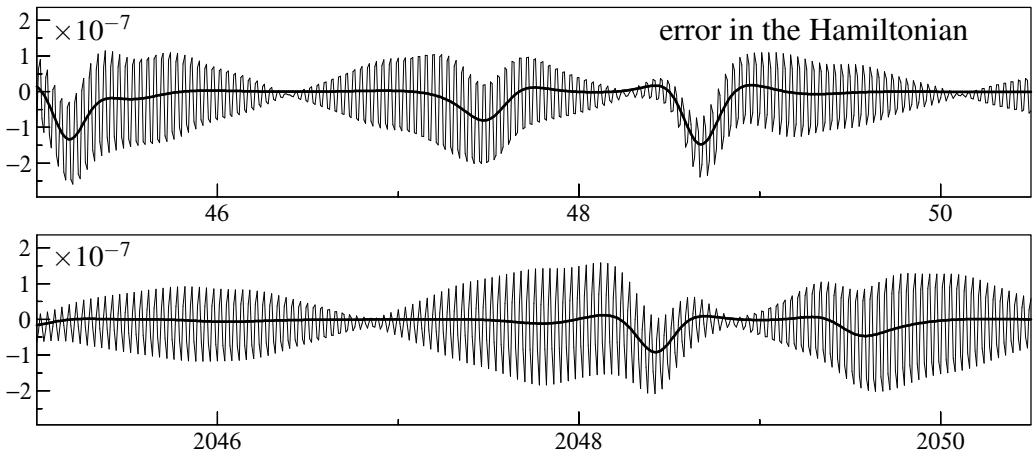


Figure 1: Numerical error in the Hamiltonian for the coupled triple pendulum problem as a function of time (two different intervals). The multistep method is applied with step size $h = 0.01$ and with two different starting approximations.

drawn in grey). We collect the coordinates of $Q_1, \ldots, Q_4$ in a vector q, and we write the constraints as $g(q) = 0$ with a quadratic function $g : \mathbb{R}^8 \to \mathbb{R}^5$. Denoting by $G(q) = g'(q)$ the matrix of partial derivatives, the equations of motion for the time-dependent position vector $q(t) \in \mathbb{R}^8$ and the Lagrange multipliers $\lambda(t) \in \mathbb{R}^5$ are given by the differential-algebraic system (with dots denoting time derivatives)

$$\ddot{q} = -\nabla U(q) - G(q)^{\mathsf{T}}\lambda, \qquad g(q) = 0 \tag{1}$$

with initial values $q(0) = q_0$, $\dot{q}(0) = \dot{q}_0$.

For its numerical integration we choose the 4-th order multistep method

$$\begin{aligned} q_{n+4} - 2q_{n+3} + 2q_{n+2} - 2q_{n+1} + q_n &= \frac{h^2}{6}\Big(7f_{n+3} - 2f_{n+2} + 7f_{n+1}\Big) \\ 0 &= g(q_{n+4}), \end{aligned} \tag{2}$$

where h is the step size, $f_n = -\nabla U(q_n) - G(q_n)^{\mathsf{T}}\lambda_n$, and $q_n \approx q(nh)$ for $n = 0, 1, 2, \ldots$. Starting approximations q_1, q_2, q_3 and λ_1, λ_2 are needed in addition to q_0. They can be obtained by a one-step Runge–Kutta method from the initial values $q(0)$ and $\dot{q}(0)$. The application of the method is as follows: For given vectors $\{q_j\}_{j \le n+3}$ and $\{\lambda_j\}_{j \le n+2}$, one first inserts q_{n+4} into $g(q_{n+4}) = 0$ and computes λ_{n+3} by modified Newton iterations, and then q_{n+4} is obtained in an explicit manner. Approximations to the velocity $p = \dot{q}$ can be obtained by finite differences. For a fourth order method we use

$$p_n = \dot{q}_n = \frac{1}{12h}\Big(8\big(q_{n+1} - q_{n-1}\big) - \big(q_{n+2} - q_{n-2}\big)\Big).$$

Figure 1 shows the total energy $H(q,p) = \frac{1}{2}\, p^{\mathsf{T}} p + U(q)$ along the numerical solution obtained with step size $h = 0.01$ on two time intervals (one in the beginning and the other much later). The thick smooth curve corresponds to very accurate starting approximations, whereas the highly oscillatory function corresponds to starting approximations that are obtained by the 2nd order implicit midpoint rule. The frequency of the oscillations is proportional to the inverse of the step size h.

Question. Classical error estimates contain a term of the form $h^4 \exp(Lt)$ with $t = nh$, where L is the Lipschitz constant of the vector field, and they are useful only on small time intervals. Numerical experiments show that the energy is well conserved and the parasitic oscillations remain small and bounded over much longer time intervals. Can this be proved rigorously?

Adiabatic invariants for highly oscillatory differential equations. As a second example we consider a chain of mass points connected by alternating stiff and soft springs (Figure 2). We let Q_k denote the displacement from the position of rest of the k-th mass point, and we assume that the potential of the stiff springs is, accord-

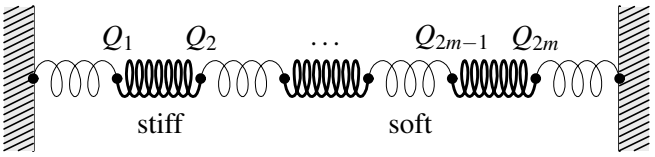


Figure 2: Chain of alternating stiff harmonic and soft nonlinear springs.

ing to Hook's law, proportional to the square of the increment due to compression or extension of the spring (with a large spring constant). The potential of the soft springs is assumed to be proportional to the fourth power of the increment (with a moderately sized spring constant). This leads to a Hamiltonian system with total energy (using $Q_0 = Q_{2m+1} = 0$)

$$H(Q,\dot{Q}) = \frac{1}{2}\sum_{k=1}^{2m} \dot{Q}_k^2 + \frac{1}{4}\sum_{j=1}^{m} \omega_j^2\,(Q_{2j} - Q_{2j-1})^2 + \sum_{j=1}^{m} (Q_{2j+1} - Q_{2j})^4,$$

where the natural frequencies ω_j are assumed to be large. After a canonical change of coordinates, where $q_j = (Q_{2j} - Q_{2j-1})/\sqrt{2}$ represents the compression/extension of a stiff spring and $q_{j+m} = (Q_{2j} + Q_{2j-1})/\sqrt{2}$ its mean position, we obtain a Hamiltonian system of the form ($j = 1,\ldots,m$)

$$\begin{aligned} \ddot{q}_j + \omega_j^2 q_j &= -\nabla_j U(q) \\ \ddot{q}_{j+m} &= -\nabla_{j+m} U(q). \end{aligned} \tag{3}$$

