Challenges in Geometric Numerical Integration

Ernst Hairer

Abstract Geometric Numerical Integration is a subfield of the numerical treatment of differential equations. It deals with the design and analysis of algorithms that preserve the structure of the analytic flow. The present review discusses numerical integrators, which nearly preserve the energy of Hamiltonian systems over long times. Backward error analysis gives important insight in the situation, where the product of the step size with the highest frequency is small. Modulated Fourier expansions permit to treat nonlinearly perturbed fast oscillators. A big challenge that remains is to get insight into the long-time behavior of numerical integrators for fully nonlinear oscillatory problems, where the product of the step size with the highest frequency is not small.

1 Geometric Numerical Integration

Ordinary differential equations arise everywhere in science and their numerical treatment is of great importance. The development took place in three periods: the numerical solution of non-stiff differential equations started in the end of the 19th century, whereas that of stiff problems began in the middle of the 20th century and was motivated by space discretizations of parabolic differential equations and by simulations of chemical reactions. With the interest in computations over long time intervals one discovered that certain methods reproduce the qualitative behavior of the exact flow much better than others. In the late 1980ties numerical analysts started to design and study (we quote from the preface of the monograph [9])

> ... numerical methods that preserve geometric properties of the flow of a differential equation: symplectic integrators for Hamiltonian systems, symmetric integrators for reversible systems, ... and methods for problems with highly oscillatory solutions.

Ernst Hairer
Sect. de mathématiques, 2-4 rue du Lièvre, Univ. de Genève, CH-1211 Genève 4, Switzerland, e-mail: Ernst.Hairer@unige.ch

This period is called "Geometric Numerical Integration" and much research has been devoted to this topic during the last decades. Special attention has been paid to the long-time integration of Hamiltonian systems and, in particular, to simulations in astronomy (planetary motion) and in molecular dynamics. The present work focuses on algorithms that nearly preserve energies (total and oscillatory) over long time intervals. Our main interest is in getting theoretical insight into their long-time behavior. We distinguish between three degrees of difficulty:

- *non oscillatory Hamiltonian systems:* the term 'non oscillatory' means that the product of the highest frequency in the system with the time step size of the numerical integrator is small. In this situation *backward error analysis* gives much insight into the long-time behavior of numerical solutions.
- *nonlinearly perturbed fast oscillators:* if several high frequency harmonic oscillators are nonlinearly coupled, then the technique of *modulated Fourier expansions* yields much information on the numerical preservation of total and oscillatory energies over long times.
- *fully nonlinear, highly oscillatory Hamiltonian problems:* this is the situation where high frequencies stem from a nonlinear part in the problem, and the numerical integrator is applied, such that the product of the time step size with the highest frequency is not small. A good understanding of the long-time behavior (e.g., near energy preservation) of numerical solutions is still missing.

In the following sections each of these types of problems is treated individually.

2 Hamiltonian Systems – Backward Error Analysis

We start by considering Hamiltonian systems

$$\dot{p} = -\nabla_q H(p,q), \qquad \dot{q} = \nabla_p H(p,q),$$

where $H(p,q)$ is a smooth scalar function (called total energy) of position variables $q \in \mathbb{R}^d$ and momenta $p \in \mathbb{R}^d$. A characteristic property of such systems is that the exact flow is *symplectic*, i.e., the derivative of the flow map φ_t with respect to initial values satisfies $(\varphi_t')^{\mathsf{T}} J \varphi_t' = J$ for the canonical structure matrix J (see [9, Chapter VI] for more details). Moreover, the total energy $H(p,q)$ is preserved along the exact flow of the system. Here, we are interested to which extent the total energy can also be preserved by a numerical solution.

