

MODIFIED TRIGONOMETRIC INTEGRATORS*

ROBERT I. MCLACHLAN[†] AND ARI STERN[‡]

Abstract. We study modified trigonometric integrators, which generalize the popular class of trigonometric integrators for highly oscillatory Hamiltonian systems by allowing the fast frequencies to be modified. Among all methods of this class, we show that the IMEX (implicit-explicit) method, which is equivalent to applying the midpoint rule to the fast, linear part of the system and the leapfrog (Störmer/Verlet) method to the slow, nonlinear part, is distinguished by the following properties: (i) it is symplectic; (ii) it is free of artificial resonances; (iii) it is the unique method that correctly captures slow energy exchange to leading order; (iv) it conserves the total energy and a modified oscillatory energy up to second order; (v) it is uniformly second-order accurate in the slow components; and (vi) it has the correct magnitude of deviations of the fast oscillatory energy, which is an adiabatic invariant. These theoretical results are supported by numerical experiments on the Fermi–Pasta–Ulam problem and indicate that the IMEX method, for these six properties, dominates the class of modified trigonometric integrators.

Key words. trigonometric integrators, geometric integrators, highly oscillatory problems, IMEX, Fermi–Pasta–Ulam

AMS subject classifications. 65P10, 70K70

DOI. 10.1137/130921118

1. Introduction.

1.1. Overview. Over the past two decades, there has been considerable interest in so-called *geometric numerical integrators*, particularly symplectic integrators for Hamiltonian systems [10, 12]. In contrast to general-purpose numerical integrators, geometric integrators are designed especially to be applied to systems with some additional underlying structure (symmetries, invariants, etc.) that the algorithm must preserve exactly, at least up to round-off error.

While there have been many successes in this area, the integration of highly oscillatory Hamiltonian systems—which feature both stiff, linear forces and soft, nonlinear forces—has remained persistently difficult, due to the simultaneous presence of fast and slow time scales. Such systems are especially prevalent, for instance, in molecular dynamics, where one must contend with strong, short-range bonding forces, as well as weak, long-range electrostatic forces.

One of the main advances has been the development and analysis of *trigonometric integrators*, which are numerical methods designed especially to integrate certain highly oscillatory systems. However, these methods have certain drawbacks: in particular, there is a trade-off between numerical stability and consistency with respect to certain dynamical features, such as the emergent multiscale phenomenon of *slow energy exchange* and the near-preservation of adiabatic invariants.

In this paper, we show that *modified trigonometric integrators*—that is, trigonometric integrators with modified oscillatory frequencies—provide a way around this

*Received by the editors May 15, 2013; accepted for publication (in revised form) March 24, 2014; published electronically June 10, 2014.

<http://www.siam.org/journals/sinum/52-3/92111.html>

[†]IFS, Massey University, Palmerston North, 5301 New Zealand (r.mclachlan@massey.ac.nz). This author was supported by a grant from the Marsden Fund of the Royal Society of New Zealand.

[‡]Department of Mathematics, Washington University in St. Louis, St. Louis, MO 63130 (astern@math.wustl.edu). This author gratefully acknowledges partial travel support from Reinout Quispel, along with an AMS–Simons travel grant and an Oberwolfach Leibniz grant.

obstacle. Naively, one might expect that perturbing the frequency would increase the error, so it may seem counterintuitive to suggest that this can actually improve numerical performance. Yet, we show that this is indeed the case: allowing for the frequency to be modified provides an additional degree of freedom, which makes it possible to sidestep the aforementioned trade-off between stability and multiscale structure preservation. Specifically, we show that a particular, unique choice of modified frequency yields an integrator that is both stable and structure preserving, and this is precisely the implicit-explicit (IMEX) method of Stern and Grinspun [15].

1.2. The numerical challenge of fast oscillations. Consider a prototypical highly oscillatory problem, given by the second-order equation

$$(1.1) \quad \ddot{q} + \Omega^2 q = g(q),$$

where $q(t) \in \mathbb{R}^d$ is a trajectory, $\Omega = \begin{pmatrix} 0 & 0 \\ 0 & \omega_I \end{pmatrix}$ is a $d \times d$ matrix with constant fast frequency $\omega \gg 1$, and $g: \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a conservative nonlinear force, so that $g = -\nabla U$ for some scalar potential $U: \mathbb{R}^d \rightarrow \mathbb{R}$. Equation (1.1) can also be written as a first-order system on $(q, p) \in \mathbb{R}^{2d}$,

$$(1.2) \quad \begin{aligned} \dot{q} &= p, \\ \dot{p} &= -\Omega^2 q + g(q), \end{aligned}$$

which are Hamilton's equations for the separable Hamiltonian

$$H(q, p) = \frac{1}{2} \|p\|^2 + \frac{1}{2} \|\Omega q\|^2 + U(q).$$

Due to this underlying Hamiltonian structure, it is desirable to use a symplectic integrator to obtain numerical solutions to (1.1)–(1.2).

One of the most popular, widely used symplectic integrators is the *Störmer/Verlet* (or *leapfrog*) method, which discretizes (1.1) by the centered finite-difference equation

$$(1.3) \quad \frac{q_{n+1} - 2q_n + q_{n-1}}{h^2} + \Omega^2 q_n = g(q_n),$$

where h denotes the time step size. An equivalent approximation for the first-order system (1.2) is given by the symmetric algorithm

$$\begin{aligned} p_{n+1/2} &= p_n + \frac{1}{2} h [-\Omega^2 q_n + g(q_n)], \\ q_{n+1} &= q_n + h p_{n+1/2}, \\ p_{n+1} &= p_{n+1/2} + \frac{1}{2} h [-\Omega^2 q_{n+1} + g(q_{n+1})], \end{aligned}$$

which is sometimes called the *velocity Verlet method*. Note that, if (1.3) is used to compute a numerical trajectory $(\dots, q_{n-1}, q_n, q_{n+1}, \dots)$, then we can still recover p_n , after the fact, by taking $p_n = \frac{q_{n+1} - q_{n-1}}{2h}$. This method corresponds to splitting the Hamiltonian into kinetic and potential components, $H(q, p) = T(p) + V(q)$, where

$$T(p) = \frac{1}{2} \|p\|^2, \quad V(q) = \frac{1}{2} \|\Omega q\|^2 + U(q),$$

and alternating between the purely kinetic flow of $T(p)$ and the purely potential flow of $V(q)$. This is an example of a *splitting method* (see McLachlan and Quispel

[13]), and since the flows of $T(p)$ and $V(q)$ are each Hamiltonian, the composition $(q_n, p_n) \mapsto (q_{n+1}, p_{n+1})$ is a symplectic map.

Despite these desirable geometric properties, however, the Störmer/Verlet method cannot integrate highly oscillatory systems efficiently. As an explicit method, it remains stable only for time steps on the order $h = \mathcal{O}(\omega^{-1})$; in particular, when $g = 0$, we have the linear stability condition $h\omega \leq 2$. Therefore, to integrate over a time interval of fixed size, the method requires $\mathcal{O}(\omega)$ time steps, and hence $\mathcal{O}(\omega)$ evaluations of the nonlinear force g , which becomes prohibitively expensive for large ω .

A typical implicit method encounters similar difficulties. For example, the *implicit midpoint method* discretizes (1.2) by the one-step algorithm

$$\begin{aligned} q_{n+1} &= q_n + h \left(\frac{p_n + p_{n+1}}{2} \right), \\ p_{n+1} &= p_n + h \left[-\Omega^2 \left(\frac{q_n + q_{n+1}}{2} \right) + g \left(\frac{q_n + q_{n+1}}{2} \right) \right], \end{aligned}$$

which is equivalent to the centered finite-difference scheme

$$\begin{aligned} (1.4) \quad \frac{q_{n+1} - 2q_n + q_{n-1}}{h^2} + \Omega^2 \left(\frac{q_{n+1} + 2q_n + q_{n-1}}{4} \right) \\ = \frac{1}{2} g \left(\frac{q_n + q_{n+1}}{2} \right) + \frac{1}{2} g \left(\frac{q_{n-1} + q_n}{2} \right) \end{aligned}$$

for the second-order equation (1.1). While this method is linearly unconditionally stable, it requires a nonlinear solve at each time step, since q_{n+1} appears inside the nonlinear force g . However, a numerical solver (e.g., Newton's method) will require several evaluations of g at each time step, so this method is also computationally expensive.

The failure of these traditional symplectic integrators has motivated the development of numerical methods designed especially for highly oscillatory Hamiltonian systems. The goal of this research has been to obtain efficient, explicit integrators that are stable and accurate for large time steps h . By “large time steps,” we mean that the step size can be chosen independently of the fast frequency ω , so that $h^{-1} = \mathcal{O}(1)$ as $\omega \rightarrow \infty$. Therefore, in sharp contrast to the Störmer/Verlet method, such integrators require only $\mathcal{O}(1)$ evaluations of the nonlinear force, rather than $\mathcal{O}(\omega)$.

1.3. Trigonometric integrators. Trigonometric integrators are designed to integrate (1.1)–(1.2) exactly when the nonlinear force g vanishes, i.e., when the system reduces to a harmonic oscillator. Let ψ and ϕ be a pair of even, real-valued *filter functions* satisfying $\psi(0) = \phi(0) = 1$, and denote $\Psi = \psi(h\Omega)$, $\Phi = \phi(h\Omega)$, and $g_n = g(\Phi q_n)$. Then the *trigonometric integrator* corresponding to the filters ψ, ϕ , is defined by the difference equation

$$(1.5) \quad q_{n+1} - 2\cos(h\Omega)q_n + q_{n-1} = h^2\Psi g_n.$$

We can extend this to a symmetric, one-step method for $(q, p) \in \mathbb{R}^{2d}$ by introducing a new filter function ψ_1 , which satisfies $\psi(\xi) = \text{sinc}(\xi)\psi_1(\xi)$, and taking $\Psi_1 = \psi_1(h\Omega)$.