Here, $U(q)$ is the potential for the soft springs expressed in the new coordinates, and we use the notation $\nabla_j U(q)$ for the partial derivative of $U(q)$ with respect

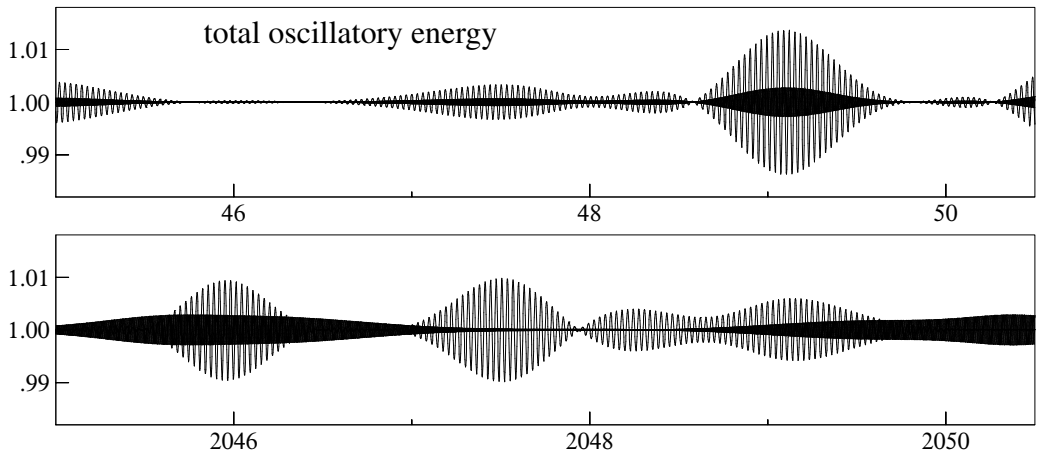


Figure 3: Total oscillatory energy for the chain of springs corresponding to frequencies proportional to 200 (thin curve) and to 1000 (thick curve).

to q_j. Neglecting the nonlinearity in (3) we are concerned with m decoupled harmonic oscillators corresponding to the m stiff springs. Their individual energy is

$$I_j(q_j,\dot q_j) = \frac{1}{2}\left(\dot q_j^2 + \omega_j^2 q_j^2\right),$$

and we let

$$I(q,\dot q) = \sum_{j=1}^{m} I_j(q_j,\dot q_j)$$

be the total oscillatory energy of the system, which is a conserved quantity in the absence of the soft springs. We are mainly interested in studying the influence of the soft springs to the long-time behaviour of $I(q,\dot q)$.

For a numerical experiment we consider the situation of Figure 2 with three stiff springs ($m = 3$), and we let their frequences be

$$\omega_1 = \frac{1}{\varepsilon}, \qquad \omega_2 = \frac{2}{\varepsilon}, \qquad \omega_3 = \frac{1.002}{\varepsilon}, \tag{4}$$

so that there is an exact resonance between the first two frequencies and a near resonance with the third one. We choose zero initial values except $q_1(0) = 1/\omega_1$, $q_4(0) = 1$, $\dot q_1(0) = 1$, and $\dot q_4(0) = 1.5$, so that the initial oscillatory energy is $I\big(q(0),\dot q(0)\big) = 1$. In Figure 3 we plot the total oscillatory energy as function of time (once on an interval of length 5.5 close to $t = 50$ and once close to $t = 2050$) for two different values of ε. The thin curve corresponds to $\varepsilon = 1/100$ and the thicker churve to $\varepsilon = 1/500$. We observe high oscillations (frequency proportional to ε^{-1} and amplitude proportional to ε) around the constant value 1.

Question. Neglecting the potential $U(q)$ in (3) the oscillatory energy $I\big(q(t),\dot q(t)\big)$ is an exact invariant of the system. Standard perturbation arguments using Gronwall's inequality yield the near-preservation of the oscillatory energy on intervals

of length $O(\varepsilon^{-1})$. This near-preservation is observed on much longer time intervals. Can this be proved rigorously?

2 Modulated Fourier expansion

Figures 1 and 3 show similar phenomena although they originate from completely different problems. We observe different time scales – high frequency oscillations superposed on a smooth slow motion. In this section we present the technique of modulated Fourier expansions, which gives much insight into the long-time behaviour for both problems.

Modulated Fourier expansion for linear multistep methods. For second order differential equations we consider linear multistep methods of the form

$$\sum_{j=0}^{k} \alpha_j q_{n+j} = h^2 \sum_{j=0}^{k} \beta_j f_{n+j}, \tag{5}$$

where $f_n = f(q_n)$ for problems $\ddot{q} = f(q)$, and $f_n = -\nabla U(q_n) - G(q_n)^{\mathsf{T}} \lambda_n$ with λ_n determined by $g(q_n) = 0$ for constrained Hamiltonian systems (1). We denote the generating polynomials of the coefficients of (5) by

$$\rho(\zeta) = \sum_{j=0}^{k} \alpha_j \zeta^j, \qquad \sigma(\zeta) = \sum_{j=0}^{k} \beta_j \zeta^j.$$

For $h \to 0$ the multistep method (5) becomes a linear difference equation, whose solution is a linear combination of ζ_j^n where $\rho(\zeta_j) = 0$. Due to the nonlinearity in the right-hand side of (5) this cannot remain true for $h > 0$. Nevertheless, it is natural to approximate the numerical solution by an expression (called modulated Fourier expansion) of the form

$$q_n \approx y(t) + \sum_{j \in I} \zeta_j^n z_j(t), \qquad t = nh, \tag{6}$$

where $y(t)$ and $z_j(t)$ are h-dependent smooth functions in the sense that, together with all their derivatives, they are bounded uniformly for $0 < h \le h_0$. The index set I corresponds to the roots of $\rho(\zeta)$ that are different from 1. For constrained Hamiltonian systems (1) it turns out that in addition to the roots of $\rho(\zeta)$, the index set I has to contain also the non-zero roots of $\sigma(\zeta)$. For the study of the long-time behaviour of linear multistep methods one has to

- study properties of the function $y(t)$, which will be an approximation to the exact solution $q(t)$ of the differential equation.
- control the parasitic components $z_j(t)$. They should remain bounded and small over long time intervals.