Numerical Experiment. As a representative example of numerical integrators we consider the *explicit Euler method*

$$\begin{aligned} p_{n+1} &= p_n - h\nabla_q H(p_n,q_n) \\ q_{n+1} &= q_n + h\nabla_p H(p_n,q_n). \end{aligned}$$

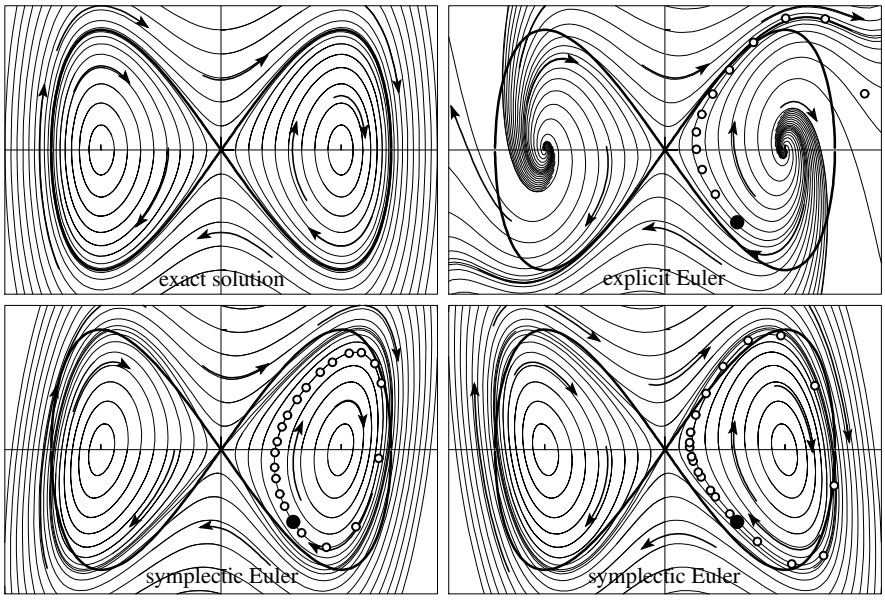


Fig. 1 Illustration of backward error analysis for the Hamiltonian $H(p,q) = \frac{1}{2}p^2 + U(q)$ with the double well potential $U(q) = \frac{1}{2}(q^2-1)^2$. The numerical solution for an initial value, indicated by a big bullet, and solutions of the modified differential equation are shown in the (q,p) phase space.

It provides approximations (p_n, q_n) to the solution of the Hamiltonian system at time $t = nh$. This method and most of the classical Runge–Kutta and multistep methods are not suited for the long-time integration of Hamiltonian systems. Even for very simple problems a linear drift in the energy $H(p_n, q_n)$ along their numerical solution can be observed.

Already very early de Vogelaere [5] noticed that each of the so-called *symplectic Euler methods*

$$\begin{aligned} p_{n+1} &= p_n - h\nabla_q H(p_{n+1}, q_n) \qquad & p_{n+1} &= p_n - h\nabla_q H(p_n, q_{n+1}) \\ q_{n+1} &= q_n + h\nabla_p H(p_{n+1}, q_n) \qquad & q_{n+1} &= q_n + h\nabla_p H(p_n, q_{n+1}) \end{aligned}$$

has a much better long-time behaviour. This can be seen in the experiment of Fig. 1, where the explicit and both symplectic Euler methods are applied to a simple Hamiltonian system (with step size $h = 0.3$). Whereas the numerical solution of the explicit Euler method spirals outwards and gives a qualitatively wrong approximation to the exact flow, those of the symplectic Euler methods remain apparently on a closed curve.

Backward Error Analysis for the Example of Figure 1. The idea of backward error analysis is the following: for a given numerical integrator search a modified differential equations, such that the exact solution of this modified equation approximates very well the numerical solution. An analysis of the modified differential

equation then gives much insight into the numerical flow. The construction of the modified differential equation is straight-forward. One makes an ansatz as a formal series in powers of the step size h, inserts its solution into the numerical method, and compares like powers of h. For a problem of the form $\dot q = p$, $\dot p = -U'(q)$ the explicit Euler method yields

$$\begin{pmatrix} \dot q \\ \dot p \end{pmatrix} = \begin{pmatrix} p \\ -U'(q) \end{pmatrix} + \frac{h}{2} \begin{pmatrix} U'(q) \\ U''(q)p \end{pmatrix} + \frac{h^2}{4} \begin{pmatrix} -2U''(q)p \\ 2U'(q)U''(q) - U'''(q)p^2 \end{pmatrix} + \cdots$$

and the symplectic Euler method (explicit in q, implicit in p) gives

$$\begin{pmatrix} \dot q \\ \dot p \end{pmatrix} = \begin{pmatrix} p \\ -U'(q) \end{pmatrix} + \frac{h}{2} \begin{pmatrix} -U'(q) \\ U''(q)p \end{pmatrix} + \frac{h^2}{12} \begin{pmatrix} 2U''(q)p \\ -2U'(q)U''(q) - U'''(q)p^2 \end{pmatrix} + \cdots .$$