TABLE 1
Filter functions corresponding to various trigonometric integrators.

Method	$\psi(\xi)$	$\phi(\xi)$	Reference
A	$\text{sinc}^2(\frac{1}{2}\xi)$	1	Gautschi [7]
B	$\text{sinc}(\xi)$	1	Deuffhard [3]
C	$\text{sinc}^2(\xi)$	$\text{sinc}(\xi)$	García-Archilla, Sanz-Serna, and Skeel [6]
D	$\text{sinc}^2(\frac{1}{2}\xi)$	$\text{sinc}(\xi)(1 + \frac{1}{3}\text{sinc}^2(\frac{1}{2}\xi))$	Hochbruck and Lubich [11]
E	$\text{sinc}^2(\xi)$	1	Hairer and Lubich [9]
G	$\text{sinc}^3(\xi)$	$\text{sinc}(\xi)$	Grimm and Hochbruck [8]

We then obtain a velocity Verlet-like algorithm,

$$\begin{aligned}
 p_n^+ &= p_n + \frac{1}{2}h\Psi_1 g_n, \\
 \begin{pmatrix} q_{n+1} \\ p_{n+1}^- \end{pmatrix} &= \begin{pmatrix} \cos(h\Omega) & h \text{sinc}(h\Omega) \\ -\Omega \sin(h\Omega) & \cos(h\Omega) \end{pmatrix} \begin{pmatrix} q_n \\ p_n^+ \end{pmatrix}, \\
 p_{n+1} &= p_{n+1}^- + \frac{1}{2}h\Psi_1 g_{n+1}.
 \end{aligned}$$

Similarly to Störmer/Verlet, if (1.5) is used to compute a numerical trajectory $(\dots, q_{n-1}, q_n, q_{n+1}, \dots)$, then p_n can be recovered by taking $\text{sinc}(h\Omega)p_n = \frac{q_{n+1} - q_{n-1}}{2h}$, as long as $\text{sinc}(h\omega) \neq 0$. Whatever the choice of ψ and ϕ , these methods reduce to Störmer/Verlet in the case $\omega = 0$, and to the exact solution of the harmonic oscillator in the case $g = 0$.

One of the simplest trigonometric integrators is the *Deuffhard/impulse method*,¹ which corresponds to the choice of filters $\psi(\xi) = \text{sinc}(\xi)$ (i.e., $\psi_1 = 1$) and $\phi = 1$. In this case, the trigonometric integrator corresponds to a splitting method: the Hamiltonian is split as $H(q, p) = H_{\text{fast}}(q, p) + U(q)$, where $H_{\text{fast}}(q, p) = \frac{1}{2}(\|p\|^2 + \|\Omega q\|^2)$. While the Deuffhard/impulse method has many desirable properties, it has one major drawback: spurious numerical resonances arise when $h\omega$ is close to a nonzero integer multiple of π , causing a loss of stability and accuracy. This resonance instability causes serious problems whenever $h\omega \geq \pi$, so effectively, the method forces us to choose $h = \mathcal{O}(\omega^{-1})$, just like the Störmer/Verlet method, making it unsuitable for integration with long time steps. (Similar resonance phenomena also plague other impulse-type methods, including multiple-time-stepping methods; cf. Biesiadecki and Skeel [1].)

In contrast to the “unfiltered” Deuffhard/impulse method, other trigonometric integrators use ϕ to filter (or “mollify”) the slow force, so as to lessen the problem of resonance instability. *Mollified impulse methods* allow the filter ϕ to be chosen arbitrarily, and then take $\psi(\xi) = \text{sinc}(\xi)\phi(\xi)$, i.e., $\psi_1 = \phi$. Like the Deuffhard/impulse method (which corresponds to the special case $\phi = 1$), these are also Hamiltonian splitting methods, where the potential U is replaced by the mollified potential $\bar{U}(q) = U(\Phi q)$. Consequently, mollified impulse methods are symplectic; in fact, it is straightforward to show that a trigonometric integrator is symplectic if and only if $\psi(\xi) = \text{sinc}(\xi)\phi(\xi)$, i.e., it is a mollified impulse method.

Various trigonometric integrators, corresponding to different choices of the filter functions ψ and ϕ , have appeared in the literature, and are summarized in Table 1.

¹Christian Lubich pointed out to us that, although the Deuffhard and impulse methods are distinct in general, they happen to coincide in the case of highly oscillatory problems of this type. This is the reason behind the slash in the name “Deuffhard/impulse.”

The alphabetical labels for these methods (A–E and G) follow the convention of Hairer, Lubich, and Wanner [10], which has since been adopted by several other authors. (We have omitted “method F,” from [10], which is a two-force method rather than a trigonometric integrator.) Of these, note that only method B (Deuffhard/impulse) and method C (mollified impulse) satisfy the symplecticity condition $\psi(\xi) = \text{sinc}(\xi)\phi(\xi)$.

1.4. Modulated Fourier expansion and slow exchange. The *modulated Fourier expansion* is a powerful technique for analyzing the dynamics of highly oscillatory systems, as well as the numerical behavior of trigonometric integrators for such systems. We give only a brief summary here; for a detailed treatment, see Hairer and Lubich [9], Hairer, Lubich, and Wanner [10].

Suppose that $q(t)$ is a solution of the highly oscillatory system (1.1). To separate out its fast- and slow-scale features, we approximate $q(t)$ asymptotically by a trajectory $x(t)$ of the form

$$(1.6) \quad x(t) = y(t) + e^{i\omega t} z(t) + e^{-i\omega t} \bar{z}(t),$$

where $y(t)$ is real valued and $z(t)$ is complex valued. Assuming that the energy of $x(t)$ is bounded on the time interval of interest, this implies that $z(t) = \mathcal{O}(\omega^{-1})$. Next, we can decompose $x = (x_0, x_1)$, $y = (y_0, y_1)$, and $z = (z_0, z_1)$, according to the blocks of Ω . Plugging $x(t)$ into (1.1), Taylor expanding $g(x)$ around y , and matching the terms on both sides up to $\mathcal{O}(\omega^{-3})$ yields the system of equations

$$(1.7) \quad \begin{aligned} \ddot{y}_0 &= g_0(y_0, \omega^{-2} g_1(y_0, 0)) + \frac{\partial^2 g_0}{\partial x_1^2}(y_0, 0)(z_1, \bar{z}_1), \\ 2i\omega \dot{z}_1 &= \frac{\partial g_1}{\partial x_1}(y_0, 0)z_1. \end{aligned}$$

(The y_1 and z_0 components can both be eliminated up to this order of accuracy.) Here, y_0 evolves on the time scale $\mathcal{O}(1)$ and describes the nonstiff dynamics of the system, while z_1 evolves on the time scale $\mathcal{O}(\omega)$ and corresponds to a multiscale phenomenon known as *slow energy exchange*. If $I_j = \frac{1}{2}p_{1,j}^2 + \frac{1}{2}\omega^2 q_{1,j}^2$ is the energy in the j th stiff component of the system, then it can be shown that, up to $\mathcal{O}(\omega^{-1})$, we have $I_j \approx 2\omega^2 |z_{1,j}|^2$. Here, we have split $q = (q_0, q_1)$ and $p = (p_0, p_1)$ into non-stiff and stiff blocks, as above, and $q_{1,j}$, $p_{1,j}$, $z_{1,j}$ denote the j th components of the corresponding vectors q_1 , p_1 , z_1 . It follows that the evolution of z_1 describes the slow exchange of energy between the stiff components, coupled through the nonlinear force. Moreover, the total stiff energy $I = \sum_j I_j \approx 2\omega^2 \|z_1\|^2$ is an adiabatic invariant, since

$$\frac{d}{dt} 2\omega^2 \|z_1\|^2 = 4\omega^2 \text{Re}\langle z_1, \dot{z}_1 \rangle = \mathcal{O}(\omega^{-1}),$$

and therefore $\dot{I} = \mathcal{O}(\omega^{-1})$. Hence, deviations in I are also $\mathcal{O}(\omega^{-1})$ over a fixed time interval.

A similar technique can be applied to analyze numerical behavior. For a trigonometric integrator with time step size h , the numerical trajectory q_n can be approximated asymptotically by $x_h(nh)$, where

$$(1.8) \quad x_h(t) = y_h(t) + e^{i\omega t} z_h(t) + e^{-i\omega t} \bar{z}_h(t).$$

As we did for the continuous dynamics, we can plug this ansatz into (1.5) and match

terms, obtaining a system of equations,

$$(1.9) \quad \begin{aligned} \delta_h^2 y_{h,0} &= g_0(y_{h,0}, \gamma \omega^{-2} g_1(y_{h,0}, 0)) + \beta \frac{\partial^2 g_0}{\partial x_1^2}(y_{h,0}, 0)(z_{h,1}, \bar{z}_{h,1}), \\ 2i\omega \dot{z}_{h,1} &= \alpha \frac{\partial g_1}{\partial x_1}(y_{h,0}, 0) z_{h,1}, \end{aligned}$$

which hold up to $\mathcal{O}(\omega^{-3})$. Here, δ_h^2 denotes the second finite-difference operator, defined by

$$\delta_h^2 y_{h,0}(t) = \frac{y_{h,0}(t+h) - 2y_{h,0}(t) + y_{h,0}(t-h)}{h^2},$$

while the constants α , β , and γ are given by

$$\alpha = \frac{\psi(h\omega)\phi(h\omega)}{\text{sinc}(h\omega)}, \quad \beta = \phi(h\omega)^2, \quad \gamma = \frac{\psi(h\omega)\phi(h\omega)}{\text{sinc}^2(\frac{1}{2}h\omega)}.$$

Comparing (1.7) and (1.9), it follows that the dynamics of $z_{h,1}$ are consistent with those for z_1 only if $\alpha = 1$. Moreover, to fully capture the coupled dynamics between y_0 and z_1 , one would also require $\beta = 1$ and $\gamma = 1$.