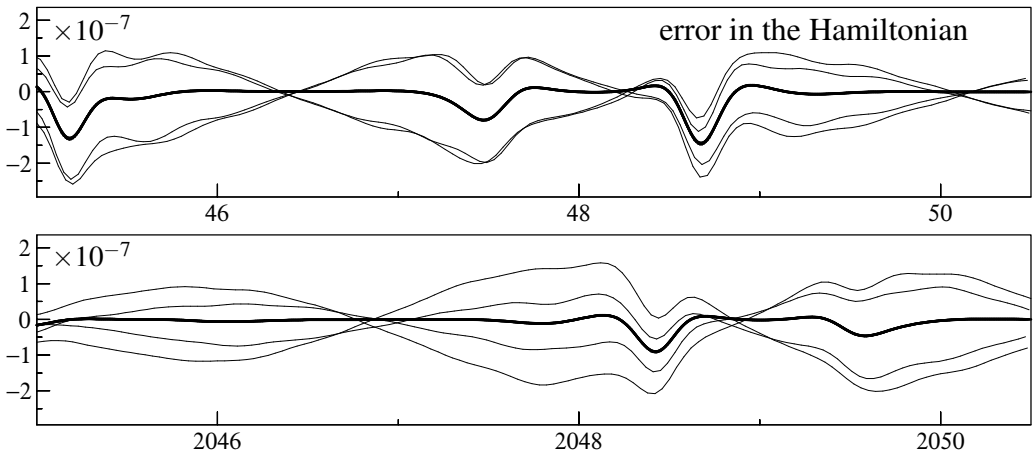


Figure 4: Numerical error in the Hamiltonian for the coupled triple pendulum problem with data as in Figure 1. Every fourth approximation is connected by a line, which gives four curves for each starting approximation.

Let us illustrate the above expansion for the numerical solution of the example of Section 1. There we have applied a multistep method with $\rho(\zeta) = (\zeta-1)^2(\zeta^2+1)$ and $\sigma(\zeta) = (7\zeta^3 - 2\zeta^2 + 7\zeta)/6$. An analysis of the functions $z_j(t)$ shows that those corresponding to zeros of $\sigma(\zeta)$ are by a factor of h smaller than those corresponding to the roots of $\rho(\zeta) = 0$. We thus have

$$q_n \approx y(t) + \mathrm{i}^n z_1(t) + (-\mathrm{i})^n z_2(t),$$

where, for real data, $y(t)$ is a real function and $z_2(t)$ is the complex conjugate of $z_1(t)$. To get an impression of the perturbation functions $z_j(t)$ we connect every fourth approximation: $q_{4n} \approx y(t) + 2\,\Re z_1(t)$ at $t = 4nh$, $q_{4n+1} \approx y(t) + 2\,\Im z_1(t)$ at $t = (4n+1)h$, $q_{4n+2} \approx y(t) - 2\,\Re z_1(t)$ at $t = (4n+3)h$, and finally $q_{4n+3} \approx y(t) - 2\,\Im z_1(t)$ at $t = (4n+3)h$. This yields four smooth curves instead of one highly oscillating curve that is obtained by connecting all q_n.

We apply the linear multistep method (2) to the coupled triple pendulum problem with the same step size and the same starting approximations as in Figure 1. Figure 4 shows the error in the Hamiltonian $H(q_n, p_n)$, where every fourth approximation is connected by a line. For the accurate starting approximation the function $z_1(t)$ is so small that all four curves coincide and give the function $H\big(y(t), \dot{y}(t)\big)$ (thick curve). For the starting approximations obtained by the implicit midpoint rule we can clearly distinguish four curves that approximate

$$H\Big(y(t) \pm 2\,\Re z_1(t), \dot{y}(t) \pm 2\,\Re \dot{z}_1(t)\Big), \quad H\Big(y(t) \pm 2\,\Im z_1(t), \dot{y}(t) \pm 2\,\Im \dot{z}_1(t)\Big).$$

This experiment illustrates that the functions $z_j(t)$ of the modulated Fourier expansion are as smooth as the approximation $y(t)$, and it encourages the study of (6) for getting insight into the long-time behaviour of multistep methods.

Modulated Fourier expansion for perturbed harmonic oscillators. The motion of a chain of mass points connected by alternating stiff and soft springs leads to a highly oscillatory differential equation, which is of the form

$$\ddot{q}_j + \omega_j^2 q_j \;=\; -\nabla_j U(\mathbf{q}), \qquad j = 0, 1, \ldots, n, \tag{7}$$

where $\mathbf{q} = (q_0, q_1, \ldots, q_n)$ with $q_j \in \mathbb{R}^{d_j}$, and ∇_j denotes the partial derivative with respect to q_j. We assume $\omega_0 = 0$ and

$$\omega_j \geq \frac{1}{\varepsilon}, \qquad 0 < \varepsilon \ll 1, \qquad j = 1, \ldots, n.$$

Suppressing the right-hand side in (7), the problem is reduced to harmonic oscillators with solution $q_j(t) = c_1 \mathrm{e}^{\mathrm{i}\omega_j t} + c_2 \mathrm{e}^{-\mathrm{i}\omega_j t}$. It seems therefore natural to approximate the solution $q(t)$ of the complete nonlinear system (7) with a linear combination of $\mathrm{e}^{\pm \mathrm{i}\omega_j t}$, where the coefficients depend smoothly on t. However, since the oscillations are here of much larger amplitude than the numerical (artificial) oscillations in computations with multistep methods, this is not sufficient. Due to the mixing of exponentials by the nonlinearity, we have to consider also products of such exponential terms. For an analysis of the solutions of (7) we thus consider an ansatz (also called modulated Fourier expansion) of the form

$$\mathbf{q}(t) \;\approx\; \mathbf{y}(t) + \sum_{\mathbf{k} \in \mathcal{K}} \mathbf{z}^{\mathbf{k}}(t)\, \mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\omega)t}. \tag{8}$$

