Long-time energy conservation of numerical methods for oscillatory differential equations

Hairer, Ernst; Lubich, Christian

SIAM J. NUMER. ANAL.
Vol. 38, No. 2, pp. 414–441

LONG-TIME ENERGY CONSERVATION OF NUMERICAL METHODS FOR OSCILLATORY DIFFERENTIAL EQUATIONS*

ERNST HAIRER† AND CHRISTIAN LUBICH‡

Abstract. We consider second-order differential systems where high-frequency oscillations are generated by a linear part. We present a frequency expansion of the solution, and we discuss two invariants of the system that determine the coefficients of the frequency expansion. These invariants are related to the total energy and the oscillatory harmonic energy of the original system.

For the numerical solution we study a class of symmetric methods that discretize the linear part without error. We are interested in the case where the product of the step size with the highest frequency can be large. In the sense of backward error analysis we represent the numerical solution by a frequency expansion where the coefficients are the solution of a modified system. This allows us to prove the near-conservation of the total and the oscillatory energy over very long time intervals.

Key words. oscillatory differential equations, long-time energy conservation, second-order symmetric methods, frequency expansion, backward error analysis, Fermi–Pasta–Ulam problem

AMS subject classifications. 65L05, 65P10

PII. S0036142999353594

1. Introduction.

Long-time near-conservation of the total energy and of adiabatic invariants in numerical solutions to Hamiltonian differential equations is important in a wide range of physical applications from molecular dynamics to nonlinear wave propagation. Backward error analysis [BG94, HaL97, Rei98] has shown that symplectic numerical integrators approximately conserve the total energy and adiabatic invariants over times that are exponentially long in the step size; more precisely, over times of length $\exp(c/h\omega)$ where ω is the highest frequency in the system. Such a result is meaningful only for $h\omega \to 0$, which is often not a practical assumption. For example, in spatially discretized wave equations $h\omega$ is the CFL number, which is not chosen small in actual computations. Recently, in [GSS99, HoL99] new symplectic or symmetric time-stepping methods have been studied which admit second-order error bounds on finite time intervals independently of the frequencies of the dominant linear part of the system. In particular for such "long-time-step methods," the case $h\omega \to 0$ is of no computational interest. The situation is reminiscent of stiff versus nonstiff differential equations, where stiff integrators are not appropriately analyzed by considering only the limit behavior $h \to 0$. In the stiff case, much insight has been gained by studying the behavior of numerical methods on well-chosen, rather simple linear and nonlinear stiff model problems.

As a first step towards an understanding of the numerical energy behavior in Hamiltonian systems when the product of the step size and the highest frequency is not a small quantity, we consider in this article the nonlinear, highly oscillatory model problem

$$\ddot{x} + \Omega^2 x = g(x), \tag{1.1}$$

*Received by the editors March 17, 1999; accepted for publication (in revised form) January 13, 2000; published electronically July 19, 2000.
http://www.siam.org/journals/sinum/38-2/35359.html

†Dept. de mathématiques, Université de Genève, CH–1211 Genève 24, Switzerland (Ernst.Hairer@math.unige.ch).

‡Mathematisches Institut, Universität Tübingen, Auf der Morgenstelle 10, D–72076 Tübingen, Germany (lubich@na.uni-tuebingen.de).

where

$$\Omega = \begin{pmatrix} 0 & 0 \\ 0 & \omega I \end{pmatrix}, \qquad \omega \gg 1 \tag{1.2}$$

(with blocks of arbitrary dimensions), and where the nonlinearity

$$g(x) = -\nabla U(x) \tag{1.3}$$

has a Lipschitz constant bounded independently of ω. This situation arises in the celebrated Fermi–Pasta–Ulam model, for which we will present numerical experiments. Clearly, in the model problem (1.1) we take strong restrictions in that the high frequencies are confined to the linear part of the problem, and that the linear part has a single high frequency. The diagonal form of Ω is not essential, since the numerical methods are invariant under a diagonalization of the matrix.

We study the long-time energy behavior of a class of symmetric numerical methods which are used with step sizes h such that the product $h\omega$ is bounded away from zero and can be arbitrarily large. The methods integrate the linear part of (1.1) exactly and reduce to the Störmer/Verlet method for $\omega = 0$. The class includes the methods of [GSS99, HoL99]. Classical symmetric methods such as the Störmer/Verlet method, the trapezoidal rule, or Numerov's method are not considered in this article. However, using the results of the present paper, their energy behavior on (1.1) for $h\omega$ in the range of linear stability is analyzed in [HaL99].

Our approach to the near-conservation of the energy is based on a frequency expansion of the solution $x(t)$ of (1.1),

$$x(t) = y(t) + \sum_{k \neq 0} e^{ik\omega t} z^k(t), \tag{1.4}$$

which is an asymptotic series where the coefficient functions $y(t)$ and $z^k(t)$ together with their derivatives are bounded independently of ω. It turns out that the system determining the coefficient functions has two (formally exact) invariants. One of these is close to the total energy

$$H(x, \dot{x}) = \tfrac{1}{2}(|\dot{x}_1|^2 + |\dot{x}_2|^2) + \tfrac{1}{2}\omega^2 |x_2|^2 + U(x), \tag{1.5}$$

where $x = (x_1, x_2)$ according to the partitioning of Ω. The other invariant is close to

$$I(x, \dot{x}) = \tfrac{1}{2}|\dot{x}_2|^2 + \tfrac{1}{2}\omega^2 |x_2|^2, \tag{1.6}$$

which represents the oscillatory energy of the system.

For the numerical solution we derive a similar frequency expansion which is valid on grid points $t = nh$. Under a nonresonance assumption on $h\omega$,

$$|\sin(\tfrac{1}{2}kh\omega)| \geq c\sqrt{h} \quad \text{for} \quad k = 1, \ldots, N \quad (N \geq 2), \tag{1.7}$$

the equations determining the coefficient functions have a similar structure to those of the continuous problem. This allows us to obtain two almost-invariants close to H and I, and rigorous estimates for the near-conservation of the total and the oscillatory energy over time intervals of size $C_N h^{-N}$. The only restriction on N comes from the above nonresonance condition. The analysis uses only the symmetry of the methods and does not require symplecticness.

In section 2 we describe the numerical methods and we present numerical experiments with the Fermi–Pasta–Ulam problem. These experiments illustrate the long-time conservation of the total and the oscillatory energy in nonresonance situations, which will later be completely explained by the theory. We also show the energy behavior of the methods near resonances. This behavior depends strongly on properties of the filter functions that determine the numerical method. We identify conditions that yield satisfactory energy conservation near resonances and the correct energy exchange between highly oscillatory components.

Section 3 gives a complete analysis of the two-dimensional linear case of (1.1) over the whole range of nonresonant, near-resonant, and exactly resonant cases. This already gives much insight into conditions determining the energy conservation in the general situation.

The frequency expansion of the analytical solution of (1.1) is introduced in section 4, that of the numerical solution in section 5. The numerical invariants are derived in section 6. The main result on the numerical long-time conservation of energy for (1.1) is formulated and proved in section 7.

2. Numerical methods and numerical experiments. In this section we present the numerical methods and we illustrate the main results of this paper with the Fermi–Pasta–Ulam problem.

2.1. The discretization. We consider the differential equation (1.1), where Ω^2 is a symmetric and positive semidefinite (not necessarily diagonal) real matrix, and we assume that initial values x_0 and $\dot{x}_0$ are given at $t_0 = 0$. By the variation-of-constants formula, the exact solution of (1.1) satisfies

$$\begin{pmatrix} x(t) \\ \dot{x}(t) \end{pmatrix} = \begin{pmatrix} \cos t\Omega & \Omega^{-1} \sin t\Omega \\ -\Omega \sin t\Omega & \cos t\Omega \end{pmatrix} \begin{pmatrix} x_0 \\ \dot{x}_0 \end{pmatrix} + \int_0^t \begin{pmatrix} \Omega^{-1} \sin(t-s)\Omega \\ \cos(t-s)\Omega \end{pmatrix} g\big(x(s)\big)\, ds. \tag{2.1}$$

(Observe that $\Omega^{-1} \sin t\Omega$ is well-defined also for singular Ω.) It is therefore natural to consider, for a fixed step size h, the explicit discretization

$$x_{n+1} = \cos h\Omega\, x_n + \Omega^{-1} \sin h\Omega\, \dot{x}_n + \tfrac{1}{2} h^2 \Psi\, g_n, \tag{2.2}$$

$$\dot{x}_{n+1} = -\Omega \sin h\Omega\, x_n + \cos h\Omega\, \dot{x}_n + \tfrac{1}{2} h \big(\Psi_0\, g_n + \Psi_1\, g_{n+1}\big), \tag{2.3}$$

where $g_n = g(\Phi x_n)$ and $\Phi = \phi(h\Omega)$, $\Psi = \psi(h\Omega)$, $\Psi_0 = \psi_0(h\Omega)$, $\Psi_1 = \psi_1(h\Omega)$ with real functions $\phi(\xi)$, $\psi(\xi)$, $\psi_0(\xi)$, $\psi_1(\xi)$ depending smoothly on ξ^2. For $g(x) \equiv 0$ this method integrates the problem (1.1) without error.

If $\phi(0) = \psi(0) = \psi_0(0) = \psi_1(0) = 1$, the method is consistent of order 2. For fixed Ω and for $h \to 0$, second-order convergence follows from classical results. In this article we are mainly interested in the situation where $h\Omega$ can take large values.

For long-time integrations, symmetric and/or symplectic methods are expected to have favorable properties. By exchanging $n \leftrightarrow n+1$ and $h \leftrightarrow -h$ in (2.2)-(2.3), it is seen that the method is symmetric for all $g(x)$ if and only if

$$\psi(\xi) = \operatorname{sinc}\xi \cdot \psi_1(\xi), \qquad \psi_0(\xi) = \cos\xi \cdot \psi_1(\xi) \tag{2.4}$$

(where $\operatorname{sinc}\xi = \sin\xi/\xi$). It can be shown by direct verification that (2.2)–(2.3) is a symplectic discretization if, in addition to (2.4), $\phi(\xi) = \psi_1(\xi)$ also holds. This condition will not be required for the analysis of our paper.

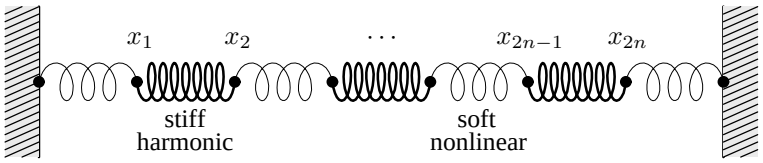

FIG. 1. *Alternating soft and stiff springs.*

Since the nonlinearity $g(x)$ in (1.1) does not depend on $\dot{x}$, we can eliminate $\dot{x}_n$ in (2.2) with the help of (2.3). In the case of a symmetric discretization we thus get the two-step recurrence

$$x_{n+1} - 2\cos h\Omega\, x_n + x_{n-1} = h^2 \Psi\, g_n. \tag{2.5}$$

The starting value x_1 is obtained from (2.2) with $n = 0$. For the case $\Omega = 0$ we recognize the well-known Störmer method.

Methods of the type (2.5) or (2.2)–(2.3) have been proposed and studied by several authors. Gautschi [Gau61] suggests taking $\psi(\xi) = \operatorname{sinc}^2(\xi/2)$. With this choice, (1.1) is integrated exactly even for $g(x) = \text{const}$. Deuflhard [Deu79] discretizes the integral in (2.1) by the trapezoidal rule and thus arrives at (2.5) with $\psi(\xi) = \operatorname{sinc}\xi$ and $\phi(\xi) = 1$. More recently, García-Archilla, Sanz-Serna, and Skeel [GSS99] introduce a function $\phi(\xi)$ in the argument of g, and they consider the case where the method is symplectic, so that $\psi(\xi) = \operatorname{sinc}\xi \cdot \phi(\xi)$. Hochbruck and Lubich [HoL99] consider $\psi(\xi) = \operatorname{sinc}^2(\xi/2)$ and general $\phi(\xi)$. The papers [GSS99] and [HoL99] derive error bounds on finite time intervals which are independent of ω and of the smoothness of the solution.

In this article, we consider general functions ϕ, ψ with $\phi(0) = 1$, $\psi(0) = 1$, which have no zeros except possibly at integral multiples of π. Since we are interested in the energy conservation of the numerical solution, we also need an approximation to the derivative if we use the two-term recurrence relation (2.5). This can be obtained by the relation (2.3) or, in the case of a symmetric method, also by the formula

$$x_{n+1} - x_{n-1} = 2h \operatorname{sinc} h\Omega\, \dot{x}_n. \tag{2.6}$$

This is possible if $h\omega$ is not a nonzero integral multiple of π. We obtain (2.6) by subtracting (2.5) from twice the formula of (2.2). For a symmetric method we obtain a formula for $\dot{x}_{n-1}$ by exchanging $n \leftrightarrow n+1$ and $h \leftrightarrow -h$ in (2.3). Subtracting the resulting formula from (2.3), we obtain the two-step recurrence

$$\dot{x}_{n+1} - \dot{x}_{n-1} = -2\Omega \sin h\Omega\, x_n + \tfrac{1}{2}h\big(\Psi_1 g_{n+1} + 2\Psi_0 g_n + \Psi_1 g_{n-1}\big). \tag{2.7}$$

Formulas (2.5) and (2.7) give a symmetric two-step method even if (2.4) is not satisfied. If $\psi(\xi) = \operatorname{sinc}^2(\xi/2)$ and $\psi_0(\xi) + \psi_1(\xi) = 2\operatorname{sinc}\xi$, then this method is exact for $g(x) = \text{const}$. The choice $\psi_1(\xi) = 0$ has been considered in [HoL99].