Of the methods listed in Table 1, only Method B, the Deuffhard/impulse method, satisfies $\alpha = 1$. However, as discussed previously, the resonance instability of this method makes it practically impossible to take large time steps. Hence, we cannot hope to model the slow-energy exchange accurately, using a trigonometric integrator, unless we sacrifice either stability or efficiency. A more fundamental problem is that $\alpha \neq \gamma$ in general, so even if we are willing to make the aforementioned trade-off, it is impossible for a trigonometric integrator to satisfy $\alpha = \beta = \gamma = 1$.

Multiforce methods provide one way around this obstacle, but as their name suggests, they require multiple evaluations of the slow force g per time step. However, we show that there is another way around this obstacle: by modifying the fast frequency ω , it is possible for a stable, efficient method to achieve $\alpha = \beta = \gamma \approx 1$, with only a single evaluation of g per time step.

Remark. Strictly speaking, (1.6) and (1.8) contain only the principal (i.e., leading-order) terms of the modulated Fourier expansion. While the constant and $e^{\pm i\omega t}$ terms are sufficient to describe slow energy exchange, other properties—including long-time conservation of the total and oscillatory energies—require further expansion in the higher-order terms $e^{\pm 2i\omega t}$, $e^{\pm 3i\omega t}$, etc. Again, we refer the reader to Hairer and Lubich [9], Hairer, Lubich, and Wanner [10] for a full account.

1.5. Overview of results. We begin, in section 2, by defining modified trigonometric integrators, and by giving a few examples. We then show how the modulated Fourier expansion can be applied to these methods, and we use this to derive consistency conditions for slow energy exchange. The main result of this section, Theorem 2.1, shows that modified trigonometric integrators can indeed satisfy the full consistency condition $\alpha = \beta = \gamma = 1$; in fact, we prove that there is a unique modified trigonometric integrator which does so, coinciding with the IMEX method of Stern and Grinspun [15]. Furthermore, Theorem 2.2 shows that IMEX conserves total energy and a modified oscillatory energy up to $\mathcal{O}(h^2)$.

Section 3 presents the results of several numerical experiments for the widely studied and dynamically rich Fermi–Pasta–Ulam test problem. We compare the numerical behavior of the trigonometric integrators listed in Table 1 with that of the

IMEX modified trigonometric integrator. These experiments demonstrate the trade-off between stability, consistency, and accuracy inherent to standard trigonometric integrators. By contrast, the IMEX modified method performs well in all of these experiments, without any observed trade-off, as predicted by Theorems 2.1 and 2.2.

Finally, one of these numerical experiments reveals that, although the total oscillatory energy $I = \sum_j I_j$ is well-conserved by all of the methods considered (both modified and unmodified), the integrators vary considerably with respect to the magnitude of deviations in this adiabatic invariant. In section 4, we analyze the deviation in total oscillatory energy by examining higher-order terms in the modulated Fourier expansion. This analysis provides a theoretical explanation for the behavior observed in the numerical experiments.

Relationship to previous work. As mentioned above, the IMEX method was introduced for highly oscillatory problems in Stern and Grinspun [15]. This earlier paper focused primarily on the variational, symplectic, and stability properties of IMEX, and on its comparison with multiple-time-stepping methods (as opposed to trigonometric integrators). It was observed that IMEX can be viewed as a Deuffhard/impulse method with modified frequency, implying the partial consistency condition $\alpha = 1$ for slow energy exchange [15, Theorem 4.1]. Theorem 2.1 is a substantial strengthening of this consistency result; the other results and numerical experiments presented in the current paper are independent of those in [15].

2. Modified trigonometric integrators.

2.1. Basic definitions. A *modified trigonometric integrator* for the highly oscillatory system (1.1) is defined by the second-order difference equation

$$(2.1) \quad q_{n+1} - 2\cos(h\tilde{\Omega})q_n + q_{n-1} = h^2\Psi g_n,$$

where $\tilde{\Omega} = \begin{pmatrix} 0 & 0 \\ 0 & \tilde{\omega}I \end{pmatrix}$ and $\tilde{\omega}$ is called the *modified frequency*. If ψ and ϕ are even, real-valued filter functions satisfying $\psi(0) = \phi(0) = 1$, we now take $\Psi = \psi(h\tilde{\Omega})$ and $\Phi = \phi(h\tilde{\Omega})$, while as before, we denote $g_n = g(\Phi q_n)$.

Although Ω is generally singular, we commit a slight abuse of notation by taking $\Omega^{-1}\tilde{\Omega}$ to mean the matrix $\begin{pmatrix} I & 0 \\ 0 & (\tilde{\omega}/\omega)I \end{pmatrix}$. Letting $\Psi = \Omega^{-1}\tilde{\Omega} \operatorname{sinc}(h\tilde{\Omega})\Psi_1$, we consider the following symmetric, one-step algorithm:

$$\begin{aligned} p_n^+ &= p_n + \frac{1}{2}h\Psi_1 g_n, \\ \begin{pmatrix} q_{n+1} \\ p_{n+1}^- \end{pmatrix} &= \begin{pmatrix} \cos(h\tilde{\Omega}) & h\Omega^{-1}\tilde{\Omega} \operatorname{sinc}(h\tilde{\Omega}) \\ -\Omega \sin(h\tilde{\Omega}) & \cos(h\tilde{\Omega}) \end{pmatrix} \begin{pmatrix} q_n \\ p_n^+ \end{pmatrix}, \\ p_{n+1} &= p_{n+1}^- + \frac{1}{2}h\Psi_1 g_{n+1}. \end{aligned}$$

As with standard trigonometric integrators, this method is symplectic when the filters satisfy $\psi_1 = \phi$; since this gives $\Psi_1 g_n = -\nabla \bar{U}(q_n)$, by the chain rule, and hence the integrator corresponds to a splitting method for the modified Hamiltonian. This symplecticity condition can also be written as $\psi(h\tilde{\omega}) = \omega^{-1}\tilde{\omega} \operatorname{sinc}(h\tilde{\omega})\phi(h\tilde{\omega})$.

If (2.1) is used to compute a numerical trajectory $(\dots, q_{n-1}, q_n, q_{n+1}, \dots)$, then the p_n can be recovered by taking $\Omega^{-1}\tilde{\Omega} \operatorname{sinc}(h\tilde{\Omega})p_n = \frac{q_{n+1} - q_{n-1}}{2h}$, as long as $\operatorname{sinc}(h\tilde{\omega}) \neq 0$; note that this is slightly different from the previous expression for a standard trigonometric integrator. If we define the modified momentum $\tilde{p}_n = \Omega^{-1}\tilde{\Omega}p_n$, then

it follows that $\text{sinc}(h\tilde{\Omega})\tilde{p}_n = \Omega^{-1}\tilde{\Omega}\text{sinc}(h\tilde{\Omega})p_n = \frac{q_{n+1}-q_{n-1}}{2h}$. Hence, the numerical algorithm in $(q, \tilde{p})$ corresponds to

$$\begin{aligned}\tilde{p}_n^+ &= \tilde{p}_n + \frac{1}{2}h\tilde{\Psi}_1g_n, \\ \begin{pmatrix} q_{n+1} \\ \tilde{p}_{n+1}^+ \end{pmatrix} &= \begin{pmatrix} \cos(h\tilde{\Omega}) & h\text{sinc}(h\tilde{\Omega}) \\ -\tilde{\Omega}\sin(h\tilde{\Omega}) & \cos(h\tilde{\Omega}) \end{pmatrix} \begin{pmatrix} q_n \\ \tilde{p}_n^+ \end{pmatrix}, \\ \tilde{p}_{n+1} &= \tilde{p}_{n+1}^+ + \frac{1}{2}h\tilde{\Psi}_1g_{n+1},\end{aligned}$$

where $\tilde{\Psi}_1 = \Omega^{-1}\tilde{\Omega}\Psi_1$, which implies $\Psi = \text{sinc}(h\tilde{\Omega})\tilde{\Psi}_1$. This is precisely a standard trigonometric integrator for the modified frequency $\tilde{\omega}$.

2.2. Examples. The first example of a modified trigonometric integrator is simply a standard trigonometric integrator, where we make the trivial choice of modified frequency $\tilde{\omega} = \omega$.

A more interesting, nontrivial example is the Störmer/Verlet method. Observe that the finite-difference scheme (1.3) can be rewritten as

$$q_{n+1} - 2(I - \tfrac{1}{2}h^2\Omega^2)q_n + q_{n-1} = h^2g(q_n).$$

If we choose $\tilde{\omega}$ such that $\sin(\frac{1}{2}h\tilde{\omega}) = \frac{1}{2}h\omega$, then it follows that $1 - \frac{1}{2}h^2\Omega^2 = 1 - 2\sin^2(\frac{1}{2}h\tilde{\Omega}) = \cos(h\tilde{\Omega})$, and therefore

$$q_{n+1} - 2\cos(h\tilde{\Omega})q_n + q_{n-1} = h^2g(q_n).$$

Hence, the Störmer/Verlet method is a modified trigonometric integrator with the above choice of $\tilde{\omega}$, and with the filters $\psi = \phi = 1$. It should be observed, though, that $\sin(\frac{1}{2}h\tilde{\omega}) = \frac{1}{2}h\omega$ has no solution when $\frac{1}{2}h\omega > 1$; we are again limited by the linear stability condition $h\omega \leq 2$, as in section 1.2.