For a multi-index $\mathbf{k} = (k_1, \ldots, k_n) \in \mathbb{Z}^n$ and the vector $\omega = (\omega_1, \ldots, \omega_n)$ of high frequencies we write $\mathbf{k} \cdot \omega = k_1\, \omega_1 + \ldots + k_n\, \omega_n$, so that $\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\omega)t}$ becomes a product of pure exponentials $\mathrm{e}^{\mathrm{i}\omega_j t}$. We also consider the norm $\|\mathbf{k}\| = |k_1| + \ldots + |k_n|$. The coefficient functions $\mathbf{y}(t)$ and $\mathbf{z}^{\mathbf{k}}(t)$ are vector-valued with the same dimension and partitioning as $\mathbf{q}(t)$ and they are assumed to be smooth. This means that together with all their derivatives they are bounded independently of ε for $0 < \varepsilon \leq \varepsilon_0$, so that the high oscillations are well separated from the slow motion in (8). The sum in (8) is over a suitably chosen finite set of multi-indices $\mathcal{K} \subset \{\mathbf{k}; \|\mathbf{k}\| \leq N\}$ with suitably chosen N. For getting insight into the long-time behaviour of the total oscillatory energy

$$I(\mathbf{q}, \dot{\mathbf{q}}) = \sum_{j=1}^{n} I_j(q_j, \dot{q}_j), \qquad I_j(q_j, \dot{q}_j) = \frac{1}{2}\Big(|\dot{q}_j|^2 + \omega_j^2\, |q_j|^2 \Big),$$

where $|q_j|$ denotes the Euclidean norm in $\mathbb{R}^{d_j}$, we are confronted with the following problems:

- find a relation between the total oscillatory energy $I(\mathbf{q}, \dot{\mathbf{q}})$ and the coefficient functions of the modulated Fourier expansion.

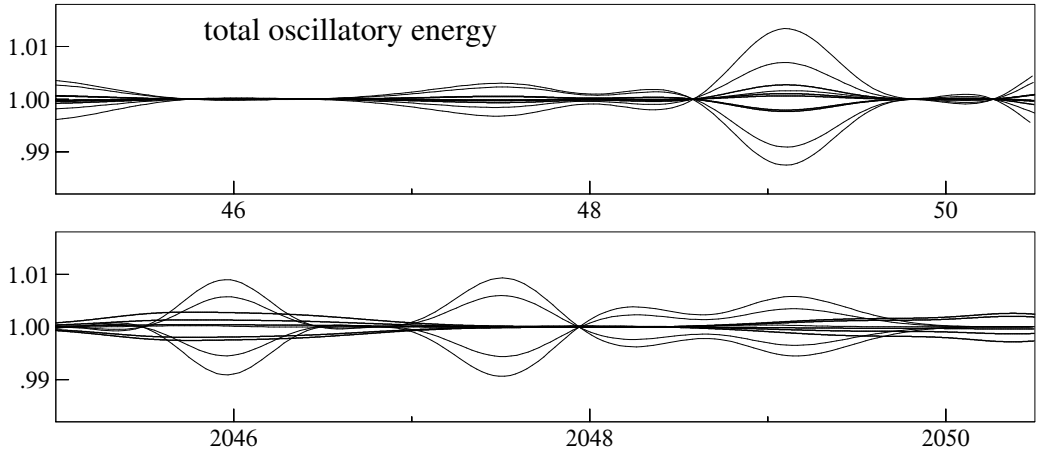


Figure 5: Total oscillatory energy connected at stroboscopic samples for the chain of springs with data as in Figure 3.

- control the coefficient functions $\mathbf{z}^{\mathbf{k}}(t)$ corresponding to high oscillations. They should remain bounded of size $\|\mathbf{z}^{\mathbf{k}}(0)\|$ over long time intervals.

To become convinced of the existence of an expansion (8) we perform a stroboscopic sampling, similar to our previous experiment with a linear multistep method (see Figure 4). We consider the example of Section 1 with three stiff springs having frequencies (4). We assume that the dominant oscillating terms in (8) are those corresponding to $\mathbf{k} = \pm\langle j\rangle$, where $\langle j\rangle = (0,\ldots,1,\ldots,0)$ is the j-th unit vector. They are given by $\mathbf{z}^{\pm\langle j\rangle}(t)\mathrm{e}^{\pm \mathrm{i}\omega_j t}$. At stroboscopic time instances $t = t_n = t_0 + 2\pi\varepsilon n$ we thus have

$$\mathbf{q}(t) \;\approx\; \mathbf{y}(t) + 2\,\Re\Big(\mathbf{z}^{\langle 1\rangle}(t)\mathrm{e}^{\mathrm{i}\omega_1 t_0} + \mathbf{z}^{\langle 2\rangle}(t)\mathrm{e}^{\mathrm{i}\omega_2 t_0} + \mathbf{z}^{\langle 3\rangle}(t)\mathrm{e}^{\mathrm{i}\omega_3 t_0}\Big),$$

which is expected to be smooth in contrast to the highly oscillatory solution (see Figure 3) of the differential equation. In Figure 5 we interpolate the oscillatory energies $\big\{I\big(\mathbf{q}(t_n),\dot{\mathbf{q}}(t_n)\big)\big\}_{n\geq 0}$ for five different values of t_0, namely $t_0 = 2\pi\varepsilon l/5$ for $l = 1,\ldots,5$. Thin curves correspond to $\varepsilon = 1/200$, thick curves to $\varepsilon = 1/1000$. It is striking how smooth these curves are (higher order terms of (8) are present but not visible in the plot), and that their qualitative behaviour is maintained on very long time intervals (on much longer intervals than shown in the figure).