2.2. Experiments with the Fermi–Pasta–Ulam problem. We consider a chain of springs, where soft nonlinear springs alternate with stiff harmonic springs (see [GGMV92] and Figure 1). The variables $x_1, \ldots, x_{2n}$ (and $x_0 = 0$, $x_{2n+1} = 0$) stand for the displacements of end-points of the springs. The movement is described by a Hamiltonian system with

$$H(x, \dot{x}) = \frac{1}{2}\sum_{i=1}^{n}\big(\dot{x}_{2i-1}^2 + \dot{x}_{2i}^2\big) + \frac{K}{2}\sum_{i=1}^{n}(x_{2i} - x_{2i-1})^2 + \sum_{i=0}^{n}(x_{2i+1} - x_{2i})^4.$$

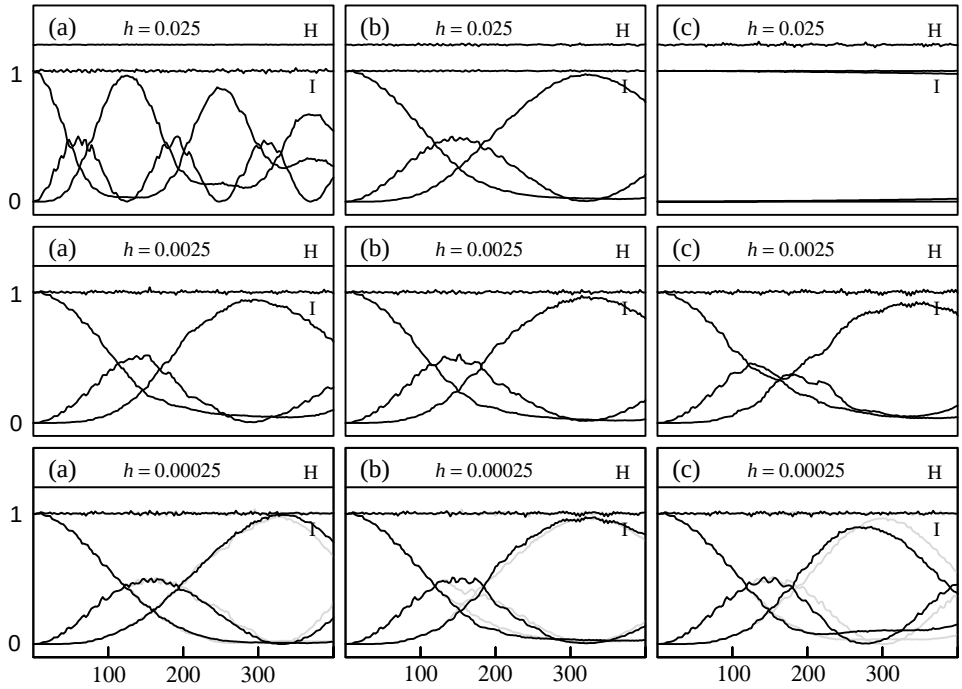

FIG. 2. *Energy exchange of stiff springs.*

Using the symplectic change of variables $u_i = (x_{2i}+x_{2i-1})/\sqrt{2}$, $v_i = (x_{2i}-x_{2i-1})/\sqrt{2}$, we get a new Hamiltonian system with

$$H(u,v,\dot u,\dot v) = \frac{1}{2}\sum_{i=1}^{n}\big(\dot u_i^2+\dot v_i^2\big) + \frac{\omega^2}{2}\sum_{i=1}^{n} v_i^2 + \frac{1}{4}\sum_{i=0}^{n}(u_{i+1}-v_{i+1}-u_i-v_i)^4, \tag{2.8}$$

where $u_0 = v_0 = u_{n+1} = v_{n+1} = 0$ and $\omega^2 = 2K$. This is exactly of the form (1.1).

For our numerical experiments we consider the case $n = 3$ (as shown in Figure 1) with $\omega = 100$. As initial values we take

$$u_1(0) = 1, \quad \dot u_1(0) = 1, \quad v_1(0) = \omega^{-1}, \quad \dot v_1(0) = 1,$$

and zero for the remaining initial values. We apply the method (2.2)–(2.3) with the following data:

(a) $\psi(\xi) = \operatorname{sinc}^2(\frac{1}{2}\xi),\ \phi(\xi) = 1$ [Gau61],
(b) $\psi(\xi) = \operatorname{sinc}\xi,\ \phi(\xi) = 1$ [Deu79],
(c) $\psi(\xi) = \operatorname{sinc}^2\xi,\ \phi(\xi) = \operatorname{sinc}\xi$ [GSS99]

with $\psi_0(\xi)$ and $\psi_1(\xi)$ given by (2.4).

We study the total energy (2.8) and the oscillatory energy

$$I = I_1 + I_2 + I_3 \qquad \text{with} \qquad I_j(v_j,\dot v_j) = \tfrac{1}{2}\big(\dot v_j^2 + \omega^2 v_j^2\big)$$

along the numerical solution on the interval $0 \le t \le 400$. With the chosen initial values we have $H = 2.000300005$ and $I_1 = 1$, $I_2 = I_3 = 0$ at $t = 0$. In Figure 2 we

have plotted, for three different step sizes and for all three methods, the numerical values for I_1, I_2, I_3, I and $H - 0.8$. We see that an exchange of energy takes place, going from the first stiff spring with energy I_1 to the second stiff spring and later to the third one. For the smallest step size we have also plotted in gray the numerical values for perturbed initial values obtained by adding 10^{-8} to $u_1(0)$, $\dot{u}_1(0)$, and $\dot{v}_1(0)$. This illustrates that the solution is very sensitive to perturbations.

In all cases we see that H and I are well preserved over the whole interval, even for step sizes where the numerical solution is completely wrong. Further experiments have shown that such a preservation holds for much longer intervals. (We tested up to $t = 100,000$.) An explanation of this phenomenon is the main objective of this paper.

2.3. Numerical experiments in near-resonant situations. When the product of the step size and the frequency $h\omega$ is close to a multiple of π, then the different methods show widely different behavior. Energy is conserved only for some choices of ψ and ϕ. Satisfactory numerical behavior also in near-resonance situations is obtained if the numerical method satisfies the following additional conditions:

$$|\psi(h\omega)| \le C_1 \operatorname{sinc}^2(\tfrac{1}{2}h\omega), \tag{2.9}$$

$$|\psi(h\omega)| \le C_2 \, |\operatorname{sinc}(h\omega)|, \tag{2.10}$$

$$|\psi(h\omega)\phi(h\omega)| \le C_3 \, \operatorname{sinc}^2(h\omega). \tag{2.11}$$

These conditions yield long-time energy conservation for all values of $h\omega$ with the exception of $h\omega$ in intervals of width $O(h)$ near integral multiples of 2π. The total energy appears to be conserved uniformly for arbitrary values of $h\omega$ if

$$\psi(h\omega) = \phi(h\omega) \operatorname{sinc}^2(h\omega) \; . \tag{2.12}$$

The necessity of these conditions is seen from an analysis of the linear case, which is given in section 3.

When $h\omega$ is close to $2m\pi$ with a positive integer m, the condition (2.9) requires a double zero of ψ at $2m\pi$. Similarly, for $h\omega$ close to an odd multiple of π, condition (2.10) there requires a simple zero of ψ. For the choice $\psi(\xi) = \operatorname{sinc}^2(\frac{1}{2}\xi)$ [Gau61, HoL99] the condition (2.9) is obviously satisfied for all values of $h\omega$, but (2.10) is violated near odd multiples of π. For $\psi(\xi) = \operatorname{sinc}\xi$ [Deu79], condition (2.10) is trivially satisfied for all $h\omega$, but condition (2.9) fails near even multiples of π. The choice $\psi(\xi) = \phi(\xi)\operatorname{sinc}\xi$ with $\phi(\xi) = \operatorname{sinc}\xi$ [GSS99] satisfies the three conditions (2.9)–(2.11) for all $h\omega$. Condition (2.12) is not satisfied by any of the methods previously proposed in the literature.

Let us illustrate the effect of the conditions (2.9), (2.10), and (2.11) on the numerical solution when $h\omega$ is close to a multiple of π. We consider the Fermi–Pasta–Ulam problem of section 2.2 with the same initial values, and we apply six different methods. Their characteristics are given in Table 1. The sign $\surd$ indicates that the corresponding condition on ψ and ϕ is satisfied. If a condition is not satisfied for all values of $h\omega$, we give the values close to which it is violated. For each of the methods (B), (C), (D) only one of the conditions (2.9)–(2.11) is not fulfilled.

In Figure 3 we show the errors of the Hamiltonian over the interval $[0, 1,000]$. We have used the step size $h = 0.01$ and ω such that $h\omega = 1.0000001 \cdot \pi$. Method (C) gives a maximal error of size $396,497$, because $\psi_1(\xi)$ given by (2.4) has a singularity at $\xi = \pi$. Methods (A) and (D) show a clear drift from the constant value of the

TABLE 1
Methods used for the numerical experiments of section 2.3.

Method	$\psi(\xi)$	$\phi(\xi)$	(2.9)	(2.10)	(2.11)
A	$\operatorname{sinc}(\xi)$	1	$2k\pi$	$\surd$	$k\pi$
B	$\operatorname{sinc}(\xi)$	$\operatorname{sinc}(\xi)$	$2k\pi$	$\surd$	$\surd$
C	$\operatorname{sinc}^2(\frac{1}{2}\xi)$	$\operatorname{sinc}^2(\xi)$	$\surd$	$(2k+1)\pi$	$\surd$
D	$\phi(\xi)\operatorname{sinc}(\xi)$	$\operatorname{sinc}(\frac{1}{2}\xi)$	$\surd$	$\surd$	$(2k+1)\pi$
E	$\phi(\xi)\operatorname{sinc}(\xi)$	$\operatorname{sinc}(\xi)$	$\surd$	$\surd$	$\surd$
F	$\operatorname{sinc}^2(\xi)$	1	$\surd$	$\surd$	$\surd$

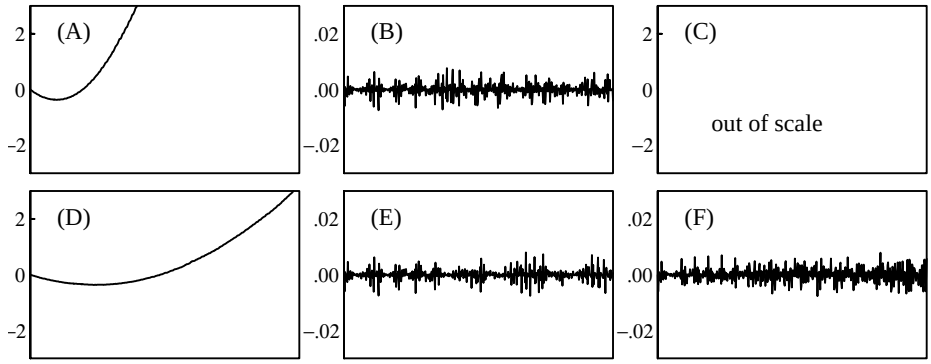

FIG. 3. *Energy conservation of different methods for* $h\omega = 1.0000001 \cdot \pi$ *and* $h = 0.01$.

Hamiltonian. Only methods (B), (E), and (F), for which all three conditions are satisfied close to π, conserve the Hamiltonian very well. For the pictures corresponding to these three methods, we have changed the scale so that the small oscillations become visible.

Figure 4 shows the same experiment, where this time ω is chosen such that $h\omega = 2.0000001 \cdot \pi$. For both situations, we get the same qualitative behavior when we plot the oscillatory energy instead of the Hamiltonian. The results of these experiments confirm that the conditions (2.9), (2.10), (2.11) cannot be omitted if we are interested in long-time energy estimates that hold uniformly in $h\omega$.

In Figure 5 the numerical results of the nonresonant case $h\omega = 2.5 \cdot \pi$ are included. All methods give satisfactory results. The most accurate results are obtained by the method (C).

In the upper pictures of Figure 6 we plot the maximal errors in the Hamiltonian as a function of $h\omega$, and we take step sizes $h = 0.1$, $h = 0.05$, and $h = 0.025$. The picture to the right corresponds to method (E) of Table 1. The picture to the left is obtained with method (F) which satisfies (2.12). Uniform convergence of the error can be observed only in this case. The lower pictures of Figure 6 show the analogue for the deviations in the oscillatory energy. Close to integral multiples of 2π this deviation is large for both methods. The same phenomenon can be observed already for linear problems (see Figure 8), for which a complete analysis is given in section 3.

2.4. Energy exchange. The energy exchange between stiff components takes place on time intervals of length $O(\omega)$. In Figure 2, this is reproduced qualitatively

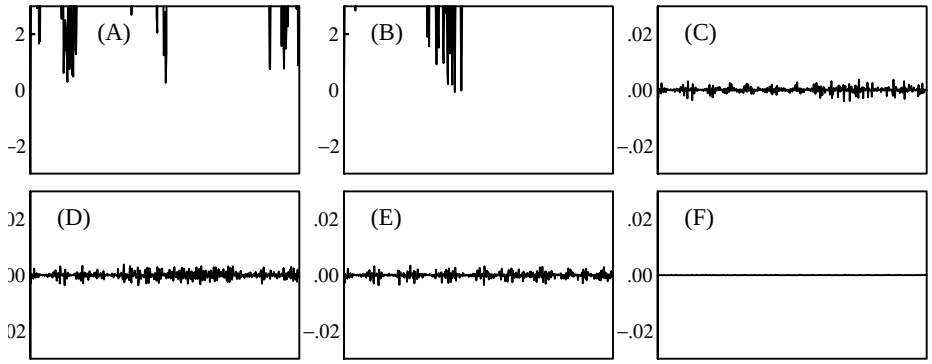

FIG. 4. *Energy conservation of different methods for* $h\omega = 2.0000001 \cdot \pi$ *and* $h = 0.01$.