Note that, although (2.1) coincides with Störmer/Verlet, the approximation of p_n is different from that used in velocity Verlet. From $\sin(\frac{1}{2}h\tilde{\omega}) = \frac{1}{2}h\omega$, we obtain

$$\frac{\tilde{\omega}}{\omega}\text{sinc}(h\tilde{\omega}) = \frac{h\tilde{\omega}\text{sinc}(h\tilde{\omega})}{h\omega} = \frac{\sin(h\tilde{\omega})}{h\omega} = \frac{2\sin(\frac{1}{2}h\tilde{\omega})\cos(\frac{1}{2}h\tilde{\omega})}{h\omega} = \cos(\tfrac{1}{2}h\tilde{\omega}).$$

Substituting this into $\Omega^{-1}\tilde{\Omega}\text{sinc}(h\tilde{\Omega})p_n = \frac{q_{n+1}-q_{n-1}}{2h}$, it follows that the momentum p_n satisfies $\cos(\frac{1}{2}h\tilde{\Omega})p_n = \frac{q_{n+1}-q_{n-1}}{2h}$.

A third, particularly interesting example is the IMEX integrator first suggested by Zhang and Skeel [16] as a “cheap” version of the implicit midpoint method, and more recently introduced and analyzed by Stern and Grinspun [15] as an attractive method in its own right for highly oscillatory problems. (An essentially similar method has also been applied to the linear Schrödinger equation; cf. Debussche and Faou [2].) Combining the left-hand side of the implicit midpoint method (1.4) with the right-hand side of the Störmer/Verlet method (1.3), and multiplying both sides by h^2 , we get the IMEX method,

$$(q_{n+1} - 2q_n + q_{n-1}) + (\tfrac{1}{2}h\Omega)^2(q_{n+1} + 2q_n + q_{n-1}) = h^2g(q_n),$$

which is only linearly implicit, hence avoids the difficulty of solving a nonlinear equation for q_{n+1} . Now, if we choose $\tilde{\omega}$ such that $\tan(\frac{1}{2}h\tilde{\omega}) = \frac{1}{2}h\omega$, then this becomes

$$(q_{n+1} - 2q_n + q_{n-1}) + \tan^2(\tfrac{1}{2}h\tilde{\Omega})(q_{n+1} + 2q_n + q_{n-1}) = h^2g(q_n),$$

or

$$q_{n+1} - 2 \cos(h\tilde{\Omega})q_n + q_{n-1} = h^2 \cos^2(\tfrac{1}{2}h\tilde{\Omega})g(q_n).$$

Hence, the IMEX method can be reframed as a modified trigonometric method, with the modified frequency $\tilde{\omega}$ satisfying $\tan(\tfrac{1}{2}h\tilde{\omega}) = \tfrac{1}{2}h\omega$, and with the filters $\psi(\xi) = \cos^2(\tfrac{1}{2}\xi)$ and $\phi = 1$. In contrast to Störmer/Verlet, it is always possible to solve for this $\tilde{\omega}$, since $\arctan$ (unlike $\arcsin$) is defined on the entire real line. Note that the IMEX method is also symplectic, since it satisfies the aforementioned condition

$$\omega^{-1}\tilde{\omega} \operatorname{sinc}(h\tilde{\omega})\phi(h\tilde{\omega}) = \frac{\frac{1}{2}h\tilde{\omega}}{\frac{1}{2}h\omega} \operatorname{sinc}(h\tilde{\omega}) = \frac{\frac{1}{2}\sin(h\tilde{\omega})}{\tan(\frac{1}{2}h\tilde{\omega})} = \cos^2(\tfrac{1}{2}h\tilde{\omega}) = \psi(h\tilde{\omega}).$$

See Stern and Grinspun [15] for a discussion of IMEX as a splitting method, which also implies its symplecticity.

2.3. Modulated Fourier expansion and slow exchange. We have seen that a modified trigonometric integrator has the same form as a standard trigonometric integrator with frequency $\tilde{\omega}$ (modulo the choice of $\tilde{p}_n = \Omega^{-1}\tilde{\Omega}p_n$ instead of p_n). Therefore, applying the modulated Fourier expansion, we get

$$(2.2) \quad \begin{aligned} \delta_h^2 y_{h,0} &= g_0(y_{h,0}, \tilde{\gamma}\tilde{\omega}^{-2}g_1(y_{h,0}, 0)) + \tilde{\beta} \frac{\partial^2 g_0}{\partial x_1^2}(y_{h,0}, 0)(z_{h,1}, \bar{z}_{h,1}), \\ 2i\tilde{\omega}\dot{z}_{h,1} &= \tilde{\alpha} \frac{\partial g_1}{\partial x_1}(y_{h,0}, 0)z_{h,1}, \end{aligned}$$

where

$$\tilde{\alpha} = \frac{\psi(h\tilde{\omega})\phi(h\tilde{\omega})}{\operatorname{sinc}(h\tilde{\omega})}, \quad \tilde{\beta} = \phi(h\tilde{\omega})^2, \quad \tilde{\gamma} = \frac{\psi(h\tilde{\omega})\phi(h\tilde{\omega})}{\operatorname{sinc}^2(\frac{1}{2}h\tilde{\omega})}.$$

Now, if we define $\tilde{I}_j = \frac{1}{2}|\tilde{p}_{1,j}|^2 + \frac{1}{2}\tilde{\omega}^2|q_{1,j}|^2$, then we have previously seen that $\tilde{I}_j \approx 2\tilde{\omega}^2|z_{1,j}|^2$. However, we are interested not in the behavior of $\tilde{I}_j$, but in that of the original stiff energies I_j . Since $\tilde{p}_{1,j} = (\tilde{\omega}/\omega)p_{1,j}$, we get

$$I_j = \frac{1}{2}|p_{1,j}|^2 + \frac{1}{2}\omega^2|q_{1,j}|^2 = \frac{\omega^2}{\tilde{\omega}^2} \left(\frac{1}{2}|\tilde{p}_{1,j}|^2 + \frac{1}{2}\tilde{\omega}^2|q_{1,j}|^2 \right) = \frac{\omega^2}{\tilde{\omega}^2} \tilde{I}_j,$$

and therefore $I_j \approx 2\omega^2|z_{1,j}|^2$.

It follows that, if (2.2) is consistent with (1.7), then the modified trigonometric integrator will be consistent for the corresponding energy exchange behavior. To compare these, let us rewrite (2.2) as

$$\begin{aligned} \delta_h^2 y_{h,0} &= g_0(y_{h,0}, \gamma\omega^{-2}g_1(y_{h,0}, 0)) + \beta \frac{\partial^2 g_0}{\partial x_1^2}(y_{h,0}, 0)(z_{h,1}, \bar{z}_{h,1}), \\ 2i\omega\dot{z}_{h,1} &= \alpha \frac{\partial g_1}{\partial x_1}(y_{h,0}, 0)z_{h,1}, \end{aligned}$$

where $\alpha = (\omega/\tilde{\omega})\tilde{\alpha}$, $\beta = \tilde{\beta}$, and $\gamma = (\omega/\tilde{\omega})^2\tilde{\gamma}$, i.e.,

$$\alpha = \frac{\omega\psi(h\tilde{\omega})\phi(h\tilde{\omega})}{\tilde{\omega}\operatorname{sinc}(h\tilde{\omega})}, \quad \beta = \phi(h\tilde{\omega})^2, \quad \gamma = \frac{\omega^2\psi(h\tilde{\omega})\phi(h\tilde{\omega})}{\tilde{\omega}^2\operatorname{sinc}^2(\frac{1}{2}h\tilde{\omega})}.$$

Hence, consistency will require $\alpha = \beta = \gamma = 1$. We now arrive at the main result of this section.

THEOREM 2.1. *The IMEX method is the unique modified trigonometric integrator satisfying $\alpha = \beta = \gamma = 1$.*

Proof. Clearly $\beta = 1$ if and only if $\phi = 1$. Substituting this into $\alpha = 1$ and solving for the filter ψ , we get $\psi(h\tilde{\omega}) = (\tilde{\omega}/\omega) \operatorname{sinc}(h\tilde{\omega})$. Therefore,

$$\gamma = \frac{\omega \operatorname{sinc}(h\tilde{\omega})}{\tilde{\omega} \operatorname{sinc}^2(\frac{1}{2}h\tilde{\omega})} = \frac{\omega \operatorname{sinc}(\frac{1}{2}h\tilde{\omega}) \cos(\frac{1}{2}h\tilde{\omega})}{\tilde{\omega} \operatorname{sinc}^2(\frac{1}{2}h\tilde{\omega})} = \frac{\omega \cos(\frac{1}{2}h\tilde{\omega})}{\tilde{\omega} \operatorname{sinc}(\frac{1}{2}h\tilde{\omega})} = \frac{\frac{1}{2}h\omega}{\tan(\frac{1}{2}h\tilde{\omega})}.$$

Hence, for $\gamma = 1$, the modified frequency must satisfy $\tan(\frac{1}{2}h\tilde{\omega}) = \frac{1}{2}h\omega$. Finally, applying this to the prior equation for ψ , we get

$$\psi(h\tilde{\omega}) = \frac{\tilde{\omega} \operatorname{sinc}(h\tilde{\omega})}{\omega} = \frac{\sin(h\tilde{\omega})}{h\omega} = \frac{2 \sin(\frac{1}{2}h\tilde{\omega}) \cos(\frac{1}{2}h\tilde{\omega})}{2 \tan(\frac{1}{2}h\tilde{\omega})} = \cos^2(\frac{1}{2}h\tilde{\omega}).$$

Therefore, $\alpha = \beta = \gamma = 1$ holds if and only if $\tan(\frac{1}{2}h\tilde{\omega}) = \frac{1}{2}h\omega$, $\psi(\xi) = \cos^2(\frac{1}{2}\xi)$, and $\phi = 1$, which is precisely the IMEX method. $\square$

Achieving consistency thus requires solving three equations (for α , β , and γ) in three unknowns (ψ , ϕ , and $\tilde{\omega}$). This is impossible for standard trigonometric integrators, since fixing $\tilde{\omega} = \omega$ results in an overdetermined system. However, allowing $\tilde{\omega}$ to be modified introduces the missing degree of freedom necessary to satisfy all three consistency conditions.