In Fig. 1, solutions of the truncated modified differential equation (with $h = 0.3$) are included. The accordance with the numerical solution is striking. For the double well potential $U(q) = \frac{1}{2}(q^2-1)^2$ the stationary points $(q,p) = (\pm 1, 0)$ turn into a spiral source for the modified equation of the explicit Euler method, whereas they remain stable centers for the symplectic Euler method. The modified equation is a Hamiltonian system only for the symplectic Euler method, in which case the Hamiltonian is given by $H_h(p,q) = \frac{1}{2}p^2 + U(q) - \frac{h}{2}U'(q)p + \frac{h^2}{12}\big(U'(q)^2 + U''(q)p^2\big) + \dots$.

Backward Error Analysis - General Situation. Consider an ordinary differential equation $\dot y = f(y)$ and a numerical integrator $y_{n+1} = \Phi_h(y_n)$. If both, the vector field f and the discrete flow Φ_h are sufficiently differentiable, one can find a (truncated) modified equation

$$\dot y = f(y) + h f_2(y) + h^2 f_3(y) + \ldots + h^{N-1} f_N(y),$$

such that

$$\| \Phi_h(y) - \varphi_{N,h}(y) \| \le C(y)\, h^{N+1},$$

where $\varphi_{N,t}$ is the exact flow of the truncated modified equation. This means that the numerical solution after one step, $y_1 = \Phi_h(y_0)$, is very close to the exact solution of the modified equation at time $t = h$ corresponding to the initial value y_0. The constant $C(y)$ depends on the truncation index and on bounds for derivatives of the vector field, but it is independent of h. If the vector field is real analytic and the integrator falls into the class of (partitioned) Runge–Kutta methods then, by choosing N proportional to h^{-1}, one can prove the estimate (see [3] and [7])

$$\| \Phi_h(y) - \varphi_{N,h}(y) \| \le h\, \alpha\, \mathrm{e}^{-\gamma/(\omega h)}$$

with constants that are independent of N and h. Here, $\gamma > 0$ only depends on the numerical integrator and $\omega > 0$ is related to the highest frequency present in the solution of the differential equation.

One of the most important applications of backward error analysis is the long-time energy preservation of symplectic integrators (see [9, Chapter IX]). In fact,

if the vector field is Hamiltonian (i.e., $f(y) = J^{-1}\nabla H(y)$) and if the discrete flow Φ_h is a symplectic transformation, then the modified differential equation is also Hamiltonian with

$$H_h(y) = H(y) + hH_2(y) + h^2H_3(y) + \ldots + h^{N-1}H_N(y).$$

This implies that $H_h(y)$ is exactly preserved along the flow $\varphi_{N,t}$ of the modified equation, and consequently $\|H_h(\Phi_h(y)) - H_h(y)\| \leq c\,h\,\mathrm{e}^{-\gamma/(\omega h)}$. Telescoping summation gives $\|H_h(y_n) - H_h(y_0)\| \leq c\,nh\,\mathrm{e}^{-\gamma/(\omega h)}$ and, since for a method of order r we have $H_j(y) = 0$ for $j = 2,\ldots,r$, this implies that

$$\|H(y_n) - H(y_0)\| \leq C h^r \qquad \text{for} \qquad nh \leq \mathrm{e}^{\gamma/(2\omega h)}.$$

Consequently, for symplectic (Runge–Kutta) methods the Hamiltonian is preserved up to an error of size $\mathcal{O}(h^r)$ on exponentially long time intervals.