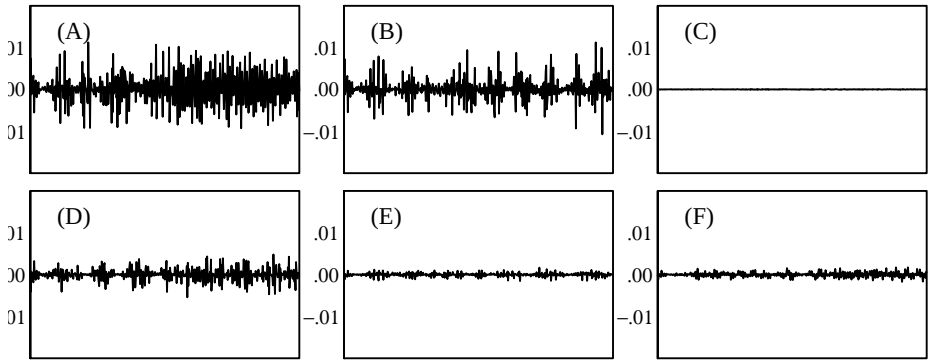

FIG. 5. *Energy conservation of different methods for* $h\omega = 2.5 \cdot \pi$ *(nonresonant case) and* $h = 0.01$.

correctly for large $h\omega$ only in the case where $\psi(\xi) = \operatorname{sinc}\xi$, $\phi(\xi) = 1$. The numerical frequency expansion of section 5 shows that the condition

$$\psi(h\omega)\,\phi(h\omega) = \operatorname{sinc}(h\omega) \tag{2.13}$$

is needed for the approximation of the energy exchange between stiff components when $h\omega$ is bounded away from zero. (Compare the equations for the z_2 component in (4.8) and (5.8).) This is a severe condition which excludes all methods considered so far with the exception of the above-mentioned method $\psi(\xi) = \operatorname{sinc}\xi$, $\phi(\xi) = 1$. On the other hand, we have seen that this method has rather poor energy conservation properties. We therefore extend the class of methods (2.5) to

$$x_{n+1} - 2\cos h\Omega\, x_n + x_{n-1} = h^2 \sum_{k=1}^{K} \Psi_k\, g(\Phi_k x_n), \tag{2.14}$$

where $\Psi_k = \psi_k(h\Omega)$, $\Phi_k = \phi_k(h\Omega)$. For consistency, the functions ψ_k, ϕ_k must satisfy

$$\sum_{k=1}^{K} \psi_k(0) = 1\ , \qquad \phi_k(0) = 1, \quad (k = 1, \ldots, K).$$

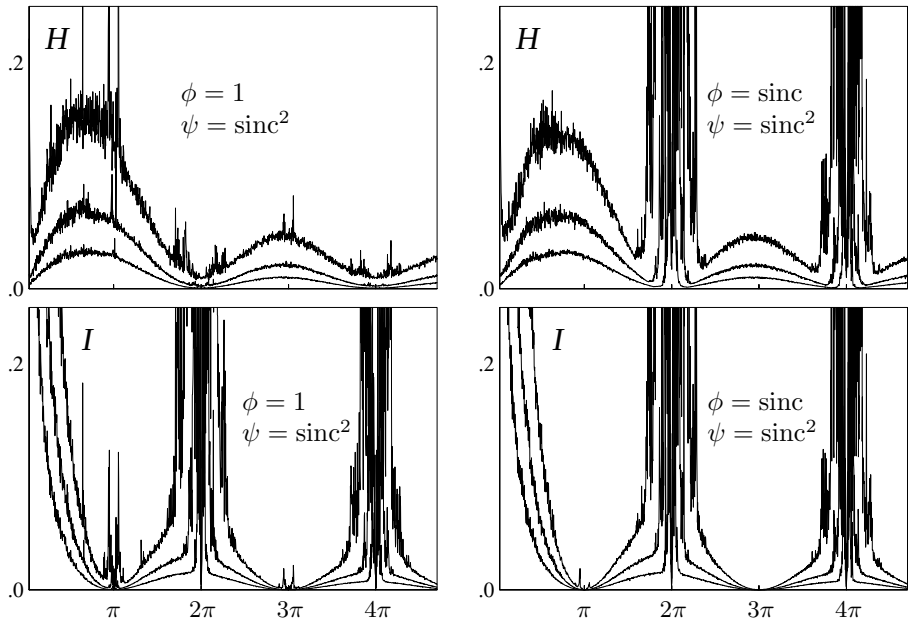

FIG. 6. *Error in the total and oscillatory energies as a function of $h\omega$ for the FPU problem.*

Conditions (2.9) and (2.10) are now needed for $\psi = \sum_k \psi_k$, and condition (2.13) is replaced with

$$\sum_{k=1}^{K} \psi_k(h\omega)\phi_k(h\omega) = \mathrm{sinc}\,(h\omega). \tag{2.15}$$

For example, the method

$$\begin{aligned} x_{n+1} - 2\cos h\Omega\, x_n + x_{n-1} &= h^2\,\mathrm{sinc}^{\,2}(h\Omega)\, g(x_n) \\ &+ h^2\,\mathrm{sinc}\,(h\Omega)\big(g(x_n) - g(\mathrm{sinc}\,(h\Omega)x_n)\big) \end{aligned} \tag{2.16}$$

with (2.6), or equivalently in one-step form

$$x_{n+1} = \cos h\Omega\, x_n + \Omega^{-1}\sin h\Omega\, \dot{x}_n + \tfrac{1}{2}h^2\,\mathrm{sinc}\,(h\Omega)\,\widetilde{g}_n, \tag{2.17}$$

$$\dot{x}_{n+1} = -\Omega\sin h\Omega\, x_n + \cos h\Omega\, \dot{x}_n + \tfrac{1}{2}h\big(\cos(h\Omega)\,\widetilde{g}_n + \widetilde{g}_{n+1}\big), \tag{2.18}$$

where $\widetilde{g}_n = g(x_n) + \mathrm{sinc}\,(h\Omega)g(x_n) - g(\mathrm{sinc}\,(h\Omega)x_n)$, shows the correct energy exchange for large $h\omega$, as for method (b) in Figure 2. In Figure 7, we compare the energy conservation for the method (2.5) with $\psi(\xi) = \mathrm{sinc}\,\xi$, $\phi(\xi) = 1$ (left), and the method (2.16) (right) for the same step sizes as in Figure 6. For (2.16), the total and oscillatory energies are well conserved over long times except for $h\omega$ in intervals of length $O(h)$ around integral multiples of π.

For ease of presentation, the following analysis will be done for methods of class (2.5), but the arguments extend in an obvious way to the class (2.14).

3. Long-time energy conservation for linear problems. We start our analysis with the case where $U(x) = \frac{1}{2}x^T Ax$ with a two-dimensional symmetric matrix A satisfying $a_{11} > 0$, so that

$$g(x) = -\nabla U(x) = -Ax$$

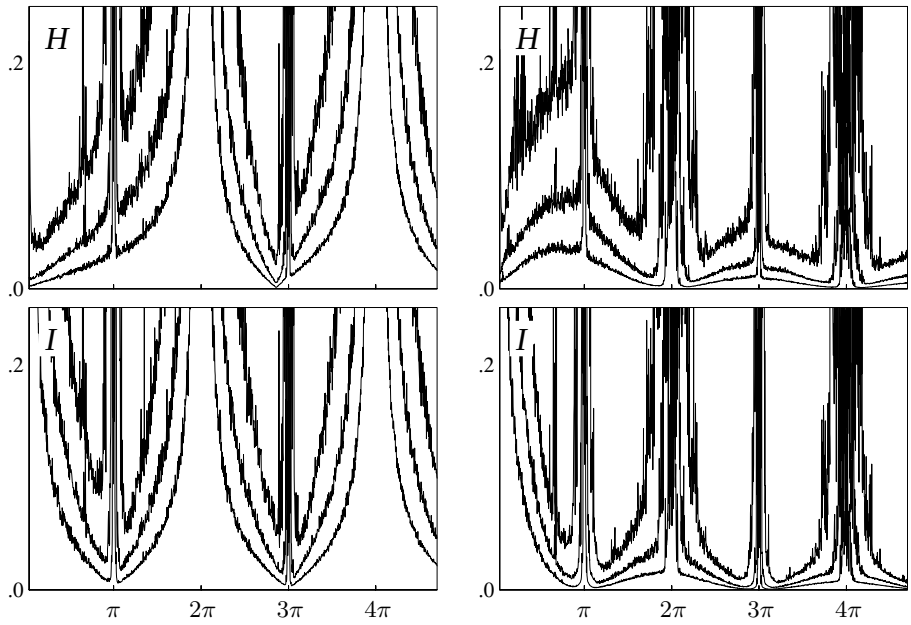

FIG. 7. *Same experiment as in Figure* 6 *with method* (2.5), $\psi = \operatorname{sinc}$, $\phi = 1$ *(left pictures), and with method* (2.16) *(right pictures).*

is linear. This already gives a lot of insight and illustrates the importance of the conditions (2.9)–(2.12). In this situation, the differential equation (1.1) becomes

$$\ddot{x} + \Omega^2 x + Ax = 0. \tag{3.1}$$

The total energy H given by (1.5) is an invariant of the system. In the following we assume that $H\big(x(0), \dot{x}(0)\big)$ is bounded uniformly in ω. This requires $x_2(0) = O(\omega^{-1})$.

3.1. Analytical solution. The exact solution of (3.1) is given by

$$x(t) = (ae^{i\mu_1 t} + \overline{a}e^{-i\mu_1 t}) \begin{pmatrix} 1 \\ O(\omega^{-2}) \end{pmatrix} + (be^{i\mu_2 t} + \overline{b}e^{-i\mu_2 t}) \begin{pmatrix} O(\omega^{-2}) \\ 1 \end{pmatrix},$$

where μ_i^2 are the eigenvalues of $\Omega^2 + A$, so that

$$\mu_1 = \pm\sqrt{a_{11}} + O(\omega^{-2}), \qquad \mu_2 = \pm\omega + O(\omega^{-1}).$$

For given initial values $x(0)$, $\dot{x}(0)$ satisfying $x_2(0) = O(\omega^{-1})$, we obtain $a = O(1)$ and $b = O(\omega^{-1})$. Consequently, we have

$$\begin{aligned} x_2(t) &= be^{i\mu_2 t} + \overline{b}e^{-i\mu_2 t} + O(\omega^{-2}), \\ \dot{x}_2(t) &= i\omega\big(be^{i\mu_2 t} - \overline{b}e^{-i\mu_2 t}\big) + O(\omega^{-2}). \end{aligned} \tag{3.2}$$

This implies $\omega^2|x_2(t)|^2 + |\dot{x}_2(t)|^2 = 4\omega^2|b|^2 + O(\omega^{-1})$, so that the quantity

$$I(x, \dot{x}) = \tfrac{1}{2}|\dot{x}_2|^2 + \tfrac{1}{2}\omega^2|x_2|^2$$

remains $O(\omega^{-1})$-close to the constant value $I\big(x(0), \dot{x}(0)\big)$ for all times t.

3.2. Numerical solution. We search for functions $\widehat{x}(t) = e^{i\mu t}v$, such that $x_n := \widehat{x}(nh)$ satisfies the numerical scheme (2.5) with $g_n = -A\Phi x_n$. This implies

$$\big(e^{i\mu h} - 2\cos h\Omega + e^{-i\mu h}\big)v \;=\; -h^2\Psi A\Phi v,$$

so that $\lambda = \cos(h\mu)$ has to be an eigenvalue of

$$\begin{pmatrix} 1 - \frac{1}{2}h^2 a_{11} & -\frac{1}{2}h^2 a_{12}\phi \\ -\frac{1}{2}h^2 a_{21}\psi & \cos(h\omega) - \frac{1}{2}h^2 a_{22}\phi\psi \end{pmatrix}, \tag{3.3}$$

and v a corresponding eigenvector. In this section we use the short notation $\psi = \psi(h\omega)$ and $\phi = \phi(h\omega)$. Since the off-diagonal elements of (3.3) are small, the eigenvalues $\lambda_1 = \cos(h\mu_1)$ and $\lambda_2 = \cos(h\mu_2)$ are close to $\alpha_1 = 1 - \frac{1}{2}h^2 a_{11}$ and $\alpha_2 = \cos(h\omega) - \frac{1}{2}h^2 a_{22}\phi\psi$, respectively. The corresponding eigenvectors are

$$v_1 = \begin{pmatrix} 1 \\ \gamma\psi \end{pmatrix}, \qquad v_2 = \begin{pmatrix} -\gamma\phi \\ 1 \end{pmatrix},$$

where $\gamma = \frac{1}{2}h^2 a_{21}/(\alpha_2 - \lambda_1)$. For $\lambda_1 \neq \lambda_2$ the general solution can thus be written as

$$\widehat{x}(t) = (ae^{i\mu_1 t} + \overline{a}e^{-i\mu_1 t})v_1 + (be^{i\mu_2 t} + \overline{b}e^{-i\mu_2 t})v_2, \tag{3.4}$$

where the complex coefficients are computed from the initial values $\widehat{x}(0) = x_0$, $\widehat{x}(h) = x_1$ with x_1 obtained from (2.2).

Inserting $x_n = \widehat{x}(nh)$ from (3.4) into (2.6) gives $\dot{x}_n = \widehat{x}\,'(nh)$, where

$$h\,\mathrm{sinc}\,(h\Omega)\,\widehat{x}\,'(t) \;=\; i\sin(\mu_1 h)(ae^{i\mu_1 t} - \overline{a}e^{-i\mu_1 t})v_1 + i\sin(\mu_2 h)(be^{i\mu_2 t} - \overline{b}e^{-i\mu_2 t})v_2.$$

To study the long-time near-conservation of $H\big(\widehat{x}(t), \widehat{x}\,'(t)\big)$ and $I\big(\widehat{x}(t), \widehat{x}\,'(t)\big)$ we distinguish two cases.

Case I: *Well-separated eigenvalues.* We assume that one of the conditions

$$1 - \cos(h\omega) \leq \tfrac{1}{4}h^2 a_{11} \qquad \text{or} \tag{3.5}$$

$$1 - \cos(h\omega) \geq h^2 a_{11} \tag{3.6}$$

is satisfied. This covers nearly all choices of $h\omega$. Only values in intervals of length $O(h)$ are excluded.

THEOREM 3.1. *Consider the numerical method* (2.2)–(2.3) *applied to* (3.1) *with* $a_{11} > 0$ *and* $x_2(0) = O(\omega^{-1})$. *Under one of the restrictions* (3.5) *or* (3.6) *on the step size, and under the conditions* (2.9), (2.10), (2.11) *on the numerical method, the energies* H *and* I *of* (1.5) *and* (1.6) *along the numerical solution satisfy*

$$H(x_n, \dot{x}_n) = H(x_0, \dot{x}_0) + O(h^2) + O(\omega^{-1}), \tag{3.7}$$

$$I(x_n, \dot{x}_n) = I(x_0, \dot{x}_0) + O(\omega^{-1}) \tag{3.8}$$

for all $n \geq 0$. *The constants symbolized by* $O(\cdot)$ *are independent of* ω, h, *and* n.