Connection between both types of modulated Fourier expansions. Comparing the modulated Fourier expansion (6) for linear multistep methods with the expansion (8) for highly oscillatory Hamiltonian systems, we notice the correspondence $\zeta_j^n \leftrightarrow \mathrm{e}^{\mathrm{i}\omega_j t}$ for $t = nh$. To a zero $\zeta_j = \mathrm{e}^{\mathrm{i}\theta_j}$ of $\rho(\zeta)$ (resp. $\sigma(\zeta)$) corresponds a frequency $\omega_j = \theta_j/h$ in the problem (7).

3 Results on the long-time behaviour

In this section we present two typical results that can be proved with the technique of modulated Fourier expansions – one on the energy preservation of linear multistep methods, the other on the near-preservation of the total oscillatory energy of highly oscillatory systems. We also give an overview on further results and hints to the literature.

Near-preservation of energy with linear multistep methods. We consider a method (5) and apply it to a constrained Hamiltonian system (1) with total energy $H(q,\dot q) = \frac{1}{2}\dot q^{\mathsf T}\dot q + U(q)$. We assume that the method coefficients satisfy

- the method is irreducible and of order p, i.e., the generating polynomials $\rho(\zeta)$ and $\sigma(\zeta)$ do not have common roots, and they satisfy

$$\frac{\rho(\zeta)}{(\log\zeta)^2} - \sigma(\zeta) = O\big((\zeta-1)^p\big) \qquad \text{for} \quad \zeta \to 1.$$

- it is symmetric, i.e., $\alpha_j = \alpha_{k-j}$ and $\beta_j = \beta_{k-j}$ for all j.
- the equation $\rho(\zeta) = 0$ has only simple roots with the exception of the double root for $\zeta = 1$; all roots are on the unit circle.
- the equation $\sigma(\zeta) = 0$ has only simple non-zero roots; all non-zero roots are on the unit circle.
- starting approximations satisfy

$$\begin{aligned} q_j - q(jh) &= O(h^{p+2}) \qquad \text{and} \quad g(q_j) = 0 \quad \text{for} \quad j = 0,\dots,k-1 \\ \lambda_j - \lambda(jh) &= O(h^p) \qquad\quad \text{for} \quad j = 1,\dots,k-2 \end{aligned}$$

  (the latter for the case of an explicit method with $\beta_k = 0$ and $\beta_{k-1} \neq 0$).

The following result is taken from [11].

Theorem 1. *The numerical solution of a linear multistep method satisfying the above assumptions preserves the total energy of the constrained Hamiltonian system (1) up to $O(h^p)$ over time $O(h^{-p-1})$:*

$$H(q_n,\dot q_n) = H(q_0,\dot q_0) + O(h^p) \qquad \textit{for} \quad nh \le h^{-p-1}.$$

The constant symbolized by O is independent of n and h subject to $nh \le h^{-p-1}$.

This result explains the excellent long-time energy preservation in the experiment of Section 1 (Figure 1). The ideas of the proof will be outlined in Section 4.

Hints to the literature. The first rigorous proofs of long-time energy preservation of linear multistep methods for second order Hamiltonian systems without constraints is given in [21], see also [24]. There, the technique of modulated Fourier expansion is further used to prove near-preservation of momentum and an at most linear growth of the global error for integrable Hamiltonian systems. These results are extended in [11] to constrained Hamiltonian systems (1). The technique of modulated Fourier expansion can also be applied to general first order Hamiltonian systems [19]. However, there one typically gets near-preservation of energy only over time intervals of length $O(h^{-2})$. This is elaborated in [10] for partitioned linear multistep methods and in [12] for general multi-value methods.

Near-preservation of the total oscillatory energy in differential equations. We consider a highly oscillatory Hamiltonian system of the form

$$\ddot{q}_j + \omega_j^2 q_j \;=\; -\nabla_j U(\mathbf{q}), \qquad j = 0, 1, \ldots, n, \tag{9}$$

where $\mathbf{q} = (q_0, q_1, \ldots, q_n)$ with $q_j \in \mathbb{R}^{d_j}$, and $\omega_0 = 0$, and we are interested in the long-time behaviour of the total oscillatory energy

$$I(\mathbf{q}, \dot{\mathbf{q}}) = \sum_{j=1}^{n} I_j(q_j, \dot{q}_j), \qquad I_j(q_j, \dot{q}_j) = \frac{1}{2}\Big(|\dot{q}_j|^2 + \omega_j^2 \, |q_j|^2 \Big). \tag{10}$$

We make the following assumptions:

- the frequencies $\omega_1, \ldots, \omega_n$ are bounded from below as
$$\omega_j \geq \frac{1}{\varepsilon}, \qquad 0 < \varepsilon \ll 1, \qquad j = 1, \ldots, n.$$
- there exist $\delta > 0$ and a set $K \subset \mathbb{R}^{d_0}$ such that the potential U has bounded derivatives of all orders in a δ-neighborhood of $K \times 0 \times \cdots \times 0$.
- the initial values $\mathbf{q}(0)$, $\dot{\mathbf{q}}(0)$ are such that
$$I\big(\mathbf{q}(0), \dot{\mathbf{q}}(0)\big) \leq E, \tag{11}$$
where the bound E is independent of ε.

The following result is taken from [15].

Theorem 2. *Consider the Hamiltonian initial value problem (9) satisfying the above assumptions. For an arbitrarily fixed integer $N \geq 1$ there then exist $C > 0$ and $\varepsilon^* > 0$ such that the following holds for $0 < \varepsilon \leq \varepsilon^*$: along the solution of (9) the oscillatory energy deviates from its starting value by no more than*

$$\big| I\big(\mathbf{p}(t), \mathbf{q}(t)\big) - I\big(\mathbf{p}(0), \mathbf{q}(0)\big) \big| \leq C \varepsilon^{3/4} \quad \textit{for} \quad 0 \leq t \leq \varepsilon^{-N}, \tag{12}$$

provided that $q_0(t)$ stays in the set K for such long times. The threshold ε^* and the constant C depend on n and N, on the energy bound E and on bounds of derivatives of the potential U.