3 Perturbed Fast Oscillators – Modulated Fourier Expansions

Whenever applicable, backward error analysis is an excellent tool for getting insight into the long-time behaviour of numerical solutions. The disadvantage is that for situations, where the product of the step size h with the highest frequency ω is not small, it does not give any information. In this section we consider nonlinearly perturbed harmonic oscillators of the form

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

where $\mathbf{q} = (q_0, q_1, \ldots, q_m)$ 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

$$\omega_j \geq \varepsilon^{-1}, \qquad 0 < \varepsilon \ll 1, \qquad j = 1,\ldots,m.$$

This system is Hamiltonian with energy

$$H(\mathbf{q},\dot{\mathbf{q}}) = \frac{1}{2}\sum_{j=0}^{m}\left(\dot{q}_j^{\mathsf{T}}\dot{q}_j + \omega_j^2 q_j^{\mathsf{T}} q_j\right) + U(\mathbf{q}).$$

With the notation Ω for the diagonal matrix with entries ω_j, and ∇U for the vector that collects all $\nabla_j U$, the differential equation can be written as $\ddot{\mathbf{q}} = -\Omega^2\mathbf{q} - \nabla U(\mathbf{q})$.

It turns out that the study of the near energy preservation of numerical integrators requires the consideration of the *oscillatory energy*

$$H_\omega(\mathbf{q},\dot{\mathbf{q}}) = \frac{1}{2}\sum_{j=1}^{m}\left(\dot{q}_j^{\mathsf{T}}\dot{q}_j + \omega_j^2 q_j^{\mathsf{T}} q_j\right),$$

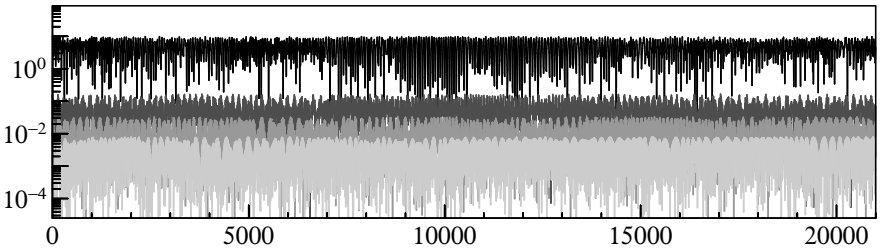


Fig. 2 Numerical energy error of the Störmer–Verlet method applied to a nonlinearly perturbed fast harmonic oscillator as function of time. The step sizes are such that $h\omega = 1.95$ (black) and $h\omega = 1, 0.5, 0.25$ (different grades of gray).

and it is essential to assume that $H_\omega(\mathbf{q}(0), \dot{\mathbf{q}}(0)) \le E$ is bounded independently of ε. The oscillatory energy is then nearly preserved along the analytic solution of the differential equation over long times [6].

Numerical Experiment with the Störmer–Verlet Method. We consider a chain with alternating soft nonlinear and stiff linear springs as described in [9, Section I.5]. It is of the above form with $m = 1$, $d_0 = d_1 = 3$, and has a quartic potential U. For our experiment we choose $\omega_1 = \omega = 50$. As numerical integrator we consider the *Störmer–Verlet method*

$$\begin{aligned} \mathbf{q}_{n+1} - 2\mathbf{q}_n + \mathbf{q}_{n-1} &= h^2\big(-\Omega^2 \mathbf{q}_n - \nabla U(\mathbf{q}_n)\big) \\ 2h\,\dot{\mathbf{q}}_n &= \mathbf{q}_{n+1} - \mathbf{q}_{n-1}, \end{aligned}$$

which is frequently used in molecular dynamics simulations. Considered as a mapping $(\mathbf{q}_n, \dot{\mathbf{q}}_n) \mapsto (\mathbf{q}_{n+1}, \dot{\mathbf{q}}_{n+1})$ it is symplectic, symmetric, and of order 2, and it is perfectly suited for computations requiring low accuracy. Note that stable numerical solutions (for the harmonic oscillator $\ddot{q} + \omega^2 q = 0$) are obtained only under the step size restriction $h\omega < 2$.

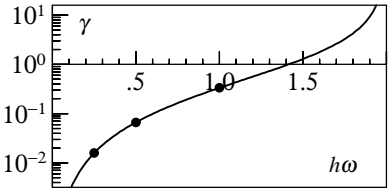
Figure 2 shows the error in the energy of the Störmer–Verlet scheme applied to the above mentioned problem of alternating soft and stiff springs. Even for a very large step size $h = 1.95/\omega$ the error (although very large) remains bounded without any drift. For more reasonable step sizes $h\omega \in \{1, 0.5, 0.25\}$ the error in the energy behaves as expected. There is no drift and the bound on the error decreases when the step size becomes smaller. This excellent long-time behaviour cannot be explained with the techniques of the previous section.