Proof. The characteristic polynomial of the matrix (3.3) is $p(\lambda) = (\lambda - \alpha_1)(\lambda - \alpha_2) - \frac{1}{4}h^4 a_{12}a_{21}\phi\psi$. Its zeros can be computed explicitly. Each of the conditions (3.5) or (3.6) together with (2.11) implies that

$$\lambda_i - \alpha_i = O(h^2\phi\psi) \qquad \text{and} \qquad \lambda_i - \alpha_i = O(h^2\omega^{-2})$$

for both eigenvalues λ_i of (3.3). Hence, for sufficiently small h, λ_1 and λ_2 are real and in the interval $[-1,1]$. The angles μ_i, defined by $\lambda_i = \cos(h\mu_i)$, are therefore also real and satisfy

$$\mu_1 = \pm\sqrt{a_{11}} + O(h^2), \qquad \mu_2 = \pm\omega + O\big(\omega^{-1}\phi\psi/\mathrm{sinc}\,(h\omega)\,\big).$$

We next estimate the coefficients $a, \overline{a}, b, \overline{b}$ in (3.4). Since the μ_i and v_i are real, the coefficients $\overline{a}, \overline{b}$ are the complex conjugates of a, b. In both situations, (3.5) and (3.6), we have $\gamma = O(1)$ and, under the assumption (2.9), we further have $\gamma\psi = O(\omega^{-2})$. Consequently, the condition $\widehat{x}(0) = x_0$ yields $a+\overline{a} = O(1)$ and $b+\overline{b} = O(\omega^{-1})$. Instead of using $\widehat{x}(h) = x_1$, it is more convenient to work with $\widehat{x}(h) - \cos h\Omega\,\widehat{x}(0) = x_1 - \cos h\Omega\, x_0$. From the estimates $\gamma\psi = O(\omega^{-1}\mathrm{sinc}\,h\omega)$, $\gamma\psi(1-\cos h\omega) = O(h^2\mathrm{sinc}\,h\omega)$, $\psi = O(\mathrm{sinc}\,(h\omega)\,\mathrm{sinc}\,(h\omega/2))$, which all follow from (2.9) and (2.10), we then get $a = O(1)$ and $b = O(\omega^{-1})$, and

$$\begin{aligned} \widehat{x}_2(t) &= be^{i\mu_2 t} + \overline{b}e^{-i\mu_2 t} + O(\omega^{-2}), \\ \widehat{x}_2'(t) &= i\omega\big(be^{i\mu_2 t} - \overline{b}e^{-i\mu_2 t}\big) + O(\omega^{-1}). \end{aligned} \tag{3.9}$$

The second relation is a consequence of the fact that $\gamma\psi = O(\omega^{-1}\,\mathrm{sinc}\,h\omega)$ and $\sin(h\mu_2) = \sin(h\omega)(1+O(\omega^{-2}))$ (which follows from $\lambda_2 = \cos h\mu_2 = \cos h\omega + O(h^2\phi\psi)$ and (2.11)). The statement (3.8) is now an immediate consequence of (3.9) and of the identity $|c+\overline{c}|^2 + |c-\overline{c}|^2 = 4|c|^2$, because the modulus of $c = be^{i\mu_2 t}$ is independent of t.

The near-conservation of the Hamiltonian can be seen similarly. If one of the conditions (3.5) and (3.6) is satisfied, we get $\gamma\phi\sin(\mu_2 h) = O(h)$ so that

$$\begin{aligned} \widehat{x}_1(t) &= ae^{i\mu_1 t} + \overline{a}e^{-i\mu_1 t} + O(\omega^{-1}), \\ \widehat{x}_1'(t) &= i\mu_1(ae^{i\mu_1 t} - \overline{a}e^{-i\mu_1 t}) + O(h^2) + O(\omega^{-1}). \end{aligned}$$

This implies $|\widehat{x}_1'(t)|^2 + \mu_1^2|\widehat{x}_1(t)|^2 = \mathrm{Const} + O(h^2) + O(\omega^{-1})$. Since $|\widehat{x}_2(t)| = O(\omega^{-1})$ and $\mu_1^2|\widehat{x}_1(t)|^2 = a_{11}|\widehat{x}_1(t)|^2 + O(h^2) = 2U\big(\widehat{x}(t)\big) + O(h^2) + O(\omega^{-1})$ for all t, this together with (3.8) proves the statement (3.7) for the total energy. ∎

Case II: *Nearly collapsing eigenvalues.* We now consider the complementary case

$$\tfrac{1}{4}h^2 a_{11} \;\le\; 1-\cos(h\omega) \;\le\; h^2 a_{11}, \tag{3.10}$$

so that the two eigenvalues of (3.3) are very close.

THEOREM 3.2. *Consider the numerical method* (2.2)–(2.3) *applied to* (3.1) *with* $a_{11} > 0$ *and* $x_2(0) = O(\omega^{-1})$. *Under the condition* (2.12) *on the numerical method, viz.,* $\psi(\xi) = \phi(\xi)\,\mathrm{sinc}^2\xi$, *we have*

$$H(x_n, \dot{x}_n) \;=\; H(x_0, \dot{x}_0) + O(h^2) + O(\omega^{-1}) \tag{3.11}$$

for all $n \ge 0$ *and uniformly in* $h\omega \ge c > 0$. *The constants symbolized by* $O(\cdot)$ *are independent of* ω, h, *and* n.

Proof. Under the condition (2.12) the numerical method satisfies (2.9), (2.10), (2.11) so that Theorem 3.1 is applicable. It therefore remains to consider the situation where $h\omega$ is restricted by (3.10).

The condition (3.10) implies $\frac{1}{2}h\sqrt{a_{11}} \le |\sin(h\omega)| \le h\sqrt{2a_{11}}$ for sufficiently small h. We now assume that $\phi\psi \ge 0$, which is satisfied by the choice (2.12). This guarantees that the eigenvalues of (3.3) are real, that $\lambda_i - \alpha_i = O(h^2\sqrt{\phi\psi})$, and that

$$\begin{aligned} \sin(h\mu_1) &= \pm\, h\sqrt{a_{11}} + O(h^3 + h\sqrt{\phi\psi}), \\ \sin(h\mu_2) &= \pm\sin(h\omega) + O(h\sqrt{\phi\psi}). \end{aligned} \tag{3.12}$$

For the special choice (2.12) we further have $|\gamma| \le 1/\sqrt{\phi\psi}$, $\gamma\psi = O(\omega^{-1})$, and $\gamma\phi = O(\omega)$, so that the coefficients of (3.4) satisfy $a = O(1)$ and $b = O(\omega^{-1})$. Moreover,

$$\gamma\big(\sin(h\mu_1) - \sin(h\mu_2)\big) = O(h) \tag{3.13}$$

holds. (Recall that $\gamma = \frac{1}{2}h^2 a_{21}/(\alpha_2 - \lambda_1)$.) This is a consequence of the estimate $h\sqrt{a_{11}}|\sin(h\mu_1) - \sin(h\mu_2)| \le |\sin^2(h\mu_1) - \sin^2(h\mu_2)| = |\lambda_2 - \lambda_1| = |\alpha_2 - \lambda_1| + O(h^2\sqrt{\phi\psi})$ and of the fact that $|\gamma| \le 1/\sqrt{\phi\psi}$.

The relations (3.12) and (3.13) thus yield

$$\begin{aligned}
\widehat{x}_1'(t) &= i\sqrt{a_{11}}\big((ae^{i\mu_1 t} - \overline{a}e^{-i\mu_1 t}) - \gamma\phi(be^{i\mu_2 t} - \overline{b}e^{-i\mu_2 t})\big) + O(h^2 + \omega^{-1}),\\
\widehat{x}_2'(t) &= i\omega\big(\gamma\psi(ae^{i\mu_1 t} - \overline{a}e^{-i\mu_1 t}) + (be^{i\mu_2 t} - \overline{b}e^{-i\mu_2 t})\big) + O(h^2 + \omega^{-1}).
\end{aligned}$$

Together with (3.4) this implies

$$\begin{aligned}
H\big(\widehat{x}(t), \widehat{x}\,'(t)\big) &= 2a_{11}|ae^{i\mu_1 t} - \gamma\phi be^{i\mu_2 t}|^2 + 2\omega^2|\gamma\psi ae^{i\mu_1 t} + be^{i\mu_2 t}|^2 + O(h^2 + \omega^{-1})\\
&= \text{const} - 4(a_{11}\gamma\phi - \omega^2\gamma\psi)\Re(a\overline{b}e^{i(\mu_1-\mu_2)t}) + O(h^2 + \omega^{-1}).
\end{aligned}$$

To prove that the t-dependent term is small, we need the relation (2.12) between ϕ and ψ. Using also (3.12) and (3.13) we obtain

$$\begin{aligned}
a_{11}\gamma\phi - \omega^2\gamma\psi &= \gamma\phi\big(a_{11} - \omega^2\,\mathrm{sinc}^2(h\omega)\big)\\
&= \gamma\phi\big(\sin^2(h\mu_1) - \sin^2(h\mu_2)\big)/h^2 + O(1) = O(1).
\end{aligned}$$

This completes the proof of Theorem 3.2, because $b = O(\omega^{-1})$. □

The proof above shows that in general

$$I\big(\widehat{x}(t), \widehat{x}\,'(t)\big) = 2\omega^2|\gamma\psi ae^{i\mu_1 t} + be^{i\mu_2 t}|^2 + O(h^2 + \omega^{-1}). \tag{3.14}$$

In the situation of Theorem 3.1 we have $a\gamma\psi = O(\omega^{-2})$ so that the first term in the right-hand expression of (3.14) becomes negligible, and long-time conservation of I can be concluded. If $a_{12} \ne 0$ and $\cos(h\omega) \approx 1 - \frac{1}{2}h^2 a_{11}$, we have $\lambda_{1,2} \approx \cos(h\omega) \mp \frac{1}{2}h^2 a_{12}\sqrt{\phi\psi}$ and $\gamma \approx 1/\sqrt{\phi\psi}$, so that $\gamma\psi \approx \sqrt{\psi/\phi}$ and $\gamma\phi \approx \sqrt{\phi/\psi}$. The initial condition $\widehat{x}(0) = x_0$ then gives

$$\gamma\psi(a + \overline{a}) \approx \tfrac{1}{2}(\gamma\psi x_{01} + x_{02}), \qquad (b + \overline{b}) \approx \tfrac{1}{2}(-\gamma\psi x_{01} + x_{02}),$$

so that none of the terms in (3.14) can be neglected. If $\mu_1 - \mu_2$ is different from an integral multiple of 2π, the expression (3.14) cannot remain close to a constant value. This result is intuitively clear, because in the situation (3.10) the two frequencies are indistinguishably close for the numerical method.

The upper pictures of Figure 8 show the maximal error in the Hamiltonian $H(x_n, \dot{x}_n)$ in dependence of $h\omega$ for the problem (3.1) with $a_{11} = 1$, $a_{12} = a_{21} = 1$, $a_{22} = 2$, and initial values $x(0) = (-0.9, \omega^{-1})^T$, $\dot{x}(0) = (1.0, 1.5)^T$. The three curves correspond to the step sizes $h = 0.2$, $h = 0.1$, and $h = 0.05$. The picture to the left it obtained with a method satisfying (2.12). Uniform convergence of the error can nicely be observed. The picture to the right corresponds to method (E) of Table 1. The lower pictures of Figure 8 plot the maximal deviation of the oscillatory energy $I(x_n, \dot{x}_n)$ as a function of $h\omega$. It confirms the analysis above, which shows that for $h\omega$ satisfying $\cos h\omega \approx 1 - \frac{1}{2}h^2 a_{11}$ the oscillatory energy cannot be well conserved.

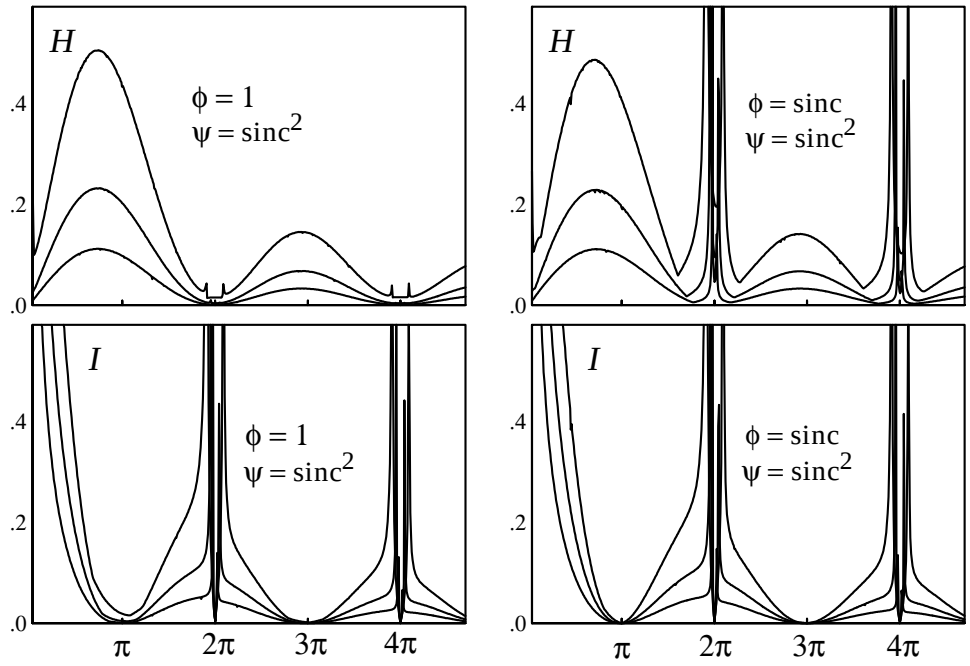

FIG. 8. *Error in the total and oscillatory energies as a function of $h\omega$ for a linear problem.*

4. Frequency expansion of the analytical solution. The main tool of our analysis for nonlinear problems is a decomposition of the solution $x(t)$ of (1.1) into a smooth part and into highly oscillatory terms with smoothly varying amplitudes. This decomposition is valid over finite time intervals. We show the existence of two almost-invariants for the coefficients of this decomposition, which are related to the total energy and the oscillatory energy of the system. A repeated use of these almost-invariants then allows us to prove the long-time near-conservation of the oscillatory energy.

4.1. The frequency expansion. We assume the nonlinearity g in (1.1) analytic on an open set D, and we consider solutions of (1.1) which satisfy

$$x(t) \in K, \tag{4.1}$$

where K is a compact subset of D. We assume further that the initial values have limited harmonic energy:

$$\tfrac{1}{2}\|\dot{x}(0)\|^2 + \tfrac{1}{2}\|\Omega x(0)\|^2 \le E, \tag{4.2}$$

where E is independent of ω.