This result explains the excellent long-time preservation of the total oscillatory energy in the experiment of Figure 3. Our proof is based on modulated Fourier expansions and will be outlined in Section 4.

Hints to the literature. Modulated Fourier expansions for the long-term analysis of highly oscillatory differential equations have first been used in [20] for the case of a single high frequency ω. Exponentially long times $t \leq e^{c\omega}$ are covered in [6]. The case of several high frequencies satisfying a non-resonance condition is studied in [7]. The result of Theorem 2, taken from [15], does not require any condition on the high frequencies. Modulated Fourier expansions have also given new insight into the FPU problem [23]. They have been successfully applied to studying long-time regularity and stability of Hamiltonian partial differential equations (nonlinear wave equation [9, 16], nonlinear Schrödinger equation [14, 17]).

Related results have been obtained in [2], [3], [4], [1] with canonical transformation techniques of Hamiltonian perturbation theory.

Modulated Fourier expansions have in addition proved very useful in analyzing numerical methods for highly oscillatory differential equations [24, 20, 7, 25] and Hamiltonian partial differential equations [8, 5, 13, 18, 22].

4 Sketch of the proof

Although Theorems 1 and 2 treat two completely different situations, their proofs with modulated Fourier expansions are closely connected. We sketch the main steps in the proof. For details we refer the reader to the original literature.

4.1 Construction of the coefficient functions

The first step of the proof consists in the construction of coefficient functions for the modulated Fourier expansion.

Linear multistep methods. To keep the presentation as simple as possible we restrict ourselves to Hamiltonian systems without constraints. We consider the method (5) and we are looking for approximations to q_n of the form

$$\widehat{q}_n = y(t) + \sum_{j \in I} \zeta_j^n z_j(t), \qquad t = nh, \tag{13}$$

where the index set I corresponds to roots of $\rho(\zeta)$ that are different from 1. To achieve this we insert this expansion into the multistep formula and compare the coefficients of ζ_j^n. Writing the Taylor series of a function as $y(t+h) = \mathrm{e}^{hD} y(t)$,

where D denotes differentiation with respect to time, this yields

$$\begin{aligned} \rho(\mathrm{e}^{hD})\, y &= h^2\, \sigma(\mathrm{e}^{hD})\, f(y) + O(h^2\|\mathbf{z}\|^2) \\ \rho(\zeta_j \mathrm{e}^{hD})\, z_j &= h^2\, \sigma(\zeta_j \mathrm{e}^{hD})\, f'(y)\, z_j + O(h^2\|\mathbf{z}\|^2) \end{aligned} \tag{14}$$

with the notation $\|\mathbf{z}\| = \max_j |z_j|$. Since $\zeta = 1$ is a double root of $\rho(\zeta) = 0$ we have $\rho(\mathrm{e}^{hD}) = h^2 D^2 \rho_0(hD)$, and for the simple roots ζ_j we have $\rho(\zeta_j \mathrm{e}^{hD}) = hD\rho_j(hD)$, where $\rho_j(0) \neq 0$ for all j. Applying the inverse of the operators $\rho_j(hD)$ to these relations yields (after truncation and omitting the remainder) the differential equations

$$\begin{aligned} \ddot{y} &= f(y) + hf_1(y) + h^2 f_2(y) + \ldots + h^{N-1} f_{N-1}(y) \\ \dot{z}_j &= h\left(a_1(y) + ha_2(y) + \ldots + h^{N-1} a_{N-1}(y)\right) z_j. \end{aligned} \tag{15}$$

This construction guarantees that for every solution of the system (15) the approximations $\widehat{q}_n$ of (13) satisfy the multistep relation with a defect of size $O(h^{N+2}) + O(h^2\|\mathbf{z}\|^2)$. Initial values for (15) are obtained from the starting approximations $q_0, q_1, \ldots, q_{k-1}$ by putting $t = 0, h, \ldots, (k-1)h$ in (13). We notice that both the coefficient functions in (15) and the initial values for the system (15) are uniquely determined.

Oscillatory differential equation. For the solution $\mathbf{q}(t) = \big(q_0(t), q_1(t), \ldots, q_n(t)\big)$ of the system (9) we consider an approximation

$$\widehat{\mathbf{q}}(t) = \mathbf{y}(t) + \sum_{\mathbf{k}\in\mathcal{K}} \mathbf{z}^{\mathbf{k}}(t)\, \mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\omega)t}. \tag{16}$$

We let $\mathbf{y}(t) = \mathbf{z}^{\mathbf{0}}(t)$, and we denote the components of the super-vector $\mathbf{z}^{\mathbf{k}}(t)$ by $z_j^{\mathbf{k}}(t)$, $j = 0, 1, \ldots, n$. To avoid technical difficulties we assume here that the frequencies $\omega_1, \ldots, \omega_n$ are non-resonant, so that $\mathbf{k}\cdot\omega \neq 0$ for all multi-indices in the set $\mathcal{K}\setminus\{\mathbf{0}\}$. Inserting this ansatz into the differential equation (9) and comparing the coefficients of $\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\omega)t}$ yields

$$\big(\omega_j^2 - (\mathbf{k}\cdot\omega)^2\big)\, z_j^{\mathbf{k}} + 2\,\mathrm{i}\,(\mathbf{k}\cdot\omega)\,\dot{z}_j^{\mathbf{k}} + \ddot{z}_j^{\mathbf{k}} = \sum_{m\geq 0} \sum_{s(\alpha)=\mathbf{k}} \frac{1}{m!}\, g_j^{(m)}(\mathbf{y})\big(\mathbf{z}^{\alpha_1}, \ldots, \mathbf{z}^{\alpha_m}\big), \tag{17}$$

where we use the notation $g_j(\mathbf{y}) = \nabla_j U(\mathbf{y})$ for the derivative of the potential with respect to y_j. The second sum is over m-tuples of non-zero multi-indices $\alpha_1, \ldots, \alpha_m$, such that $s(\alpha) = \alpha_1 + \ldots + \alpha_m = \mathbf{k}$.