A Theorem on the Numerical Energy Preservation. To simplify the notation we present a result for the case of only one high frequency ($m = 1$), and we denote $\omega = \omega_1$. For the numerical solution, obtained by the Störmer–Verlet method, we assume that the step size satisfies $0 < c_0 \le h\omega \le c_1 < 2$ and that the numerical non-resonance condition

$$\big|\sin(\tfrac{1}{2}kh\widetilde{\omega})\big| \ge c\sqrt{h}, \qquad k = 1, \dots, N$$

Fig. 3 The function $\gamma(h\omega)$ appearing in the modified energies:

$$\gamma(h\omega) = \frac{\frac{1}{4}(h\omega)^2}{1-\frac{1}{4}(h\omega)^2}$$



holds for some $N \geq 2$ and $c > 0$, where $h\widetilde{\omega}$ is defined by the relation $\sin(\frac{1}{2}h\widetilde{\omega}) = \frac{1}{2}h\omega$. The latter condition is in fact a definition of N. We further suppose that the numerical solution stays in a region on which all derivatives of U are bounded. With the modified energies

$$H^*(\mathbf{q},\dot{\mathbf{q}}) = H(\mathbf{q},\dot{\mathbf{q}}) + \tfrac{1}{2}\gamma(h\omega)\|\dot{q}_1\|^2$$
$$H^*_\omega(\mathbf{q},\dot{\mathbf{q}}) = H_\omega(\mathbf{q},\dot{\mathbf{q}}) + \tfrac{1}{2}\gamma(h\omega)\|\dot{q}_1\|^2,$$

where $\gamma(h\omega)$ is given in Fig. 3, it then holds that

$$\begin{aligned} H^*(\mathbf{q}_n,\dot{\mathbf{q}}_n) &= H^*(\mathbf{q}_0,\dot{\mathbf{q}}_0) + \mathscr{O}(h) \\ H^*_\omega(\mathbf{q}_n,\dot{\mathbf{q}}_n) &= H^*_\omega(\mathbf{q}_0,\dot{\mathbf{q}}_0) + \mathscr{O}(h) \end{aligned} \qquad \text{for} \qquad 0 \leq nh \leq h^{-N+1}.$$

The constants symbolised by $\mathscr{O}$ are independent of n, h, ω under the above conditions. Along the numerical solution the expression $\|\dot{q}_1\|^2$ is highly oscillatory and of size $\mathscr{O}(1)$. Moreover, its time average over intervals of length T is nearly constant:

$$\frac{h}{T}\sum_{|jh|\leq T}\|\dot{q}_{n+j,1}\|^2 = \frac{1}{1+\gamma(h\omega)}H^*_\omega(\mathbf{q}_0,\dot{\mathbf{q}}_0) + \mathscr{O}(h).$$

This result is taken from [9, Chapter XIII.8]. The long-time behaviour of Fig. 2 can now be explained. Due to the near preservation of the modified energies, the dominant error term comes from $\frac{1}{2}\gamma(h\omega)\|\dot{q}_1\|^2$. For $h\omega = 1.95$ we have $\frac{1}{2}\gamma(h\omega) \approx 10$, which explains the large (but bounded) energy error for this particular step size. For the step sizes, for which $h\omega \in \{1, 0.5, 0.25\}$, this expression is much smaller (see Fig. 3) and precisely corresponds to the observation of Fig. 2.

Idea of the Proof. A detailed proof of the above result can be found in [9, Chapter XIII.8]. An extension to the multi-frequency case is presented in [4]. We shortly mention here the two main ingredients (for the multi-frequency case):

- *Modifying the frequencies.* For vanishing potential U the Störmer–Verlet discretization reduces to a linear three-term recursion with exact solution $q_{n,j} = c_{j,1}\,\mathrm{e}^{\mathrm{i}\widetilde{\omega}_j nh} + c_{j,2}\,\mathrm{e}^{-\mathrm{i}\widetilde{\omega}_j nh}$, where the modified frequencies $\widetilde{\omega}_j$ are given by

$$1 - \frac{(h\omega)^2}{2} = \cos\bigl(h\widetilde{\omega}\bigr) \qquad \text{or, equivalently,} \qquad \sin\Bigl(\frac{h\widetilde{\omega}}{2}\Bigr) = \frac{h\omega}{2}.$$

The Störmer–Verlet scheme thus becomes the trigonometric integrator

$$\mathbf{q}_{n+1} - 2\cos\bigl(h\widetilde{\Omega}\bigr)\mathbf{q}_n + \mathbf{q}_{n-1} = -h^2\nabla U(\mathbf{q}_n).$$

which is easier to analyse, because the linear part exactly integrates a harmonic oscillator with modified frequencies.