2.4. Long-time nearconservation of total energy and modified oscillatory energy. Away from resonances, standard trigonometric integrators nearly conserve the total energy $H(q, p)$ and stiff oscillatory energy $I(q, p)$, up to order $\mathcal{O}(h)$ (Hairer, Lubich, and Wanner [10, Chapter XIII, Theorem 7.1]). In fact, they note that this result can be refined further: under the same assumptions, trigonometric integrators nearly conserve the related quantities

$$H(q, p) - \rho q_1^T g_1(q), \quad J(q, p) - \rho q_1^T g_1(q),$$

each up to order $\mathcal{O}(h^2)$, where

$$\rho = \frac{\psi(h\omega)}{\operatorname{sinc}^2(\frac{1}{2}h\omega)} - 1,$$

and where $J(q, p) = I(q, p) - q_1^T g_1(q)$ is called the *modified oscillatory energy*. In particular, we have $\rho = 0$ for Gautschi-type methods with $\psi(\xi) = \operatorname{sinc}^2(\frac{1}{2}\xi)$ (e.g., Methods A and D in Table 1), so it follows that these methods exhibit even better long-time energy behavior, with $H(q, p)$ and $J(q, p)$ nearly conserved up to order $\mathcal{O}(h^2)$. (See Hairer, Lubich, and Wanner [10, Chapter XIII, Exercise 8].)

We now show that this improved long-time energy behavior also holds for the IMEX method. Observe that, since IMEX corresponds to a trigonometric integrator with frequency $\tilde{\omega}$ in the modified coordinates $(q, \tilde{p})$, it follows that the total and modified oscillatory energies,

$$\tilde{H}(q, \tilde{p}) - \tilde{\rho} \tilde{q}_1^T g_1(q), \quad \tilde{J}(q, \tilde{p}) - \tilde{\rho} \tilde{q}_1^T g_1(q),$$

are nearly conserved up to order $\mathcal{O}(h^2)$. Here, $\tilde{H}$, $\tilde{J}$, and $\tilde{\rho}$ are defined just as above, with $\tilde{\omega}$ in place of ω . For the IMEX method, note also that

$$\frac{\psi(h\tilde{\omega})}{\operatorname{sinc}^2(\frac{1}{2}h\tilde{\omega})} = \frac{\cos^2(\frac{1}{2}h\tilde{\omega})}{\operatorname{sinc}^2(\frac{1}{2}h\tilde{\omega})} = \frac{(\frac{1}{2}h\tilde{\omega})^2}{\tan^2(\frac{1}{2}h\tilde{\omega})} = \frac{(\frac{1}{2}h\tilde{\omega})^2}{(\frac{1}{2}h\omega)^2} = \frac{\tilde{\omega}^2}{\omega^2},$$

so $\tilde{\rho} = \tilde{\omega}^2/\omega^2 - 1$. The following theorem expresses the true energies $H(q, p)$ and $J(q, p)$ in terms of their modified counterparts $\tilde{H}(q, \tilde{p})$ and $\tilde{J}(q, \tilde{p})$, thereby yielding near conservation of both quantities up to $\mathcal{O}(h^2)$.

Note that the IMEX method avoids the undesirable phenomenon of resonance instability, since $h\tilde{\omega}$ is bounded away from nonzero integer multiples of π whenever $h\omega$ is bounded. (An alternative proof for the stability of this method is given in Stern and Grinspun [15].) Therefore, this convergence result can be stated without placing nonresonance restrictions on $h\omega$.

THEOREM 2.2. *For the IMEX method,*

$$\begin{aligned} H(q, p) &= [\tilde{H}(q, \tilde{p}) - \tilde{\rho} q_1^T g_1(q)] - \tilde{\rho} J(q, p), \\ J(q, p) &= \frac{\omega^2}{\tilde{\omega}^2} [\tilde{J}(q, \tilde{p}) - \tilde{\rho} q_1^T g_1(q)]. \end{aligned}$$

Consequently, both H and J are nearly conserved up to $\mathcal{O}(h^2)$ as $h \rightarrow 0$ for any fixed $h\omega$.

Proof. The modified Hamiltonian $\tilde{H}(q, \tilde{p})$ only differs from $H(q, p)$ in replacing $I(q, p)$ by $\tilde{I}(q, \tilde{p}) = \tilde{\omega}^2/\omega^2 I(q, p) = (\tilde{\rho} + 1)I(q, p)$. Therefore,

$$\begin{aligned} H(q, p) &= \tilde{H}(q, \tilde{p}) - \tilde{I}(q, \tilde{p}) + I(q, p) \\ &= \tilde{H}(q, \tilde{p}) - \tilde{\rho} I(q, p) \\ &= \tilde{H}(q, \tilde{p}) - \tilde{\rho} [J(q, p) + q_1^T g_1(q)] \\ &= [\tilde{H}(q, \tilde{p}) - \tilde{\rho} q_1^T g_1(q)] - \tilde{\rho} J(q, p), \end{aligned}$$

which proves the first equality. By a similar calculation,

$$\begin{aligned} J(q, p) &= \tilde{J}(q, \tilde{p}) - \tilde{I}(q, \tilde{p}) + I(q, p) \\ &= [\tilde{J}(q, \tilde{p}) - \tilde{\rho} q_1^T g_1(q)] - \tilde{\rho} J(q, p), \end{aligned}$$

which rearranges to

$$J(q, p) = (\tilde{\rho} + 1)^{-1} [\tilde{J}(q, \tilde{p}) - \tilde{\rho} q_1^T g_1(q)] = \frac{\omega^2}{\tilde{\omega}^2} [\tilde{J}(q, \tilde{p}) - \tilde{\rho} q_1^T g_1(q)],$$

yielding the second equality. Since we have already seen that $\tilde{J}(q, \tilde{p}) - \tilde{\rho} q_1^T g_1(q)$ is nearly conserved up to $\mathcal{O}(h^2)$, it follows that the same is true of $J(q, p)$. Finally, since $\tilde{H}(q, \tilde{p}) - \tilde{\rho} q_1^T g_1(q)$ and $J(q, p)$ are nearly conserved up to $\mathcal{O}(h^2)$, the first equality implies that so is $H(q, p)$. $\square$

3. Numerical experiments.

3.1. The Fermi–Pasta–Ulam problem. Due to its rich multiscale coupling behavior, a variant of the Fermi–Pasta–Ulam (FPU) problem has become a popular highly oscillatory test problem for numerical integrators. The original FPU problem is due to Fermi, Pasta, and Ulam [4], while the version considered here is due to Galgani et al. [5], and appears extensively in Hairer, Lubich, and Wanner [10, I.5 and XIII].

Suppose we have 2ℓ unit point masses, connected together in series by alternating weak cubic and stiff linear springs. Denote the displacements of the point masses by $q_1, \dots, q_{2\ell} \in \mathbb{R}$, where the endpoints $q_0 = q_{2\ell+1} = 0$ are fixed, and let $p_i = \dot{q}_i$ for

$i = 1, \dots, 2n$. In these variables, the FPU system has the Hamiltonian

$$H(q, p) = \frac{1}{2} \sum_{i=1}^{\ell} (p_{2i-1}^2 + p_{2i}^2) + \frac{\omega^2}{4} \sum_{i=1}^{\ell} (q_{2i} - q_{2i-1})^2 + \sum_{i=0}^{\ell} (q_{2i+1} - q_{2i})^4.$$

To put this into the standard form of a highly oscillatory problem, we follow Hairer, Lubich, and Wanner [10, p. 22] in defining the coordinate transformation

$$\begin{aligned} x_{0,i} &= \frac{q_{2i} + q_{2i-1}}{\sqrt{2}}, & x_{1,i} &= \frac{q_{2i} - q_{2i-1}}{\sqrt{2}}, \\ y_{0,i} &= \frac{p_{2i} + p_{2i-1}}{\sqrt{2}}, & y_{1,i} &= \frac{p_{2i} - p_{2i-1}}{\sqrt{2}}, \end{aligned}$$

so that the Hamiltonian becomes

$$\begin{aligned} H(x, y) &= \frac{1}{2} \sum_{i=1}^{\ell} (y_{0,i}^2 + y_{1,i}^2) + \frac{\omega^2}{2} \sum_{i=1}^{\ell} x_{1,i}^2 \\ &\quad + \frac{1}{4} \left[(x_{0,1} - x_{1,1})^4 + \sum_{i=1}^{\ell-1} (x_{0,i+1} - x_{1,i+1} - x_{0,i} + x_{1,i})^4 + (x_{0,\ell} + x_{1,\ell})^4 \right], \end{aligned}$$

which has the desired form.

Following the numerical examples in Hairer, Lubich, and Wanner [10], we consider an instance of the FPU problem with $\ell = 3$, and where the initial conditions are given by

$$x_{0,1}(0) = 1, \quad y_{0,1}(0) = 1, \quad x_{1,1}(0) = \omega^{-1}, \quad y_{1,1}(0) = 1,$$

with all other initial values set to zero. In terms of the stiff energies $I_j = \frac{1}{2}(y_{1,j}^2 + \omega^2 x_{1,j}^2)$, where $j = 1, 2, 3$, these conditions initialize the FPU system with $I_1 = 1$ and $I_2 = I_3 = 0$. As the system evolves dynamically, the phenomenon of slow energy exchange causes this energy to be transferred among I_1 , I_2 , and I_3 , on the time scale $\mathcal{O}(\omega^{-1})$, while the total stiff energy $I = I_1 + I_2 + I_3$ remains nearly constant.

3.2. Resonance stability and oscillatory energy deviation. Figure 1 depicts the maximum deviation in frequency-scaled oscillatory energy ωI , over the time interval $[0, 1000]$, for a range of different frequencies. As discussed in section 1.4, deviations in I are $\mathcal{O}(\omega^{-1})$, whereas those in ωI are $\mathcal{O}(1)$, making the latter more appropriate for comparison across frequencies. (To our knowledge, the use of ωI rather than I for numerical experiments originated in O’Neale and McLachlan [14].) The time step size is fixed at $h = 0.02$, while $h\omega/\pi$ ranges over $(0, 4.5]$.