THEOREM 4.1. *Under the assumptions* (4.2) *and* (4.1) *for* $0 \le t \le T$*, the solution* $x(t)$ *of* (1.1) *has for arbitrary* $N \ge 2$ *an expansion of the form*

$$x(t) = y(t) + \sum_{0<|k|<N} e^{ik\omega t} z^k(t) + R_N(t), \tag{4.3}$$

where the remainder term and its derivative are bounded by

$$R_N(t) = O(\omega^{-N-2}) \qquad \text{and} \qquad \dot{R}_N(t) = O(\omega^{-N-1}). \tag{4.4}$$

The real functions $y = (y_1, y_2)$ and the complex functions $z^k = (z_1^k, z_2^k)$ are bounded, together with all their derivatives, by

$$\begin{aligned} &y_1 = O(1), \qquad z_1^1 = O(\omega^{-3}), \quad z^k = O(\omega^{-k-2}), \quad k = 2, \dots, N-1, \\ &y_2 = O(\omega^{-2}), \quad z_2^1 = O(\omega^{-1}), \end{aligned} \tag{4.5}$$

and we have $z^{-k} = \overline{z^k}$. They are unique up to terms of size $O(\omega^{-N-2})$. The constants symbolized by the O-notation are independent of ω and t with $0 \le t \le T$ (but depend on E, N, T, and on the order of the derivative).

Proof. To determine the smooth functions $y(t)$, $z(t) = z^1(t)$, and $z^2(t), \dots, z^{N-1}(t)$, we put

$$\widetilde{x}(t) := y(t) + \sum_{0<|k|<N} e^{ik\omega t} z^k(t), \tag{4.6}$$

insert this function into (1.1), expand the nonlinearity around $y(t)$, and compare the coefficients of $e^{ik\omega t}$. With the notation $g^{(m)}(y)z^\alpha = g^{(m)}(y)(z^{\alpha_1}, \dots, z^{\alpha_m})$ for a multiindex $\alpha = (\alpha_1, \dots, \alpha_m)$, we obtain the following system of differential equations:

$$\begin{pmatrix} \ddot{y}_1 \\ \omega^2 y_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \ddot{y}_2 \end{pmatrix} = g(y) + \sum_{s(\alpha)=0} \frac{1}{m!}\, g^{(m)}(y) z^\alpha, \tag{4.7}$$

$$\begin{pmatrix} -\omega^2 z_1 \\ 2i\omega \dot{z}_2 \end{pmatrix} + \begin{pmatrix} 2i\omega \dot{z}_1 + \ddot{z}_1 \\ \ddot{z}_2 \end{pmatrix} = \sum_{s(\alpha)=1} \frac{1}{m!}\, g^{(m)}(y) z^\alpha, \tag{4.8}$$

$$\begin{pmatrix} -k^2\omega^2 z_1^k \\ (1-k^2)\omega^2 z_2^k \end{pmatrix} + \begin{pmatrix} 2ki\omega \dot{z}_1^k + \ddot{z}_1^k \\ 2ki\omega \dot{z}_2^k + \ddot{z}_2^k \end{pmatrix} = \sum_{s(\alpha)=k} \frac{1}{m!}\, g^{(m)}(y) z^\alpha. \tag{4.9}$$

Here the sums range over all $m \ge 1$ and all multi-indices $\alpha = (\alpha_1, \dots, \alpha_m)$ with integers α_j satisfying $0 < |\alpha_j| < N$, which have a given sum $s(\alpha) = \sum_{j=1}^m \alpha_j$.

For large ω, the dominating terms in these differential equations are given by the left-most expressions. However, since the central terms involve higher derivatives, we are confronted with singular perturbation problems. We are interested in smooth functions y, z, z^k that satisfy the system up to a defect of size $O(\omega^{-N})$. In the spirit of Euler's derivation of the Euler–Maclaurin summation formula (see, e.g., [HaW96]) we remove the disturbing higher derivatives by using iteratively the differentiated equations (4.7)–(4.9). This leads to a system

$$\begin{aligned} &\ddot{y}_1 = \mathcal{F}_1(\dot{y}_1, y, z^1, \dots, z^{N-1}, \omega^{-1}), && \dot{z}_2 = \omega^{-1}\mathcal{F}_2(\dot{y}_1, y, z^1, \dots, z^{N-1}, \omega^{-1}), \\ &z_1 = \omega^{-2}\mathcal{G}_1(\dot{y}_1, y, z^1, \dots, z^{N-1}, \omega^{-1}), && y_2 = \omega^{-2}\mathcal{G}_2(\dot{y}_1, y, z^1, \dots, z^{N-1}, \omega^{-1}), \\ &z_1^k = \omega^{-2}\mathcal{G}_1^k(\dot{y}_1, y, z^1, \dots, z^{N-1}, \omega^{-1}), && z_2^k = \omega^{-2}\mathcal{G}_2^k(\dot{y}_1, y, z^1, \dots, z^{N-1}, \omega^{-1}), \end{aligned}$$

where $\mathcal{F}_j, \mathcal{G}_j, \mathcal{G}_j^k$ are formal series in powers of ω^{-1}. Since we get formal algebraic relations for y_2, z_1, z^k, we can further eliminate these variables in the functions $\mathcal{F}_j, \mathcal{G}_j, \mathcal{G}_j^k$. We finally obtain for y_2, z_1, z^k the algebraic relations

$$\begin{aligned} z_1 &= \omega^{-2}\big(G_{10}(y_1, \dot{y}_1, z_2) + \omega^{-1} G_{11}(y_1, \dot{y}_1, z_2) + \cdots\big) \\ y_2 &= \omega^{-2}\big(G_{20}(y_1, \dot{y}_1, z_2) + \omega^{-1} G_{21}(y_1, \dot{y}_1, z_2) + \cdots\big) \\ z_1^k &= \omega^{-2}\big(G_{10}^k(y_1, \dot{y}_1, z_2) + \omega^{-1} G_{11}^k(y_1, \dot{y}_1, z_2) + \cdots\big) \\ z_2^k &= \omega^{-2}\big(G_{20}^k(y_1, \dot{y}_1, z_2) + \omega^{-1} G_{21}^k(y_1, \dot{y}_1, z_2) + \cdots\big) \end{aligned} \tag{4.10}$$

and a system of real second-order differential equations for y_1 and complex first-order differential equations for z_2:

$$\begin{aligned} \ddot{y}_1 &= F_{10}(y_1, \dot{y}_1, z_2) + \omega^{-1} F_{11}(y_1, \dot{y}_1, z_2) + \cdots \\ \dot{z}_2 &= \omega^{-1}\big(F_{20}(y_1, \dot{y}_1, z_2) + \omega^{-1} F_{21}(y_1, \dot{y}_1, z_2) + \cdots\big). \end{aligned} \tag{4.11}$$

At this point we can forget the above derivation and we can take it as a motivation for the ansatz (4.10)–(4.11), which we truncate after the $O(\omega^{-N})$ terms. Inserting this ansatz and its first and second derivatives into (4.7)–(4.9) and comparing like powers of ω^{-1} yields recurrence relations for the functions F^k_{jl}, G^k_{jl}. This shows that these functions, together with their derivatives, are all bounded on compact sets.

We determine initial values for (4.11) such that the function $\widetilde{x}(t)$ of (4.6) satisfies $\widetilde{x}(0) = x(0)$ and $\dot{\widetilde{x}}(0) = \dot{x}(0)$. Because of the special structure of the ansatz (4.10)–(4.11), this gives a system

$$\begin{aligned} x_1(0) &= y_1(0) + O(\omega^{-2}), \\ x_2(0) &= z_2(0) + \overline{z}_2(0) + O(\omega^{-2}), \\ \dot{x}_1(0) &= \dot{y}_1(0) + O(\omega^{-1}), \\ \dot{x}_2(0) &= i\omega z_2(0) - i\omega \overline{z}_2(0) + O(\omega^{-1}), \end{aligned} \tag{4.12}$$

which, by the implicit function theorem, yields (locally) unique initial values $y_1(0)$, $\dot{y}_1(0)$, $z_2(0)$. The assumption (4.2) implies that $z_2(0) = O(\omega^{-1})$. It further follows from the boundedness of F_{2l} that $z_2(t) = O(\omega^{-1})$ for $0 \le t \le T$. By looking closer at the structure of the function G^k_{jl} it can be seen that it contains at least k times the factor z_2. This implies the stated bounds for all other functions.

We still have to estimate the remainder $R_N(t) = x(t) - \widetilde{x}(t)$. For this we consider the solution of (4.10)–(4.11) with initial values (4.12). By construction, these functions satisfy the system (4.7)–(4.9) up to a defect of $O(\omega^{-N})$. This gives a defect of size $O(\omega^{-N})$, when the function $\widetilde{x}(t)$ of (4.6) is inserted into (1.1). Hence on a finite time interval $0 \le t \le T$, we obtain $R_N(t) = O(\omega^{-N})$ and $\dot{R}_N(t) = O(\omega^{-N})$. To obtain the slightly sharper bounds (4.4), we apply the above proof with N replaced by $N + 2$. □

4.2. The Hamiltonian of the frequency expansion. Now consider the situation where $g(x) = -\nabla U(x)$, so that (1.1) is a Hamiltonian system

$$\ddot{x} + \Omega^2 x = -\nabla U(x) \tag{4.13}$$

with Hamiltonian

$$H(x, \dot{x}) = \tfrac{1}{2}\dot{x}^T\dot{x} + \tfrac{1}{2}x^T\Omega^2 x + U(x), \tag{4.14}$$

where $U(x)$ is assumed to be analytic. Let $v^k(t) = e^{ik\omega t} z^k(t)$ $(0 < |k| < N)$ and note that by (4.7)–(4.9) these functions satisfy

$$\ddot{v}^k + \Omega^2 v^k = -\sum_{s(\alpha)=k} \frac{1}{m!}\, U^{(m+1)}(y) v^\alpha + O(\omega^{-N}). \tag{4.15}$$

Here, the sum is again over all $m \ge 1$ and all multi-indices $\alpha = (\alpha_1, \ldots, \alpha_m)$ with integers α_j $(0 < |\alpha_j| < N)$ which have a given sum $s(\alpha) = \sum_{j=1}^m \alpha_j$, and we write $v^\alpha = (v^{\alpha_1}, \ldots, v^{\alpha_m})$. Further we denote $V = (v^1, v^{-1}, \ldots, v^{N-1}, v^{-N+1})$ and we let

$$\mathcal{U}(y, V) = U(y) + \sum_{s(\alpha)=0} \frac{1}{m!}\, U^{(m)}(y) v^\alpha. \tag{4.16}$$

From (4.7) and (4.15) it follows that the vector (y, V) satisfies the system

$$\ddot y + \Omega^2 y = -\nabla_y \mathcal{U}(y,V) + O(\omega^{-N}), \tag{4.17}$$

$$\ddot v^k + \Omega^2 v^k = -\nabla_{v^{-k}} \mathcal{U}(y,V) + O(\omega^{-N}), \tag{4.18}$$

which, neglecting the $O(\omega^{-N})$ terms, is Hamiltonian with

$$\begin{aligned}\mathcal{H}(y,\dot y,V,\dot V) &= \tfrac12\big(\dot y^T\dot y + y^T\Omega^2 y\big)\\ &\quad + \tfrac12 \sum_{0<|k|<N}\Big((\dot v^{-k})^T\dot v^k + (v^{-k})^T\Omega^2 v^k\Big) + \mathcal{U}(y,V).\end{aligned} \tag{4.19}$$

THEOREM 4.2. *Under the assumptions* (4.2) *and* (4.1) *for* $0 \le t \le T$, *we have*

$$\mathcal{H}(y(t),\dot y(t),V(t),\dot V(t)) = \mathcal{H}(y(0),\dot y(0),V(0),\dot V(0)) + O(\omega^{-N}), \tag{4.20}$$

$$\mathcal{H}(y(t),\dot y(t),V(t),\dot V(t)) = H(x(t),\dot x(t)) + O(\omega^{-1}). \tag{4.21}$$

The constants symbolized by $O(\cdot)$ *are independent of* ω *and* t *with* $0 \le t \le T$ *but depend on* E, N, *and* T.

Proof. Multiplying (4.17) and (4.18) with $\dot y^T$ and $(\dot v^{-k})^T$, respectively, gives

$$\dot y^T(\ddot y + \Omega^2 y) + \sum_{0<|k|<N}(\dot v^{-k})(\ddot v^k + \Omega^2 v^k) = -\frac{d}{dt}\,\mathcal{U}(y,V) + O(\omega^{-N}).$$

Integrating from 0 to t and using $v^{-k} = \overline{v^k}$ then yields (4.20).

By the bounds of Theorem 4.1, we have for $0 \le t \le T$

$$\mathcal{H}(y,\dot y,V,\dot V) = \tfrac12|\dot y_1|^2 + |\dot v_2^1|^2 + \omega^2|v_2^1|^2 + U(y) + O(\omega^{-1}). \tag{4.22}$$

On the other hand, we have from (4.14) and (4.3) that

$$H(x,\dot x) = \tfrac12|\dot y_1|^2 + \tfrac12|\dot v_2^1 + \dot v_2^{-1}|^2 + \tfrac12\omega^2|v_2^1 + v_2^{-1}|^2 + U(y) + O(\omega^{-1}). \tag{4.23}$$

Using $v_2^1 = e^{i\omega t}z_2^1$ and $\dot v_2^1 = e^{i\omega t}(\dot z_2^1 + i\omega z_2^1)$ together with $v_2^{-1} = \overline{v_2^1}$, it follows from $\dot z_2^1 = O(\omega^{-1})$ that $\dot v_2^1 + \dot v_2^{-1} = i\omega(v_2^1 - v_2^{-1}) + O(\omega^{-1})$ and $|\dot v_2^1| = \omega|v_2^1| + O(\omega^{-1})$. Inserted into (4.22) and (4.23) this yields the statement (4.21). ∎

4.3. Another almost-invariant. Besides the Hamiltonian $\mathcal{H}(y,\dot y,V,\dot V)$, the coefficients of the frequency expansion have another almost-invariant. It depends only on the oscillating part and it is given by

$$\mathcal{I}(V,\dot V) = -i\omega \sum_{0<|k|<N} k\,(v^{-k})^T\dot v^k. \tag{4.24}$$

This almost-invariant turns out to be close to the energy of the harmonic oscillator,

$$I(x,\dot x) = \tfrac12|\dot x_2|^2 + \tfrac12\omega^2|x_2|^2. \tag{4.25}$$

THEOREM 4.3. *Under the assumptions* (4.2) *and* (4.1) *for* $0 \le t \le T$, *we have*

$$\mathcal{I}(V(t),\dot V(t)) = \mathcal{I}(V(0),\dot V(0)) + O(\omega^{-N}), \tag{4.26}$$

$$\mathcal{I}(V(t),\dot V(t)) = I(x(t),\dot x(t)) + O(\omega^{-1}). \tag{4.27}$$

The constants symbolized by $O(\cdot)$ are independent of ω and t with $0 \le t \le T$ but depend on E, N, and T.