- *Modulated Fourier expansion.* For the problem of this section the Störmer–Verlet method is a nonlinear perturbation of a three-term relation, for which the solution is a linear combination of exponentials $\mathrm{e}^{\pm \mathrm{i}\widetilde{\omega}_j nh}$. It is therefore natural to approximate the numerical solution $\mathbf{q}_n$ of the complete discretisation as a linear combination of products of $\mathrm{e}^{\pm \mathrm{i}\widetilde{\omega}_j nh}$ (called *modulated Fourier expansion*)

$$\mathbf{q}_n = \mathbf{y}(t) + \sum_{\mathbf{k}\in\mathscr{K}} \mathbf{z}^{\mathbf{k}}(t)\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\widetilde{\omega})t} \qquad \text{with} \qquad t = nh.$$

Here, $\mathbf{k} = (k_1,\ldots,k_n)$ is a multi-index, $\widetilde{\omega} = (\widetilde{\omega}_1,\ldots,\widetilde{\omega}_n)$ is the vector of modified high frequencies, $\mathbf{k}\cdot\widetilde{\omega} = k_1\widetilde{\omega}_1 + \ldots + k_n\widetilde{\omega}_n$, and $\mathscr{K}$ is a suitable finite index set. The coefficient functions $\mathbf{y}(t)$ and $\mathbf{z}^{\mathbf{k}}(t)$ are vector-valued with the same dimension and partitioning as $\mathbf{q}_n$, and they are assumed to be smooth. This means that together with all their derivatives they are bounded independently of ε for $0 < \varepsilon \le \varepsilon_0$ so that with this ansatz high oscillations are well separated from the slow motion. The proof is typically in three steps:

  - Construction of the modulation functions $\mathbf{y}(t)$ and $\mathbf{z}^{\mathbf{k}}(t)$ as the solution of a differential-algebraic system (on short intervals of length $\mathscr{O}(1)$).
  - Proof of the existence of formal invariants of the differential-algebraic system, which are close to the total and oscillatory energies (on short intervals).
  - Concatenation of estimates on short intervals to get the near energy preservation of long time intervals.

  Modulated Fourier expansions for the long-term analysis of (analytical and numerical) solutions of highly oscillatory differential equations have been introduced in [8] for the case of a single high frequency ω. The case of several high frequencies satisfying a non-resonance condition is studied in [4]. They are extensively treated in Chapters XIII and XIV of the monograph [9]. Related results for the analytic solution have been obtained in [1], [2] with canonical transformation techniques of Hamiltonian perturbation theory.

4 Fully Nonlinear, Highly Oscillatory Hamiltonian Problems

What happens, when neither backward error analysis can be applied nor the problem can be cast into the form of a perturbed fast oscillator? Let us consider a molecular dynamics model, which consists of an N-body problem

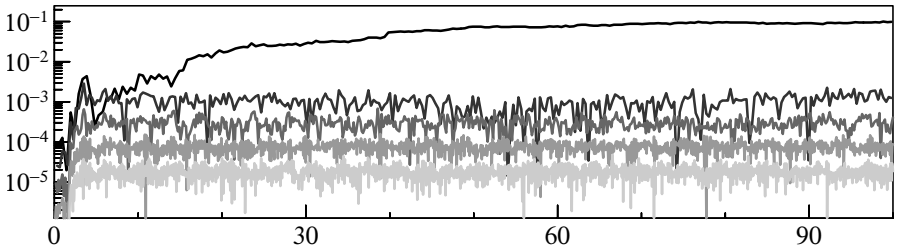


Fig. 4 Numerical energy error of the Störmer–Verlet method applied to the molecular dynamics model of Section 4. The step sizes are $h = 0.06$ (black) and $h = 0.04, 0.02, 0.01, 0.005$ (different grades of gray).