The “spikes” seen at nonzero integer values of $h\omega/\pi$ correspond to resonance instability. Note that the energy blowup is particularly severe for Method B (the Deuffhard/impulse method), while Methods C and E have resonances only at even values of $h\omega/\pi$. Only Method G and the IMEX method display no resonance spikes at all.

Away from the resonance instabilities, the energy deviation behavior is also interesting. For all of the methods considered, ωI appears to be $\mathcal{O}(1)$ away from the resonance spikes. Indeed, as long as the methods are stable (i.e., the solutions remain bounded), we have $I = \mathcal{O}(\omega^{-1})$, so all of the methods nearly conserve the adiabatic invariant I . (This holds true whether or not the method is consistent for the individual oscillatory energies I_j .)

However, the methods behave quite differently with respect to the magnitude of the deviations in oscillatory energy. For the reference solution, the maximum deviation

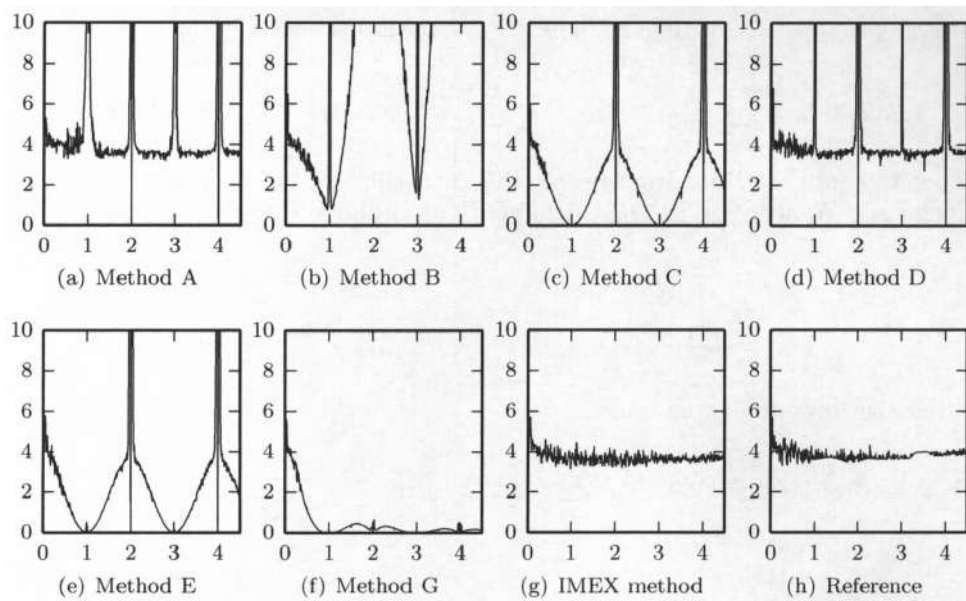


FIG. 1. Maximum deviation of scaled oscillatory energy ωI on the time interval $[0, 1000]$ versus $h\omega/\pi$ ($h = 0.02$).

in ωI is nearly constant with respect to ω , with an approximate numerical value of 4. Methods B, C, and E display a significant decrease in oscillatory energy deviation near odd integer values of $h\omega/\pi$, indicating that the adiabatic invariant is conserved *too well*, compared to the reference solution. (This artificial “antiresonance” can be seen as a sort of numerical damping.) This behavior is even more dramatic for Method G, where the energy deviation is artificially low for nearly all values of $h\omega/\pi$, not just values close to odd integers. Of the methods considered, the IMEX method is the only one which correctly captures the magnitude of these deviations in oscillatory energy. This phenomenon will be revisited and analyzed in section 4, where we will show that this behavior is governed by higher-order terms in the modulated Fourier expansion.

3.3. Long-time near conservation of total energy. Figure 2 shows the maximum deviation in total energy (i.e., in the Hamiltonian) over the time interval $[0, 1000]$, for a range of different frequencies. (This is in contrast to Figure 1, which depicted only the oscillatory energy component of the Hamiltonian.) The reference plot is omitted, as the exact solution preserves total energy exactly. As in Figure 1, Methods A–D again exhibit “spikes” in the energy error at resonant frequencies. Notably, this is not the case for Method E—despite the fact that it exhibited resonance spikes for the oscillatory energy alone—nor for Method G nor IMEX. Furthermore, observe that Methods A, D, and IMEX conserve energy much more closely than the other methods (at least away from resonances), by roughly an order of magnitude. This is consistent with the discussion in section 2.4, including the result in Theorem 2.2, which stated that these methods conserve total energy up to $\mathcal{O}(h^2)$, whereas the remaining methods only do so up to $\mathcal{O}(h)$.

Figure 3 illustrates the relationship of total energy conservation to step size, plotting the log ratio of the deviation in total energy for $h = 0.02$ and $h = 0.04$. As anticipated by the theoretical results in section 2.4, including Theorem 2.2, we see that the

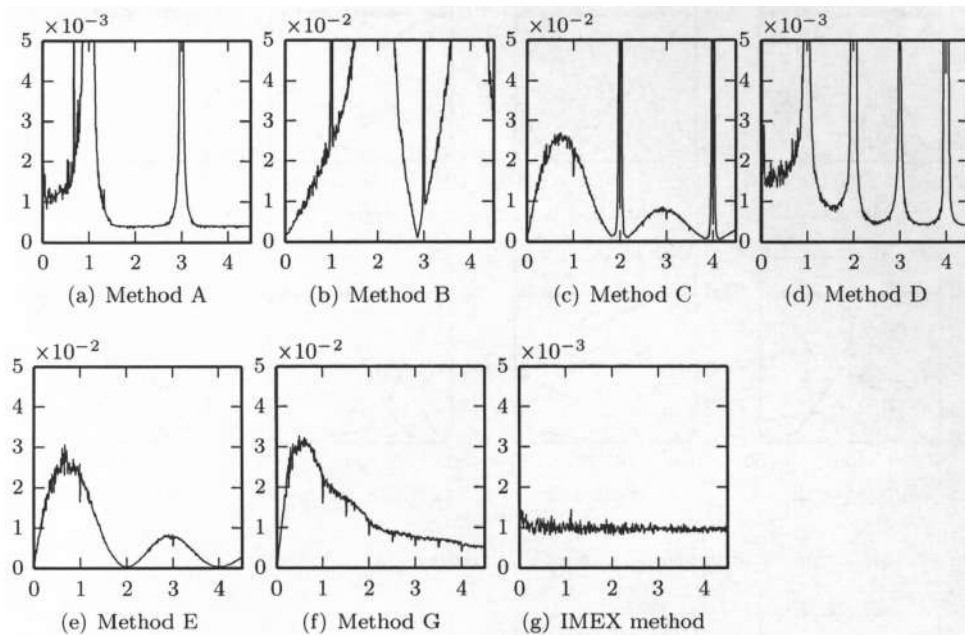


FIG. 2. Maximum deviation of total energy on the time interval $[0, 1000]$ versus $h\omega/\pi$ ($h = 0.02$). The y-axis is scaled 10 times smaller for Methods A, D, and IMEX due to the smaller energy deviations for these methods.

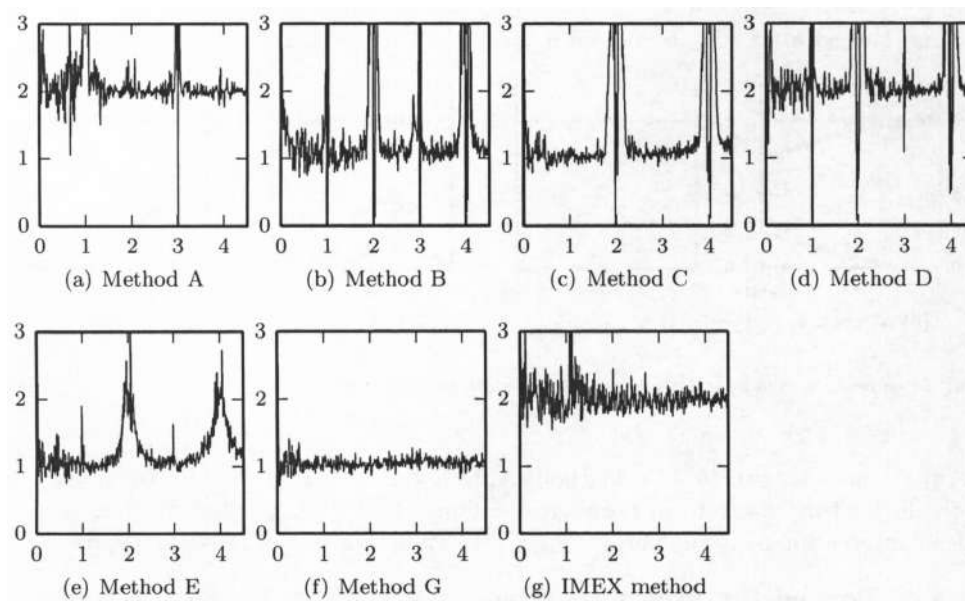


FIG. 3. Log ratio of the maximum deviations in total energy on the time interval $[0, 1000]$, for $h = 0.02$ and $h = 0.04$, plotted against $h\omega/\pi$. Away from resonances, Methods A, D, and IMEX conserve total energy up to $\mathcal{O}(h^2)$, while the remaining trigonometric integrators only conserve total energy up to $\mathcal{O}(h)$.