Proof. With the vector $Z = (z^1, z^{-1}, \dots, z^{N-1}, z^{-N+1})$ it holds that $\mathcal{U}(y, Z) = \mathcal{U}(y, V)$ and also $\nabla_{z^{-k}}\mathcal{U}(y, Z) = e^{-ik\omega t}\nabla_{v^{-k}}\mathcal{U}(y, V)$. Differentiating the identity $0 = \mathcal{U}(y(t), V(t)) - \mathcal{U}(y(t), Z(t))$ with respect to t yields

$$\begin{aligned} 0 &= \sum_{0<|k|<N} \Big((\dot v^{-k})^T \nabla_{v^{-k}}\mathcal{U}(y(t), V(t)) - (\dot z^{-k})^T \nabla_{z^{-k}}\mathcal{U}(y(t), Z(t)) \Big) \\ &= \sum_{0<|k|<N} \big(\dot v^{-k} - e^{-ik\omega t}\dot z^{-k} \big)^T \nabla_{v^{-k}}\mathcal{U}(y(t), V(t)) \\ &= -\sum_{0<|k|<N} ik\omega (v^{-k})^T \nabla_{v^{-k}}\mathcal{U}(y(t), V(t)), \end{aligned} \tag{4.28}$$

because $v^{-k} = e^{-ik\omega t} z^{-k}$ and hence $\dot v^{-k} = e^{-ik\omega t}(\dot z^{-k} - ik\omega z^{-k})$.

The proof of Theorem 4.3 is now very similar to that of Theorem 4.2. We multiply the relation (4.18) with $-i\omega k (v^{-k})^T$ instead of $(\dot v^{-k})^T$. Summing up yields, with the use of (4.28),

$$-i\omega \sum_{0<|k|<N} k\, (v^{-k})^T \big(\ddot v^k + \Omega^2 v^k \big) = O(\omega^{-N}). \tag{4.29}$$

The derivative of $\mathcal{I}(V, \dot V)$, given by (4.24), is

$$\frac{d}{dt}\mathcal{I}(V, \dot V) = -i\omega \sum_{0<|k|<N} k \Big((v^{-k})^T \ddot v^k + (\dot v^{-k})^T \dot v^k \Big). \tag{4.30}$$

In the sums $\sum_k k(v^{-k})^T \Omega^2 v^k$ and $\sum_k k(\dot v^{-k})^T \dot v^k$, the terms with k and $-k$ cancel. Hence, the statement (4.26) follows from (4.29) and (4.30).

Using $\dot v^k = e^{ik\omega t}(\dot z^k + ik\omega z^k) = ik\omega v^k + O(\omega^{-1})$, it follows from the bounds of Theorem 4.1 that

$$\mathcal{I}(V, \dot V) = 2\omega^2 |v_2^1|^2 + O(\omega^{-1}).$$

On the other hand, using the arguments of the proof of Theorem 4.2, we have

$$I(x, \dot x) = \tfrac{1}{2}|\dot v_2^1 + \dot v_2^{-1}|^2 + \tfrac{1}{2}\omega^2 |v_2^1 + v_2^{-1}|^2 + O(\omega^{-1}) = 2\omega^2 |v_2^1|^2 + O(\omega^{-1}).$$

This proves the second statement of the theorem. ∎

COROLLARY 4.4. *If $x(t) \in K$ for $0 \le t \le \omega^N$, then*

$$I(x(t), \dot x(t)) = I(x(0), \dot x(0)) + O(\omega^{-1}) + O(t\omega^{-N}) \ .$$

The constants symbolized by $O(\cdot)$ are independent of ω and t with $0 \le t \le \omega^N$ but depend on E and N.

Proof. With a fixed $T > 0$, let V_j denote the vector of frequency expansion terms that correspond with starting values $(x(jT), \dot x(jT))$. For $t = (n+\theta)T$ with $0 \le \theta < 1$, we have by (4.27)

$$\begin{aligned} &I(x(t), \dot x(t)) - I(x(0), \dot x(0)) \\ &\quad = \mathcal{I}(V_n(\theta T), \dot V_n(\theta T)) + O(\omega^{-1}) - \mathcal{I}(V_0(0), \dot V_0(0)) + O(\omega^{-1}) \\ &\quad = \mathcal{I}(V_n(\theta T), \dot V_n(\theta T)) - \mathcal{I}(V_n(0), \dot V_n(0)) \\ &\qquad + \sum_{j=0}^{n-1} \Big(\mathcal{I}(V_{j+1}(0), \dot V_{j+1}(0)) - \mathcal{I}(V_j(0), \dot V_j(0)) \Big) + O(\omega^{-1}) \ . \end{aligned}$$

We note that $\mathcal{I}(V_{j+1}(0),\dot V_{j+1}(0))-\mathcal{I}(V_j(0),\dot V_j(0)) = O(\omega^{-N})$, because, by the uniqueness statement of Theorem 4.1, we have $V_{j+1}(0) = V_j(T) + O(\omega^{-N})$ and $\dot V_{j+1}(0) = \dot V_j(T) + O(\omega^{-N})$, and we have the bound (4.26) of Theorem 4.3. The same argument applies to $\mathcal{I}(V_n(\theta T),\dot V_n(\theta T)) - \mathcal{I}(V_n(0),\dot V_n(0))$. This yields the result. ∎

Remark. It is already known from the articles [BGG87] and [BGG89] that the oscillatory energy $I(x(t),\dot x(t))$ is nearly preserved over long times. The proofs in [BGG87] and [BGG89] are completely different. They use coordinate transforms from Hamiltonian perturbation theory and show that I is nearly preserved over time intervals which grow exponentially with ω. By carefully tracing the N-dependence of the constants in the $O(\omega^{-N})$-terms, it is possible to obtain near-conservation of I over exponentially long time intervals also within the present framework of frequency expansions.

5. Frequency expansion of the numerical solution. In this section we show that the numerical solution (2.2), (2.3) for nonlinear problems (1.1) has a frequency expansion similar to that of the analytical solution. Following the idea of backward analysis and motivated by the results of section 4 we look for a function

$$\widehat x(t) = y(t) + \sum_{0<|k|<N} e^{ik\omega t} z^k(t) \tag{5.1}$$

(with smooth $y(t)$ and $z^k(t)$ depending on h)[1] such that, up to a small defect,

$$\widehat x(t+h) - 2\cos h\Omega\ \widehat x(t) + \widehat x(t-h) \;=\; h^2\Psi\, g(\Phi\widehat x(t)), \quad \widehat x(0) = x_0,\ \widehat x(h) = x_1. \tag{5.2}$$

We assume throughout this section that

$$h\omega \ge d > 0, \tag{5.3}$$

and that the numerical solution Φx_n remains in a compact subset of the region where $g(x)$ is analytic, i.e.,

$$\Phi x_n \in K \qquad \text{for } \ 0 \le nh \le T. \tag{5.4}$$

5.1. Functional calculus. For the computation of the functions $y(t)$ and $z^k(t)$ the following functional calculus is convenient. Let f be an entire complex function bounded by $|f(\zeta)| \le C\, e^{\gamma|\zeta|}$. Then

$$f(hD)x(t) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}\, h^k\, x^{(k)}(t)$$

converges for every function x which is analytic in a disk of radius $r > \gamma h$ around t. We note that $(hD)^k x(t) = h^k x^{(k)}(t)$ for $k = 0,1,2,\dots$ and $\exp(hD)x(t) = x(t+h)$. If f_1 and f_2 are two such entire functions, then

$$f_1(hD)f_2(hD)x(t) = (f_1\cdot f_2)(hD)x(t)$$

whenever both sides exist. In particular, we have

$$x(t+h) - 2\cos h\Omega\, x(t) + x(t-h) = (e^{hD} - 2\cos h\Omega + e^{-hD})x(t).$$

[1] To avoid an overloaded notation with hats, we use the same letters y and z^k as for the analytical solution. We hope that this does not cause confusion.

We therefore introduce the operator

$$\begin{aligned}\mathcal{L}(hD) &= e^{hD} - 2\cos h\Omega + e^{-hD} \;=\; 2\big(\cos(ihD) - \cos h\Omega\big) \\ &= 4\,\sin\big(\tfrac{1}{2}h\Omega + \tfrac{1}{2}ihD\big)\sin\big(\tfrac{1}{2}h\Omega - \tfrac{1}{2}ihD\big)\end{aligned} \tag{5.5}$$

which, for $h \to 0$, is an approximation to $h^2(D^2 + \Omega^2)$.

We next study the application of such an operator to functions of the form $e^{i\omega t}z(t)$. By Leibniz's rule of calculus we have $(hD)^k e^{i\omega t}z(t) = e^{i\omega t}(hD + i\omega h)^k z(t)$. After a short calculation this also yields

$$f(hD)e^{i\omega t}z(t) = e^{i\omega t}f(hD + i\omega h)z(t) = e^{i\omega t}\sum_{k=0}^{\infty}\frac{f^{(k)}(i\omega h)}{k!}\,h^k\,z^{(k)}(t). \tag{5.6}$$

5.2. Modified equations for the coefficient functions of the frequency expansion. With the operator $\mathcal{L}(hD)$ of (5.5) the condition (5.2) becomes

$$\mathcal{L}(hD)\widehat{x}(t) = h^2\Psi g\big(\Phi\widehat{x}(t)\big). \tag{5.7}$$

Inserting the ansatz (5.1), expanding the right-hand side of (5.7) into a Taylor series around $\Phi y(t)$, and comparing the coefficients of $e^{ik\omega t}$ yield for the functions $y(t)$ and $z^k(t)$

$$\begin{aligned}\mathcal{L}(hD)y &= h^2\Psi\Big(g(\Phi y) + \sum_{s(\alpha)=0}\frac{1}{m!}\,g^{(m)}(\Phi y)(\Phi z)^\alpha\Big), \\ \mathcal{L}(hD + ik\omega h)z^k &= h^2\Psi\sum_{s(\alpha)=k}\frac{1}{m!}\,g^{(m)}(\Phi y)(\Phi z)^\alpha.\end{aligned} \tag{5.8}$$

Here, $\alpha = (\alpha_1, \ldots, \alpha_m)$ is a multi-index as in the proof of Theorem 4.1, $s(\alpha) = \sum_{j=1}^m \alpha_j$, and $(\Phi z)^\alpha$ is an abbreviation for the m-tupel $(\Phi z^{\alpha_1}, \ldots, \Phi z^{\alpha_m})$. To get smooth functions $y(t)$ and $z^k(t)$ which solve (5.8) up to a small defect, we look at the dominating terms in the Taylor expansions of $\mathcal{L}(hD)$ and $\mathcal{L}(hD + ik\omega h)$. With the abbreviations $s_k = \sin(\frac{1}{2}kh\omega)$ and $c_k = \cos(\frac{1}{2}kh\omega)$ we have

$$\begin{aligned}\mathcal{L}(hD) &= \begin{pmatrix}0 & 0\\ 0 & 4s_1^2\end{pmatrix} - \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}(ihD)^2 + \cdots, \\ \mathcal{L}(hD + ih\omega) &= \begin{pmatrix}-4s_1^2 & 0\\ 0 & 0\end{pmatrix} + 2s_2\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}(ihD) - c_2\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}(ihD)^2 + \cdots, \\ \mathcal{L}(hD + ikh\omega) &= \begin{pmatrix}-4s_k^2 & 0\\ 0 & 4s_{k-1}s_{k+1}\end{pmatrix} + 2s_{2k}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}(ihD) \\ &\qquad - c_{2k}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}(ihD)^2 + \cdots.\end{aligned} \tag{5.9}$$

The situation is now more complicated than in (4.7)–(4.9) for the frequency expansion of the analytical solution, because several of the coefficients in (5.9) may vanish due to numerical resonance. We here confine the discussion to the nonresonant case. We assume that h and ω^{-1} lie in a subregion of the (h, ω^{-1})-plane of small parameters for which there exists a positive constant c such that

$$|\sin(\tfrac{1}{2}kh\omega)| \geq c\sqrt{h} \qquad \text{for} \quad k = 1, \ldots, N, \quad \text{with} \quad N \geq 2. \tag{5.10}$$

The condition excludes that $h\omega$ is $o(\sqrt{h})$ close to integral multiples of π. For given h and ω, the condition imposes a restriction on N. In the following, N is a fixed integer such that (5.10) holds.