To get smooth coefficient functions (with derivatives bounded independently of ε) we determine the dominating term in the left-hand expression, we put the other terms to the right-hand side, and we iteratively eliminate the higher derivatives. After suitable truncation this yields a second order differential equation for the case $(j, \mathbf{k}) = (0, \mathbf{0})$, first order differential equations for $\mathbf{k} = \pm\langle j\rangle$ with $j \geq 1$, and algebraic relations for all other situations.

When dividing the equation by $(\mathbf{k}\cdot\omega)$ or by $\left(\omega_j^2-(\mathbf{k}\cdot\omega)^2\right)$, one may encounter small denominators. To avoid this difficulty, this requires either non-resonance assumptions for the frequencies or the introduction of modified frequencies.

For every solution of the resulting differential-algebraic system, the function $\widehat{\mathbf{q}}(t)$ of (16) satisfies the equation (9) up to a small defect. Unique initial values for the functions $z_j^{\mathbf{k}}$ are obtained from $\mathbf{q}_0, \dot{\mathbf{q}}_0$ by putting $\widehat{\mathbf{q}}(0)=\mathbf{q}_0, \dot{\widehat{\mathbf{q}}}(0)=\dot{\mathbf{q}}_0$.

4.2 Formal invariants for the modified equations

The second step heavily relies on the fact that we are concerned with Hamiltonian differential equations. For this situation we derive invariants for the coefficient functions of the modulated Fourier expansion.

Linear multistep methods. We again consider the relation (14), but instead of multiplying it by the inverse of a factor of $\rho(\zeta_j e^{hD})$, we multiply it by the inverse of the operator $\sigma(\zeta_j e^{hD})$. Expanding $\sigma(\zeta_j e^{hD})^{-1}\rho(\zeta_j e^{hD})$ into a series of powers of h this yields (notice that $\sigma(\zeta_j)\neq 0$ and ζ_j is a simple zero of $\rho(\zeta)$)

$$c_{j1}\,\mathrm{i}h\dot{z}_j + c_{j2}\,h^2\ddot{z}_j + c_{j3}\,\mathrm{i}h^3\dddot{z}_j + \ldots \;=\; h^2 f'(y)\,z_j + O(h^2\|\mathbf{z}\|^2) \qquad (18)$$

with real coefficients c_{jl}, and $c_{j1}\neq 0$. We then take the scalar product with the complex conjugate vector $\bar{z}_j$. For Hamiltonian systems we have $f(y)=-\nabla U(y)$, so that $f'(y)$ is a symmetric matrix and $\Im(\bar{z}_j^{\mathsf{T}} f'(y) z_j)=0$. Furthermore, we have $2\Re(\bar{z}_j^{\mathsf{T}}\dot{z}_j)=\frac{d}{dt}\|z_j\|^2$, and it turns out that the imaginary part of the whole left-hand side of (18) can be written as a total differential of an expression, which after division by $c_{j1}h$ and suitable truncation, is of the form

$$K_j(y,z_j)=\|z_j\|^2+\bar{z}_j^{\mathsf{T}}\big(hE_{j,1}(y)+\ldots+h^{N-1}E_{j,N-1}(y)\big)\,z_j.$$

Consequently, up to a small truncation error the function $K_j(y,z_j)$ is an invariant of the system (15) for the coefficient functions of the modulated Fourier expansion. This is essential for a long-time bound of the parasitic components in symmetric multistep methods.

Oscillatory differential equation. For the system (9) one can recover a Hamiltonian structure in the equations (17) for the coefficient functions. With the scalar function

$$\mathcal{U}(\mathbf{z})=U(\mathbf{y})+\sum_{m=1}^{N}\sum_{j_1,\ldots,j_m=0}^{n}\sum_{\mathbf{k}_1+\ldots+\mathbf{k}_m=\mathbf{0}}\frac{1}{m!}\nabla^m_{j_1,\ldots,j_m}U(\mathbf{y})\big(z_{j_1}^{\mathbf{k}_1},\ldots,z_{j_m}^{\mathbf{k}_m}\big),$$

the right-hand side of (17), up to a defect due to the truncation of the series, can be written as $\nabla_j^{-\mathbf{k}}\mathcal{U}(\mathbf{z})$, where $\nabla_j^{-\mathbf{k}}$ denotes the derivative with respect to $z_j^{-\mathbf{k}}$. The

important observation is now that the extended potential $\mathcal{U}(\mathbf{z})$ is invariant under the action of the one-parameter group $S(\theta)\,\mathbf{z} = (\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\omega)\theta} z_j^{\mathbf{k}})$, i.e.,

$$\mathcal{U}\big(S(\theta)\,\mathbf{z}\big) = \mathcal{U}(\mathbf{z}) \qquad \text{for all } \theta \in \mathbb{R}.$$

By Noether's theorem this implies the existence of a conserved quantity of the system (17). It turns out that this conserved quantity is close to the total oscillatory energy of the Hamiltonian equation (9).

4.3 From short to long time intervals

In Section 4.1 we have constructed smooth coefficient functions for the modulated Fourier expansion. The approximation (13), when inserted in the multistep formula, yields a small defect of size $O(h^2\delta^2)$ as long as $\|\mathbf{z}\| \le \delta$. Similarly, the approximation (16), when inserted into the oscillatory differential equation (9), yields a small defect of size ε^N. In both situations this implies that on a short interval of size $O(1)$ the approximation coincides with the exact solution up to an error of the size of the defect.

An application of Gronwall-type estimates then yields bounds on the difference $\widehat{q}(t) - q(t)$, respectively $\widehat{\mathbf{q}}(t) - \mathbf{q}(t)$, that grow exponentially with time. Such estimates cannot be improved in general, because the formulation of our two problems admits chaotic problems, where perturbations in initial values propagate exponentially with time. However, they are not useful for a proof of Theorems 1 and 2.

At this point we need the existence of invariants (Section 4.2) for the system defining the coefficient functions of the modulated Fourier expansion. These invariants have to be close to the quantities that we want to have under control. We want to get bounds for the spurious oscillations in the application of symmetric multistep methods and for the variation of the total oscillatory energy in highly oscillatory differential equations.