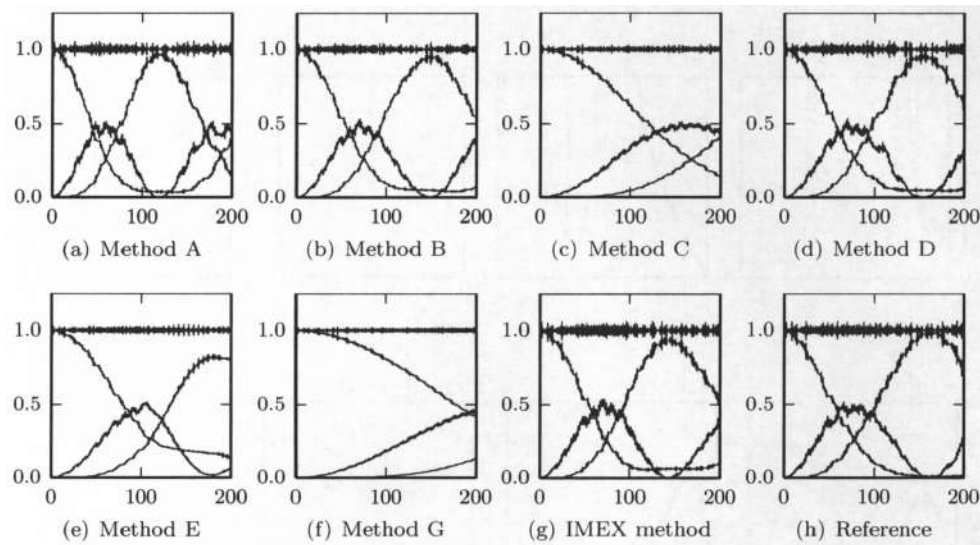


FIG. 4. Slow exchange of individual and total oscillatory energies versus time ($\omega = 50$, $h = 0.03$).

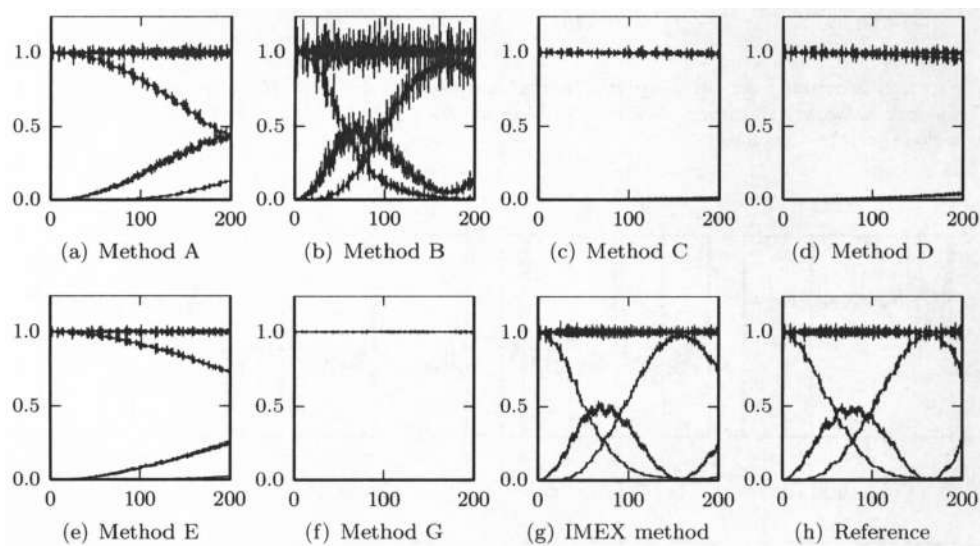


FIG. 5. Slow exchange of individual and total oscillatory energies versus time ($\omega = 50$, $h = 0.1$).

energy deviations are $\mathcal{O}(h^2)$ for Methods A, D, and IMEX, and $\mathcal{O}(h)$ for the remaining methods (at least away from resonances). Only the IMEX method exhibits second-order conservation of total energy, while also remaining free of resonance spikes.

3.4. Slow energy exchange. Figures 4, 5, and 6 depict the phenomenon of slow energy exchange for the FPU problem, following similar numerical experiments in Hairer, Lubich, and Wanner [10]. Each plot contains four curves, corresponding to the three stiff energies, I_1, I_2, I_3 , along with the adiabatic invariant $I = I_1 + I_2 + I_3 \approx 1$.

In Figure 4, the parameters $\omega = 50$ and $h = 0.03$ correspond to a moderate choice of time step size: since $h\omega/\pi \approx 0.48$, this is prior to the onset of resonance instability

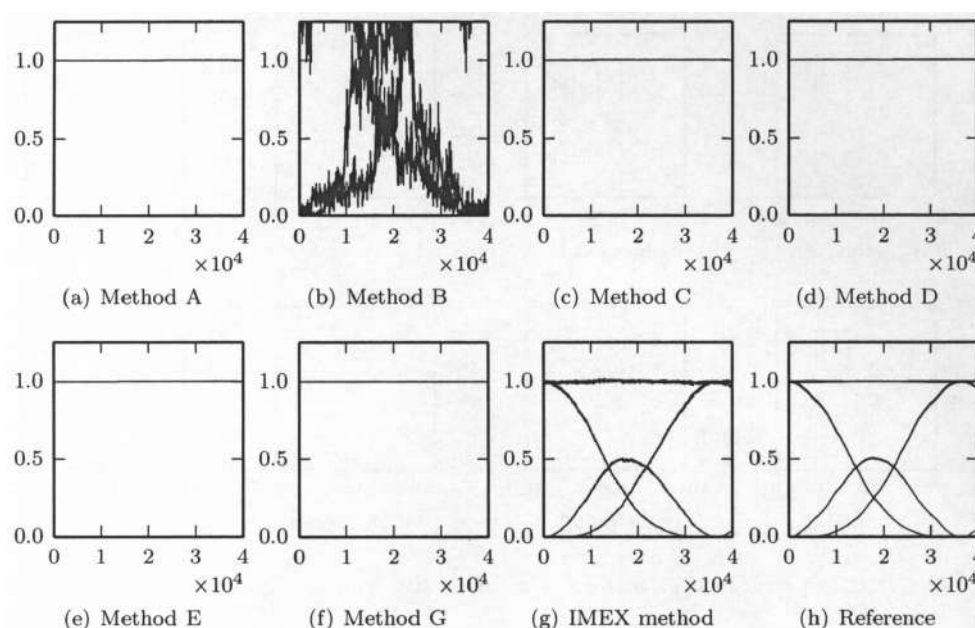


FIG. 6. *Slow exchange of individual and total oscillatory energies versus time ($\omega = 10000$, $h = 0.1$).*

at nonzero integer values. Methods B, D, and the IMEX method give qualitatively correct energy exchange behavior on the time interval $[0, 200]$. By contrast, the exchange occurs too quickly for Method A, and too slowly for Methods C, E, and G (the latter quite dramatically).

In Figure 5, the fast frequency remains $\omega = 50$, but we take a significantly larger time step size of $h = 0.1$ (with $h\omega/\pi \approx 1.59$). Method B and the IMEX method still capture the correct rate of energy exchange on the time interval $[0, 200]$, while for the other methods, the exchange occurs much too slowly. Notice that we are also beginning to see the effects of oscillatory energy deviation, as in Figure 1. Indeed, the excessive “noise” visible for Method B is due to the wide resonance band at $h\omega/\pi = 2$, while the pronounced lack of noise in Method G is due to its artificially low deviations in oscillatory energy. Only the IMEX method displays the correct oscillatory energy behavior, capturing both the rate of exchange and the magnitude of deviations.

Next, in Figure 6, we depict the behavior of these methods as they approach their high-frequency limit, keeping $h = 0.1$ but taking $\omega = 10000$ (hence $h\omega/\pi \approx 318$). Since ω has been scaled by a factor of 200 compared to the previous experiments, the time interval must also be scaled correspondingly, so we look at energy exchange over the interval $[0, 40000]$. As before, only Method B and the IMEX method capture the correct rate of exchange, while for the other methods, the exchange occurs so slowly that it cannot be seen at all on the time scale considered. Method B is again hampered by resonance instability, as in Figure 5, which manifests as excess noise in the oscillatory energy plot. Of the methods considered, only the IMEX method captures the correct energy behavior in the high-frequency limit.

3.5. Global error in slow components. Finally, in Figures 7 and 8, we investigate the global error in the slow position x_0 and slow momentum y_0 , plotted against the time step size h . (These are the only components of real concern; indeed, trigono-

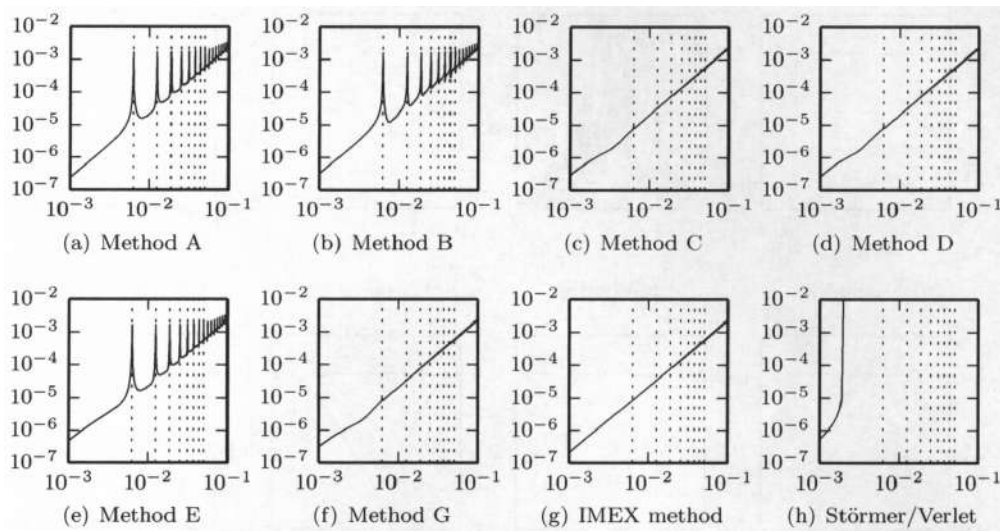


FIG. 7. Global error in the slow position x_0 at the first time step after $t = 1$ versus time step size ($\omega = 1000$).