THEOREM 5.1. *Under the limited-energy condition* (4.2)*, under the nonresonance condition* (5.10)*, under the conditions* (5.3), (5.4)*, and under the conditions* (2.9) *and* (2.10) *on the numerical method* (2.2)–(2.4)*, the numerical solution is of the form*

$$x_n = y(t) + \sum_{0<|k|<N} e^{ik\omega t} z^k(t) + \begin{pmatrix} O(t^2 h^N) \\ O(\psi(h\omega) t^2 h^N) \end{pmatrix}$$

uniformly for $0 \le t = nh \le T$*, where the functions* $y = (y_1, y_2)$ *and* $z^k = (z_1^k, z_2^k)$ *satisfy* (5.8) *up to a defect of* $O(h^{N+2})$ *in their first components, and* $O(\psi(h\omega)h^{N+2})$ *in their second components. Together with all their derivatives these functions are bounded by*

$$\begin{aligned} &y_1 = O(1), && z_1^1 = O(h\phi(h\omega)\omega^{-1}), && z_1^k = O(h\phi(h\omega)^k\omega^{-k}), \\ &y_2 = O(\omega^{-2}), && z_2^1 = O(\omega^{-1}), && z_2^k = O(h\psi(h\omega)\phi(h\omega)^k\omega^{-k}) \end{aligned} \tag{5.11}$$

for $k = 2, \dots, N-1$*. We have* $z^{-k} = \overline{z^k}$*, and the constants symbolized by the O-notation are independent of* ω *and* h *but depend on* E*,* N*, and* T*.*

Proof. Under assumption (5.10), the first nonvanishing coefficients in (5.9) are the dominant ones, and the derivation of the defining relations for y and z^k is the same as for the analytical solution in Theorem 4.1. We insert (5.9) into (5.8) and we eliminate recursively the higher derivatives. This motivates the following ansatz for the computation of the functions y and z^k:

$$\begin{aligned} \ddot{y}_1 &= f_{10}(\cdot/\cdot) + \sqrt{h}\, f_{11}(\cdot/\cdot) + \cdots, \\ \dot{z}_2 &= \frac{\psi(h\omega)h}{s_2}\Big(f_{20}(\cdot/\cdot) + \sqrt{h}\, f_{21}(\cdot/\cdot) + \cdots\Big), \\ z_1 &= \frac{h^2}{s_1^2}\Big(g_{10}(\cdot/\cdot) + \sqrt{h}\, g_{11}(\cdot/\cdot) + \cdots\Big), \\ y_2 &= \frac{\psi(h\omega)h^2}{s_1^2}\Big(g_{20}(\cdot/\cdot) + \sqrt{h}\, g_{21}(\cdot/\cdot) + \cdots\Big), \\ z_1^k &= \frac{h^2}{s_k^2}\Big(g_{10}^k(\cdot/\cdot) + \sqrt{h}\, g_{11}^k(\cdot/\cdot) + \cdots\Big), \\ z_2^k &= \frac{\psi(h\omega)h^2}{s_{k+1}s_{k-1}}\Big(g_{20}^k(\cdot/\cdot) + \sqrt{h}\, g_{21}^k(\cdot/\cdot) + \cdots\Big), \end{aligned} \tag{5.12}$$

where the functions depend smoothly on the variables y_1, $\dot{y}_1$, z_2 and on the bounded parameters $\sqrt{h}/s_k$, s_k, c_k, and $\psi(h\omega)$. Inserting this ansatz and its derivatives into (5.8) and comparing like powers of $\sqrt{h}$ yields recurrence relations for the functions f_{jl}^k, g_{jl}^k. The functions g_{jl}^k (for $k \ge 1$) contain at least k times the factor $\phi(h\omega)z_2$, and f_{2l} contains at least once this factor. Since the series in (5.12) need not converge, we truncate them after the $(\sqrt{h})^{2(N+2)}$ terms.

We next determine the initial values $y_1(0)$, $\dot{y}_1(0)$, and $z_2(0)$ such that $\widehat{x}(0)$ and $\widehat{x}(h)$ of (5.1) coincide with the starting values x_0 and x_1 of the numerical scheme (x_1 is computed from x_0 and $\dot{x}_0$ via the formula (2.2) with $n = 0$). Using the nonresonance

assumption (5.10), the condition $\widehat{x}(0) = x_0 = (x_{01}, x_{02})$ becomes

$$\begin{aligned} x_{01} &= y_1(0) + O\big(h\phi(h\omega)z_2(0)\big), \\ x_{02} &= z_2(0) + \overline{z}_2(0) + O\big(h^2\psi(h\omega)/s_1^2\big) + O\big(h\psi(h\omega)\phi(h\omega)z_2(0)\big). \end{aligned} \tag{5.13}$$

The formula for the first component of (2.2), $x_{11} - x_{01} = h\dot{x}_{01} + \frac{1}{2}h^2 g_1(\Phi x_0)$, together with $\widehat{x}_1(h) - \widehat{x}_1(0) = h\dot{y}_1(0) + O(h^2) + O(h\phi(h\omega)z_2(0))$ implies that

$$\dot{x}_{01} = \dot{y}_1(0) + O(h) + O\big(\phi(h\omega)z_2(0)\big). \tag{5.14}$$

For the second component we have $x_{12} - \cos h\omega\, x_{02} = h\,\mathrm{sinc}\, h\omega\, \dot{x}_{02} + O(h^2\psi(h\omega))$ from (2.2), and $\widehat{x}_2(h) - \cos h\omega\, \widehat{x}_2(0) = (1 - \cos h\omega)y_2(0) + O(h^2\psi(h\omega)) + i \sin h\omega\big(z_2(0) - \overline{z}_2(0)\big) + O\big(h\psi(h\omega)\phi(h\omega)z_2(0)\big)$, which after division by $h\,\mathrm{sinc}\, h\omega$ yields

$$\begin{aligned} \dot{x}_{02} \;=\; i\omega\big(z_2(0) - \overline{z}_2(0)\big) \;&+\; O(h\psi(h\omega)/\mathrm{sinc}\, h\omega) \\ &+\; O\big(\psi(h\omega)\phi(h\omega)z_2(0)/\mathrm{sinc}\, h\omega\big). \end{aligned} \tag{5.15}$$

The four equations (5.13), (5.14), (5.15) constitute a nonlinear system for the four quantities $y_1(0)$, $\dot{y}_1(0)$, $\omega\big(z_2(0) + \overline{z}_2(0)\big)$, and $\omega\big(z_2(0) - \overline{z}_2(0)\big)$. By the implicit function theorem and using the limited-energy assumption (4.2), we get a locally unique solution for sufficiently small h, if the conditions (2.9) and (2.10) are satisfied.

The initial value for z_2 satisfies $z_2(0) = O(\omega^{-1})$, and it follows from (2.10) that $h\psi(h\omega)/s_2 = O(\omega^{-1})$, so that $\dot{z}_2 = O(\omega^{-1}z_2)$ by (5.12). This implies $z_2(t) = O(\omega^{-1})$ for $t \le T$. The other estimates (5.11) are directly obtained from (5.12). Consequently, the values $\widehat{x}(nh)$ inserted into the numerical scheme (2.5) yield a defect of size $O(h^{N+2})$:

$$\widehat{x}((n+1)h) - 2\cos h\Omega\, \widehat{x}(nh) + \widehat{x}((n-1)h) = h^2\Psi\left(g(\Phi\widehat{x}(nh)) + O(h^N)\right). \tag{5.16}$$

Standard convergence estimates then show that on bounded time intervals $x_n - \widehat{x}(nh)$ is of size $O(t^2h^N)$ in the first component and of size $O(\psi(h\omega)t^2h^N)$ in the second component. This completes the proof of Theorem 5.1. □

5.3. Frequency expansion of the derivative approximation. Under the condition (5.10) we have $h\omega \ne k\pi$ for integer k, so that the derivative approximation $\dot{x}_n$ is given by (2.6). We now define $\widehat{x}'(t)$ by the continuous analogue

$$\widehat{x}(t+h) - \widehat{x}(t-h) = 2h\,\mathrm{sinc}\,(h\Omega)\,\widehat{x}'(t). \tag{5.17}$$

Using condition (2.10), Theorem 5.1 implies that

$$\dot{x}_n = \widehat{x}'(nh) + O(t^2h^{N-1}) + O(h^{N+1})$$

on bounded time intervals. We next write the function $\widehat{x}'(t)$ as

$$\widehat{x}'(t) = y'(t) + \sum_{0<|k|<N} e^{ik\omega t} z'^k(t). \tag{5.18}$$

Inserting the relation (5.1) into $-i\sin(ihD)\widehat{x}(t) = h\,\mathrm{sinc}\,(h\Omega)\widehat{x}'(t)$, which is equivalent to (5.17), and comparing the coefficients of $e^{ik\omega t}$ we obtain

$$\begin{aligned} \mathrm{sinc}\,(ihD)\,\dot{y}_1 \;&=\; y_1', \\ \mathrm{sinc}\,(ihD)\,\dot{y}_2 \;&=\; \mathrm{sinc}\,(h\omega)y_2', \\ (ih)^{-1}\,\sin(ihD - k\omega h)\,z_1^k \;&=\; z_1'^k, \\ (ih)^{-1}\,\sin(ihD - k\omega h)\,z_2^k \;&=\; \mathrm{sinc}\,(h\omega)z_2'^k. \end{aligned} \tag{5.19}$$

In particular, we get for z_2^1 that

$$z_2'^1 \;=\; i\omega\cos(ihD)\,z_2^1 - i\omega\,\frac{\cos\omega h}{\sin\omega h}\,\sin(ihD)\,z_2^1. \tag{5.20}$$

THEOREM 5.2. *Under the assumptions of Theorem* 5.1*, the numerical solution* $\dot x_n$*, given by* (2.6)*, satisfies*

$$\dot x_n = y'(t) + \sum_{0<|k|<N} e^{ik\omega t} z'^k(t) + O(t^2h^{N-1}) + O(h^{N+1})$$

uniformly for $0 \le t = nh \le T$*, where the functions* $y' = (y_1', y_2')$ *and* $z'^k = (z_1'^k, z_2'^k)$*, together with all their derivatives, are bounded by*

$$\begin{array}{lll} y_1' = \dot y_1 + O(h^2), & z_1'^1 = O(\omega^{-1}), & z_1'^k = O(\omega^{-k}), \\ z_2'^1 = i\omega z_2^1 + O(\omega^{-1}), & y_2' = O(\omega^{-1}), & z_2'^k = O(\omega^{-k}), \quad k = 2,\dots,N-1. \end{array} \tag{5.21}$$

The constants symbolized by the O-notation are independent of ω *and* h *but depend on* E*,* N*, and* T*.*

Proof. The estimates follow from (5.19) and from Theorem 5.1. For y_2', $z_1'^1$, and $z_2'^1$ we use the formulas (5.12) to get the sharper result. ∎

5.4. Energy along the numerical solution. In the Hamiltonian case $g(y) = -\nabla U(y)$, the total energy H and the oscillatory energy I are related to the frequency expansion coefficients as follows.

LEMMA 5.3. *If the coefficients of the frequency expansions for* x_n *and* $\dot x_n$ *satisfy* (5.11) *and* (5.21)*, respectively, then*

$$H(\widehat x, \widehat x') = \tfrac12\,|\dot y_1|^2 + 2\omega^2|z_2^1|^2 + U(y) + O(\omega^{-1}) + O(h^2), \tag{5.22}$$

$$I(\widehat x, \widehat x') = 2\omega^2|z_2^1|^2 + O(\omega^{-1}). \tag{5.23}$$

Proof. By definition (4.25) we have $I(\widehat x, \widehat x') = \frac12|\widehat x_2'|^2 + \frac12\omega^2|\widehat x_2|^2$. Since

$$\begin{aligned} \omega\widehat x_2(t) &= \omega\big(e^{i\omega t}z_2^1(t) + e^{-i\omega t}z_2^{-1}(t)\big) + O(\omega^{-1}) \qquad \text{by (5.11)}, \\ \widehat x_2'(t) &= i\omega\big(e^{i\omega t}z_2^1(t) - e^{-i\omega t}z_2^{-1}(t)\big) + O(\omega^{-1}) \qquad \text{by (5.21)}, \end{aligned}$$

the statement (5.23) follows from the fact that $|v+\overline v|^2 + |v-\overline v|^2 = 4|v|^2$. The formula (5.22) can be proved in the same way. ∎

6. Almost-invariants of the numerical frequency expansion. In this section we show that, in the Hamiltonian case $g(y) = -\nabla U(y)$, the coefficients of the frequency expansion of the numerical solution have invariants that can be obtained as in section 4. We denote $V = (v^1, v^{-1}, \dots, v^{N-1}, v^{-N+1})$ with $v^k(t) = e^{ik\omega t}z^k(t)$, where z^k are the coefficients of the frequency expansion (5.1). Similar to (4.16) we consider the function

$$\mathcal U(y, V) = U(\Phi y) + \sum_{s(\alpha)=0} \frac{1}{m!}\,U^{(m)}(\Phi y)(\Phi v)^\alpha, \tag{6.1}$$

where the sum is taken over all $m \ge 1$ and all multi-indices $\alpha = (\alpha_1, \dots, \alpha_m)$ with nonvanishing integral components for which $s(\alpha) = \sum_j \alpha_j = 0$ and $0 < |\alpha_j| < N$. It

then follows from Theorem 5.1 that the coefficients y and v^k satisfy

$$\Phi\,\mathcal{L}(hD)y = -h^2\Psi\left(\nabla_y \mathcal{U}(y,V) + O(h^N)\right), \tag{6.2}$$

$$\Phi\,\mathcal{L}(hD)v^k = -h^2\Psi\left(\nabla_{v^{-k}} \mathcal{U}(y,V) + O(\phi(h\omega)h^N)\right). \tag{6.3}$$

The factor $\phi(h\omega)$ in the defect of (6.3) is due to the presence of the factor $\phi(h\omega)z_2$ in the relations (5.12) defining the z-components.

The similarity of these relations to (4.17), (4.18) allows us to obtain invariants that are the analogues of $\mathcal{H}$ and $\mathcal{I}$ of section 4.