Let us explain the ideas for estimating the parasitic oscillations of multistep methods over long times. We consider a grid $0 = t_0 < t_1 < t_2 < \ldots$, such that $t_{m+1} - t_m = O(1)$. On the m-th subinterval we consider the coefficient functions $y^{[m]}(t)$, $z_j^{[m]}(t)$ of the modulated Fourier expansion which are defined by the condition $\widehat{q}(t_m) = q_m$, where q_m denotes the numerical solution of the multistep method at time $t = t_m$. From Section 4.2 we know that

$$\Big\| K_j\big(y^{[m]}(t_m), z_j^{[m]}(t_m)\big) - K_j\big(y^{[m]}(t_{m+1}), z_j^{[m]}(t_{m+1})\big)\Big\| \le \delta_1,$$

where δ_1 is a small quantity. On the small interval $[t_m, t_{m+1}]$ the difference between the modulated Fourier expansion $\widehat{q}_n$ and the exact numerical values q_n is bounded by the defect in the multistep formula. By the uniqueness of the coefficients of

the Fourier expansion (Section 4.1) we therefore have

$$\left\| K_j\left(y^{[m]}(t_{m+1}), z_j^{[m]}(t_{m+1})\right) - K_j\left(y^{[m+1]}(t_{m+1}), z_j^{[m+1]}(t_{m+1})\right)\right\| \leq \delta_2,$$

where δ_2 is proportional to the defect. Summing up these two estimates, using the triangle inequality and a telescopic summation, we obtain

$$\left\| K_j\left(y^{[m]}(t_m), z_j^{[m]}(t_m)\right) - K_j\left(y^{[0]}(t_0), z_j^{[0]}(t_0)\right)\right\| \leq m(\delta_1 + \delta_2).$$

Since $K_j(y, z_j)$ is close to $\|z_j\|^2$, this proves that the parasitic components $z_j(t)$ remain small over very long time intervals.

References

[1] D. Bambusi, *Birkhoff normal form for some nonlinear PDEs*, Comm. Math. Phys. **234** (2003), 253–285.

[2] G. Benettin, L. Galgani, and A. Giorgilli, *Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. Part I*, Comm. Math. Phys. **113** (1987), 87–103.

[3] ———, *Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. Part II*, Comm. Math. Phys. **121** (1989), 557–601.

[4] J. Bourgain, *Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations*, Geom. Funct. Anal. **6** (1996), 201–230.

[5] D. Cohen and L. Gauckler, *One-stage exponential integrators for nonlinear Schrödinger equations over long times*, BIT **52** (2012), 877–903.

[6] D. Cohen, E. Hairer, and C. Lubich, *Modulated Fourier expansions of highly oscillatory differential equations*, Foundations of Comput. Math. **3** (2003), 327–345.

[7] ———, *Numerical energy conservation for multi-frequency oscillatory differential equations*, BIT **45** (2005), 287–305.

[8] ———, *Conservation of energy, momentum and actions in numerical discretizations of nonlinear wave equations*, Numer. Math. **110** (2008), 113–143.

[9] ———, *Long-time analysis of nonlinearly perturbed wave equations via modulated Fourier expansions*, Arch. Ration. Mech. Anal. **187** (2008), 341–368.

[10] P. Console and E. Hairer, *Long-term stability of symmetric partitioned linear multistep methods*, Current challenges in stability issues for numerical differential equations, Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2013, pp. 1–36.

[11] P. Console, E. Hairer, and C. Lubich, *Symmetric multistep methods for constrained Hamiltonian systems*, Numerische Mathematik (2013).

[12] R. D'Ambrosio and E. Hairer, *Long-term stability of multi-value methods for ordinary differential equations*, Submitted for publication (2013).

[13] E. Faou, L. Gauckler, and C. Lubich, *Plane wave stability of the split-step Fourier method for the nonlinear Schrödinger equation*, Preprint (2013).

[14] ———, *Sobolev stability of plane wave solutions to the cubic nonlinear Schrödinger equation on a torus*, Comm. Partial Differential Equations (2013).

[15] L. Gauckler, E. Hairer, and C. Lubich, *Energy separation in oscillatory Hamiltonian systems without any non-resonance condition*, Comm. Math. Phys. (2013).

[16] L. Gauckler, E. Hairer, C. Lubich, and D. Weiss, *Metastable energy strata in weakly nonlinear wave equations*, Comm. Partial Differential Equations **37** (2012), no. 8, 1391–1413.

[17] L. Gauckler and C. Lubich, *Nonlinear Schrödinger equations and their spectral semi-discretizations over long times*, Found. Comput. Math. **10** (2010), 141–169.

[18] L. Gauckler and C. Lubich, *Splitting integrators for nonlinear Schrödinger equations over long times*, Found. Comput. Math. **10** (2010), 275–302.

[19] E. Hairer, *Backward error analysis for multistep methods*, Numer. Math. **84** (1999), 199–232.

[20] E. Hairer and C. Lubich, *Long-time energy conservation of numerical methods for oscillatory differential equations*, SIAM J. Numer. Anal. **38** (2001), 414–441.

[21] ———, *Symmetric multistep methods over long times*, Numer. Math. **97** (2004), 699–723.

[22] ———, *Spectral semi-discretisations of weakly nonlinear wave equations over long times*, Found. Comput. Math. **8** (2008), 319–334.

[23] ———, *On the energy distribution in Fermi–Pasta–Ulam lattices*, Arch. Ration. Mech. Anal. **205** (2012), no. 3, 993–1029.

[24] E. Hairer, C. Lubich, and G. Wanner, *Geometric numerical integration. Structure-preserving algorithms for ordinary differential equations*, 2nd ed., Springer Series in Computational Mathematics 31, Springer-Verlag, Berlin, 2006.

[25] J. M. Sanz-Serna, *Modulated Fourier expansions and heterogeneous multi-scale methods*, IMA J. Numer. Anal. **29** (2009), 595–605.

Author's address:

Ernst Hairer
Université de Genève
Section de mathématiques
2-4, rue du Lièvre, 1211 Genève 4, Switzerland
e-mail Ernst.Hairer@unige.ch

Christian Lubich
Universität Tübingen
Mathematisches Institut
Auf der Morgenstelle 10, 72076 Tübingen, Germany
e-mail Lubich@na.uni-tuebingen.de