$$\ddot{\mathbf{q}} = -\nabla U(\mathbf{q}), \qquad U(\mathbf{q}) = \sum_{i=2}^{N} \sum_{j=1}^{i-1} V(\|q_i - q_j\|)$$

interacting with the Lennard–Jones potential $V(r) = r^{-12} - 2r^{-6}$. We assume particles $q_i \in \mathbb{R}^2$ in a plane, and an initial configuration consisting of $N = 100$ particles which are at randomly perturbed points of the lattice $\{(l,m)\,; l,m = 1,\ldots,10\}$. Initial velocities are taken to be zero.

Figure 4 shows the error in the energy of the Störmer–Verlet method applied to the N-body problem. For the step size $h = 0.06$ the error increases and is soon out of scale. However, for $h \leq 0.04$ the energy is well preserved, and the error decreases as expected, when the step size becomes smaller.

To check whether this behaviour can be explained either by the backward error analysis of Section 2 or by the technique of modulated Fourier expansions of Section 3 we have to compute the highest frequencies of the solutions in the system. Along the numerical solution we have computed the Hessian matrix $\nabla^2 U(\mathbf{q})$ and its eigenvalues. The dominant eigenvalues are negative and they represent $-\omega^2$, where ω corresponds to the frequencies in the system. In Figure 5 we present the five largest frequencies obtained in this way. We see that they are slightly smaller than the value $\omega = 25$.

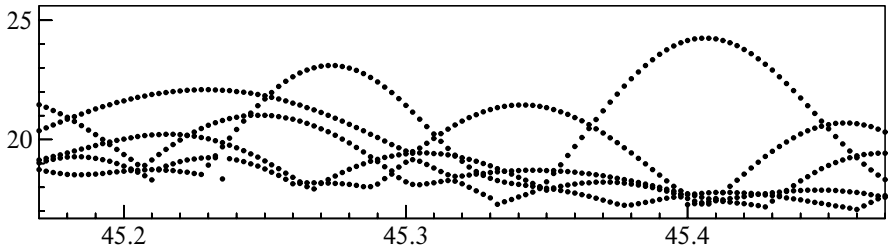


Fig. 5 The five largest frequencies corresponding to the solution of the N-body problem of Section 4 as a function of time.

- Backward error analysis does not give any information on the long-time behaviour. The reason is that for the step sizes used in the experiment of Figure 4 we have $h\omega = 1.5$ (for $h = 0.06$), and $h\omega = 1$ (for $h = 0.04$), which are not small.
- The technique of modulated Fourier expansions requires that the differential equation is a perturbation of a set of harmonic oscillators. This means that the large frequencies of the system have to be nearly constant. Figure 5 shows that this is by far not the case.

An explanation of the good energy preservation observed in Figure 4 is still missing.

Acknowledgements This review is an update of a talk given at the "giornata INdAM" in June 2007 (Pisa). We acknowledge the support over many years of the Fonds National Suisse, Project No. 200020-144313/1.

References

1. 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.
2. _____, *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.
3. G. Benettin and A. Giorgilli, *On the Hamiltonian interpolation of near to the identity symplectic mappings with application to symplectic integration algorithms*, J. Statist. Phys. **74** (1994), 1117–1143.
4. D. Cohen, E. Hairer, and C. Lubich, *Numerical energy conservation for multi-frequency oscillatory differential equations*, BIT **45** (2005), 287–305.
5. R. de Vogelaere, *Methods of integration which preserve the contact transformation property of the Hamiltonian equations*, Tech. report, Dept. Math. Univ. of Notre Dame, Notre Dame, Ind., 1956.
6. L. Gauckler, E. Hairer, and C. Lubich, *Energy separation in oscillatory Hamiltonian systems without any non-resonance condition*, Comm. Math. Phys. **321** (2013), 803–815.
7. E. Hairer and C. Lubich, *The life-span of backward error analysis for numerical integrators*, Numer. Math. **76** (1997), 441–462, Erratum: http://www.unige.ch/math/folks/hairer/.
8. _____, *Long-time energy conservation of numerical methods for oscillatory differential equations*, SIAM J. Numer. Anal. **38** (2000), 414–441.
9. 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.