6.1. First invariant. As in section 4.2 we multiply (6.2) and (6.3) by $\dot y^T$ and $(\dot v^{-k})^T$, respectively, and we thus obtain

$$\dot y^T\Psi^{-1}\Phi h^{-2}\mathcal{L}(hD)y + \sum_{0<|k|<N} (\dot v^{-k})^T\Psi^{-1}\Phi h^{-2}\mathcal{L}(hD)v^k + \frac{d}{dt}\mathcal{U}(y,V) = O(h^N).$$

Since we know bounds on z^k and on its derivatives (Theorem 5.1), we switch to the quantities z^k and we get the equivalent relation

$$\begin{aligned}\dot y^T\Psi^{-1}\Phi h^{-2}\mathcal{L}(hD)y \;+ \sum_{0<|k|<N} &(\dot z^{-k} - ik\omega z^{-k})^T\Psi^{-1}\Phi h^{-2}\mathcal{L}(hD+ik\omega h)z^k \\ &+\frac{d}{dt}\,\mathcal{U}(y,V) = O(h^N).\end{aligned} \tag{6.4}$$

We shall show that the left-hand side is the total derivative of an expression that depends only on y, z^k and its derivatives. Indeed, the term $\dot y^T y^{(2l)}$ can be written as

$$\dot y^T y^{(2l)} \;=\; \frac{d}{dt}\Big(\dot y^T y^{(2l-1)} - \ddot y^T y^{(2l-2)} + \cdots \mp (y^{(l-1)})^T y^{(l+1)} \pm \frac{1}{2}\,(y^{(l)})^T y^{(l)}\Big).$$

Similarly, we get for $z = z^k$ and $\overline z = z^{-k}$ that

$$\begin{aligned}
\operatorname{Re}\dot{\overline z}^T z^{(2l)} \;&=\; \operatorname{Re}\frac{d}{dt}\Big(\dot{\overline z}^T z^{(2l-1)} - \cdots \mp (\overline z^{(l-1)})^T z^{(l+1)} \pm \frac{1}{2}\,(\overline z^{(l)})^T z^{(l)}\Big),\\
\operatorname{Re}\overline z^T z^{(2l+1)} \;&=\; \operatorname{Re}\frac{d}{dt}\Big(\overline z^T z^{(2l)} - \cdots \pm (\overline z^{(l-1)})^T z^{(l+1)} \mp \frac{1}{2}\,(\overline z^{(l)})^T z^{(l)}\Big),\\
\operatorname{Im}\dot{\overline z}^T z^{(2l+1)} \;&=\; \operatorname{Im}\frac{d}{dt}\Big(\dot{\overline z}^T z^{(2l)} - \ddot{\overline z}^T z^{(2l-1)} + \cdots \mp (\overline z^{(l)})^T z^{(l+1)}\Big),\\
\operatorname{Im}\overline z^T z^{(2l+2)} \;&=\; \operatorname{Im}\frac{d}{dt}\Big(\overline z^T z^{(2l+1)} - \dot{\overline z}^T z^{(2l)} + \cdots \pm (\overline z^{(l)})^T z^{(l+1)}\Big).
\end{aligned}$$

Hence, there exists a function $\widehat{\mathcal{H}}_0[y,Z](t)$, which depends on the values at t of the functions y and $Z = (z^1, z^{-1}, \dots, z^{N-1}, z^{-N+1})$ and of their first N derivatives, such that (6.4) reads

$$\frac{d}{dt}\widehat{\mathcal{H}}_0[y,Z](t) = O(h^N).$$

This yields immediately the first statement of the following result.

THEOREM 6.1. *Under the assumptions of Theorem* 5.1*, the coefficient functions* y *and* $Z = (z^1, z^{-1}, \dots, z^{N-1}, z^{-N+1})$ *satisfy*

$$\widehat{\mathcal{H}}_0[y,Z](t) \;=\; \widehat{\mathcal{H}}_0[y,Z](0) \;+\; O(th^N)$$

for $0 \le t \le T$. *Moreover,*

$$\widehat{\mathcal{H}}_0[y,Z](t) \;=\; \tfrac{1}{2}\,|\dot y_1(t)|^2 + U(\Phi y(t)) + \mu(h\omega)\,2\omega^2|z_2^1(t)|^2 + O(h^2),$$

where $\mu(h\omega) = \operatorname{sinc}(h\omega)\phi(h\omega)/\psi(h\omega)$.

Proof. The formula for $\widehat{\mathcal{H}}_0$ is obtained from the formulas (5.9) for $\mathcal{L}(hD + ikh\omega)$ together with the estimates of Theorem 5.1. ∎

Remark. Symplectic discretizations have $\psi(\xi) = \operatorname{sinc}\xi \cdot \phi(\xi)$, so that $\mu(h\omega) = 1$.

6.2. Second invariant. As in the proof of Theorem 4.3 we have for the function $\mathcal{U}(y,V)$ of (6.1) that

$$0 \;=\; -\sum_{0<|k|<N} ik\omega (v^{-k})^T \nabla_{v^{-k}} \mathcal{U}\big(y(t),V(t)\big).$$

Consequently, it follows from (6.3) that

$$-i\omega \sum_{0<|k|<N} k(v^{-k})^T \Psi^{-1}\Phi h^{-2}\mathcal{L}(hD)v^k \;=\; O(\phi(h\omega)h^N).$$

Written in the z variables, this becomes

$$-i\omega \sum_{0<|k|<N} k(z^{-k})^T \Psi^{-1}\Phi h^{-2}\mathcal{L}(hD + ik\omega h)z^k \;=\; O(\phi(h\omega)h^N). \tag{6.5}$$

As in section 6.1, the left-hand expression can be written as the total derivative of a function $\widehat{\mathcal{I}}_0[Z](t)$ which depends on the values at t of the function Z and its first N derivatives:

$$\frac{d}{dt}\widehat{\mathcal{I}}_0[Z](t) = O(\phi(h\omega)h^N).$$

THEOREM 6.2. *Under the assumptions of Theorem* 5.1*, the coefficient functions* $Z = (z^1, z^{-1}, \dots, z^{N-1}, z^{-N+1})$ *satisfy*

$$\widehat{\mathcal{I}}_0[Z](t) \;=\; \widehat{\mathcal{I}}_0[Z](0) \;+\; O(t\phi(h\omega)h^N)$$

for $0 \le t \le T$. *Moreover,*

$$\widehat{\mathcal{I}}_0[Z](t) \;=\; \mu(h\omega)\Big(2\omega^2|z_2^1(t)|^2 + O(h^2)\Big),$$

with $\mu(h\omega)$ *as in Theorem* 6.1.

Proof. From (6.5) and the estimates of Theorem 5.1, we obtain

$$\widehat{\mathcal{I}}_0[Z](t) = \mu(h\omega)\,2\omega^2|z_2^1(t)|^2 + O(\phi(h\omega)h^2).$$

Because of condition (2.10), this yields the stated formula for $\widehat{\mathcal{I}}_0$. ∎

7. Long-time energy conservation of the numerical discretization. We are now able to prove the main result of this paper. This shows that the total energy H and the oscillatory energy I are nearly conserved over time intervals of length $C_N h^{-N}$, for any N for which the nonresonance condition (5.10) is satisfied.

For the convenience of the reader we restate our assumptions:

- the limited-energy condition (4.2): $\frac{1}{2}\|\dot x(0)\|^2 + \frac{1}{2}\|\Omega x(0)\|^2 \le E$;

- the boundedness condition (5.4) for the numerical solution sequence: Φx_n stays in a compact subset of the domain of analyticity of g;
- the condition (5.3): $h\omega \geq d > 0$;
- the conditions (2.9) and (2.10) on the numerical method:

$$|\psi(h\omega)| \leq C_1 \operatorname{sinc}^2(\tfrac{1}{2}h\omega) \ , \qquad |\psi(h\omega)| \leq C_2 \, |\operatorname{sinc}(h\omega)|;$$

- the nonresonance condition (5.10): for some $N \geq 2$,

$$|\sin(\tfrac{1}{2}kh\omega)| \geq c\sqrt{h} \quad \text{for} \ \ k = 1, \ldots, N.$$

THEOREM 7.1. *Under the above conditions, the numerical solution of* (1.1) *obtained by the method* (2.2)–(2.4) *satisfies*

$$\begin{aligned} H(x_n, \dot{x}_n) &= H(x_0, \dot{x}_0) + O(h) + O(th^N), \\ I(x_n, \dot{x}_n) &= I(x_0, \dot{x}_0) + O(h) + O(th^N), \end{aligned}$$

where $t = nh$. The constants symbolized by $O(\cdot)$ are independent of n and of h and ω satisfying the above conditions but depend on N.

Proof. (a) If we consider the linear combinations $\widehat{\mathcal{H}} = \widehat{\mathcal{H}}_0 - (1 - 1/\mu)\widehat{\mathcal{I}}_0$ and $\widehat{\mathcal{I}} = \widehat{\mathcal{I}}_0/\mu$, it follows from Theorem 6.1 and Theorem 6.2 that

$$\begin{aligned} \widehat{\mathcal{H}}[y, Z](t) &= \widehat{\mathcal{H}}[y, Z](0) + O(th^N), \\ \widehat{\mathcal{I}}[Z](t) &= \widehat{\mathcal{I}}[Z](0) + O(th^N). \end{aligned} \tag{7.1}$$

Moreover, by Theorem 6.1 and Theorem 6.2, together with Lemma 5.3 we have

$$\begin{aligned} \widehat{\mathcal{H}}[y, Z](t) &= H\big(\widehat{x}(t), \widehat{x}'(t)\big) + O(h^2) + O(\omega^{-1}), \\ \widehat{\mathcal{I}}[Z](t) &= I\big(\widehat{x}(t), \widehat{x}'(t)\big) + O(\omega^{-1}), \end{aligned} \tag{7.2}$$

where again $\widehat{x}(t)$ is defined by the frequency expansion (5.1) with coefficients $y(t)$ and $Z(t)$, and $\widehat{x}'(t)$ is defined by (5.17). The relations (7.1) and (7.2) hold only on finite time intervals $0 \leq t \leq T$, on which the frequency expansion is defined.

(b) We now apply the above relations repeatedly on intervals of length h, for frequency expansions corresponding to different starting values. As long as $(x_n, \dot{x}_n)$ satisfies the limited-energy condition (4.2) (possibly with a larger constant E), Theorem 5.1 gives us frequency expansion coefficients $y_n(t), Z_n(t)$ corresponding to starting values $(x_n, \dot{x}_n)$. Because of the uniqueness (up to $O(h^{N+1})$) of the coefficients of the frequency expansion, the following diagram commutes up to terms of size $O(h^{N+1})$:

$$\begin{array}{ccc} (x_n, \dot{x}_n) & \longleftrightarrow & (y_n(0), \dot{y}_n(0), Z_n(0), \dot{Z}_n(0)) \\ & & \Big\downarrow \text{ flow} \\ \Big\downarrow \begin{array}{l}\text{numerical} \\ \text{method}\end{array} & & (y_n(h), \dot{y}_n(h), Z_n(h), \dot{Z}_n(h)) \\ & & \qquad = \ (\text{up to } O(h^{N+1})) \\ (x_{n+1}, \dot{x}_{n+1}) & \longleftrightarrow & (y_{n+1}(0), \dot{y}_{n+1}(0), Z_{n+1}(0), \dot{Z}_{n+1}(0)). \end{array} \tag{7.3}$$

The construction of the coefficient functions via (5.12) shows that higher derivatives of (y_n, Z_n) at h and (y_{n+1}, Z_{n+1}) at 0 also differ by only $O(h^{N+1})$. We thus have

from (7.3) and (7.1)

$$\begin{aligned}\widehat{\mathcal{H}}[y_{n+1}, Z_{n+1}](0) &= \widehat{\mathcal{H}}[y_n, Z_n](h) + O(h^{N+1}) \\ &= \widehat{\mathcal{H}}[y_n, Z_n](0) + O(h^{N+1}).\end{aligned}$$

Using this relation repeatedly, we obtain

$$\widehat{\mathcal{H}}[y_n, Z_n](0) \;=\; \widehat{\mathcal{H}}[y_0, Z_0](0) + O(nh^{N+1}).$$

Moreover, from (7.2) we have the following for the coefficient functions corresponding to the starting values $(x_n, \dot{x}_n)$ and $(x_0, \dot{x}_0)$:

$$\begin{aligned}\widehat{\mathcal{H}}[y_n, Z_n](0) &= H\big(\widehat{x}_n(0), \widehat{x}'_n(0)\big) + O(h^2) + O(\omega^{-1}), \\ \widehat{\mathcal{H}}[y_0, Z_0](0) &= H\big(\widehat{x}_0(0), \widehat{x}'_0(0)\big) + O(h^2) + O(\omega^{-1}).\end{aligned}$$

Since $\widehat{x}_n(0) = x_n$ by construction, and $\widehat{x}'_n(0) = \dot{x}_n + O(h^{N+1})$ by Theorem 5.2, we obtain

$$\begin{aligned}H(x_n, \dot{x}_n) - H(x_0, \dot{x}_0) &= H(\widehat{x}_n(0), \widehat{x}'_n(0)) - H(\widehat{x}_0(0), \widehat{x}'_0(0)) + O(h^{N+1}) \\ &= \widehat{\mathcal{H}}[y_n, Z_n](0) - \widehat{\mathcal{H}}[y_0, Z_0](0) + O(h^2) + O(\omega^{-1}) \\ &= O(nh^{N+1}) + O(h^2) + O(\omega^{-1}),\end{aligned}$$

which gives the desired bound for the deviation of the total energy along the numerical solution. The same argument applies to $I(x_n, \dot{x}_n)$. ∎

Acknowledgment. We are grateful to Sebastian Reich for drawing our attention to the Fermi–Pasta–Ulam problem.

REFERENCES

[BGG87] 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), pp. 87–103.

[BGG89] 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* II, Comm. Math. Phys., 121 (1989), pp. 557–601.

[BG94] 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), pp. 1117–1143.

[Deu79] P. Deuflhard, *A study of extrapolation methods based on multistep schemes without parasitic solutions*, Z. Angew. Math. Phys., 30 (1979), pp. 177–189.

[GGMV92] L. Galgani, A. Giorgilli, A. Martinoli, and S. Vanzini, *On the problem of energy equipartition for large systems of the Fermi-Pasta-Ulam type: Analytical and numerical estimates*, Phys. D, 59 (1992), pp. 334–348.

[GSS99] B. García-Archilla, J.M. Sanz-Serna, and R. D. Skeel, *Long-time-step methods for oscillatory differential equations*, SIAM J. Sci. Comput., 20 (1998), pp. 930–963.

[Gau61] W. Gautschi, *Numerical integration of ordinary differential equations based on trigonometric polynomials*, Numer. Math., 3 (1961), pp. 381–397.

[HaL97] E. Hairer and Ch. Lubich, *The life-span of backward error analysis for numerical integrators*, Numer. Math., 76 (1997), pp. 441–462.

[HaL99] E. Hairer and Ch. Lubich, *Energy conservation by Störmer-type numerical integrators*, in Proceedings of 18th Dundee Biennial Conference, D. F. Griffiths and G. A. Watson, eds., Numerical Analysis 1999, CRC Press, Boca Raton, FL, 2000, pp. 169–190.

[HaW96] E. Hairer and G. Wanner, *Analysis by Its History*, Undergrad. Texts Math., Springer-Verlag, New York, 1996.

[HoL99] M. Hochbruck and Ch. Lubich, *A Gautschi-type method for oscillatory second-order differential equations*, Numer. Math. 83 (1999), pp. 403–426.

[Rei98] S. Reich, *Dynamical Systems, Numerical Integration, and Exponentially Small Estimates*, Habilitationsschrift, FU Berlin, 1998.