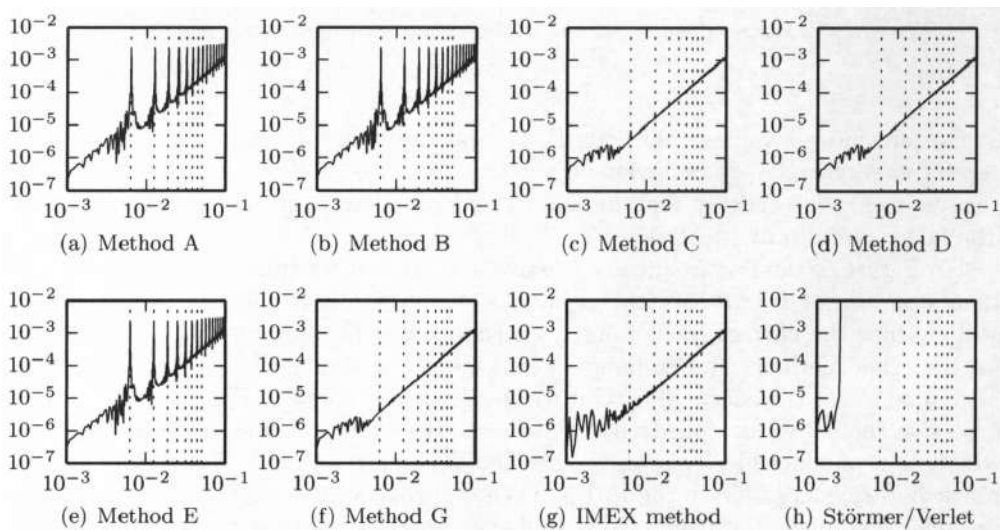


FIG. 8. Global error in the slow momentum y_0 at the first time step after $t = 1$ versus time step size ($\omega = 1000$).

metric methods are designed precisely for problems where we are not interested in resolving the fast oscillations.) For both figures, the error plotted is the Euclidean distance between the numerical solution and a reference solution, taken at the first time step after $t = 1$. Dotted vertical lines indicate values for which $h\omega$ is an integer multiple of 2π , where some of the methods suffer from resonance instabilities that manifest as spikes in the global error. Note that the Störmer/Verlet blows up due to linear instability near $h = 2/\omega = 2 \times 10^{-3}$, illustrating its unsuitability for this type of highly oscillatory problem.

Using the modulated Fourier expansion, the analysis of Hairer, Lubich, and Wanner [10, Chapter XIII, Theorem 4.1] shows that each of the trigonometric methods is

second order, as long as $h\omega$ is bounded away from an integer multiple of π . However, near these resonance points, the order of accuracy reduces to one unless the filter functions satisfy certain conditions. (See also Grimm and Hochbruck [8].) Among the standard trigonometric integrators, Methods C, D, and G satisfy these conditions, and hence are second order uniformly in $h\omega$. By contrast, Methods A, B, and E do not satisfy these conditions, and the resulting error spikes (visible in Figures 7 and 8) lead to a reduction in their uniform order of accuracy.

By construction, however, the IMEX method always has $h\tilde{\omega}$ bounded away from nonzero integer multiples of π . Indeed, since $h\tilde{\omega} = 2 \arctan(\frac{1}{2}h\omega)$, we have $-\pi < h\tilde{\omega} < \pi$, and $h\tilde{\omega}$ approaches $\pm\pi$ only in the limit as $h\omega$ approaches $\pm\infty$. It follows that $h\tilde{\omega}$ is bounded away from nonzero integer multiples of π whenever $h\omega$ is bounded. Hence, comparing the modulated Fourier coefficients, the argument of Hairer, Lubich, and Wanner [10] implies that the global error in the slow components for the IMEX method is second order, uniformly in $h\omega$ on any bounded region $|h\omega| \leq M$.

Combined with the previous results, we remark that only Method G and the IMEX method are uniformly second order and free of resonance instabilities. Of these two methods, however, only IMEX captures the correct stiff energy behavior.

4. Analysis of oscillatory energy deviations. We observed in section 3.2 that the oscillatory energy $I(x, \dot{x})$ is nearly conserved, for long times, by all the methods considered. This is proved in Hairer, Lubich, and Wanner [10, Chapter XIII, Theorem 7.1], where it is shown that the deviations in oscillatory energy are $\mathcal{O}(h)$. However, from Figure 1, it is also apparent that the magnitude of these deviations, for certain methods, is very different from the correct value displayed in the reference solution. This can be explained by carrying the modulated Fourier expansion to one more term.

For the exact solution, the system of modulated Fourier coefficients is known to have a formal invariant $\mathcal{I} = -i\omega(\bar{u}^T \dot{u} - u^T \dot{\bar{u}}) + \mathcal{O}(\omega^{-6})$, where $u = e^{i\omega t} z$. A higher-order expansion of this invariant appears in [10, Chapter XIII, equation 6.12], and the $\mathcal{O}(\omega^{-6})$ estimate is obtained by observing from [10, Chapter XIII, equation 5.3] that the remainder is a product of $-i\omega$ with two additional factors, which are, respectively, $\mathcal{O}(\omega^{-4})$ and $\mathcal{O}(\omega^{-3})$. Now, with $\dot{u} = e^{i\omega t}(\dot{z} + i\omega z) = i\omega e^{i\omega t} z + \mathcal{O}(\omega^{-2})$, this gives

$$\mathcal{I} = 2\omega^2 \|z_1\|^2 + \mathcal{O}(\omega^{-2}).$$

This should be compared to the oscillatory energy, which is

$$I(x, \dot{x}) = \frac{1}{2} \|\dot{x}_1\|^2 + \frac{1}{2} \omega^2 \|x_1\|^2.$$

To do this, consider the expansions

$$\begin{aligned} x_1 &= y_1 + e^{i\omega t} z_1 + e^{-i\omega t} \bar{z}_1 + \mathcal{O}(\omega^{-4}), \\ \dot{x}_1 &= \dot{y}_1 + e^{i\omega t}(\dot{z}_1 + i\omega z_1) + e^{-i\omega t}(\dot{\bar{z}}_1 - i\omega \bar{z}_1) + \mathcal{O}(\omega^{-4}). \end{aligned}$$

Inserting $y_1 = \mathcal{O}(\omega^{-2})$, $z_1 = \mathcal{O}(\omega^{-1})$, $\dot{z}_1 = \mathcal{O}(\omega^{-2})$, we get the estimate

$$\begin{aligned} I(x, \dot{x}) &= \frac{1}{2} \omega^2 \|z_1 - \dot{z}_1\|^2 + \frac{1}{2} \omega^2 \|y_1 + e^{i\omega t} z_1 + e^{-i\omega t} \bar{z}_1\|^2 + \mathcal{O}(\omega^{-2}) \\ &= 2\omega^2 (\|z_1\|^2 + y_1^T \operatorname{Re}(e^{i\omega t} z_1)) + \mathcal{O}(\omega^{-4}) \\ &= \mathcal{I} + 2\omega^2 y_1^T \operatorname{Re}(e^{i\omega t} z_1) + \mathcal{O}(\omega^{-2}). \end{aligned}$$

Since $\mathcal{I}$ is (nearly) conserved over long times, the second term controls the $\mathcal{O}(\omega^{-1})$ deviations in the oscillatory energy I . This term contains two fluctuating components, corresponding to the evolution of $y_1 = \omega^{-2}g_1(y)$ and $2i\omega\dot{z}_1 = g'_1(y)z = g_{1,1}(y_0, 0)z_1 + \mathcal{O}(\omega^{-2})$. Since the latter is controlled by the formal invariant $\mathcal{I}$, it follows that

$$|I - \mathcal{I}| \leq \frac{\sqrt{\mathcal{I}}}{\sqrt{2}}\omega^{-1}\|g_{1,1}(y_0, 0)\| + \mathcal{O}(\omega^{-2}).$$

Repeating the above estimates for modified trigonometric integrators gives the formal invariant

$$\mathcal{I}_h = 2\omega^2\|z_{h,1}\| + \mathcal{O}(\omega^{-2})$$

for $h \rightarrow 0$, $h\omega$ fixed, which is related to I by

$$I(x_n, \dot{x}_n) = \mathcal{I}_h + 2\omega^2 y_{h,1}^T \operatorname{Re}(e^{i\omega t} z_{h,1}) + \mathcal{O}(\omega^{-2}).$$

From the modulated Fourier expansion, we have

$$y_{h,1} = \gamma\phi(h\omega)^{-1}y_1 + \mathcal{O}(\omega^{-3}), \quad \dot{z}_{h,1} = \alpha\dot{z}_1 + \mathcal{O}(\omega^{-2}).$$

Thus, although the $z_{h,1}$ factor may have an incorrect evolution on the $\mathcal{O}(\omega^{-1})$ time scale if $\alpha \neq 1$, in the neighborhood of a particular solution (y, z) on the $\mathcal{O}(1)$ time scale, it does not affect the deviations in I . Rather, these are controlled by the first factor, and are therefore correct to leading order if and only if $\gamma = \phi$, i.e., $\omega^2\psi(h\tilde{\omega}) = \tilde{\omega}^2 \operatorname{sinc}^2(\frac{1}{2}h\tilde{\omega})$.

For standard trigonometric integrators, this is true for the Gautschi-type methods with $\psi(\xi) = \operatorname{sinc}^2(\frac{1}{2}h\omega)$, i.e., for Method A (with $\phi = 1$) and Method D (with $\phi \neq 1$). The IMEX method also satisfies this consistency condition, since $\gamma = \phi = 1$. For the remaining methods, the observed deviations in Figure 1 are correct up to the factor γ/ϕ calculated above. For example, Methods C and E have $\psi(\xi) = \operatorname{sinc}^2(\xi)$, so

$$\frac{\gamma(\xi)}{\phi(\xi)} = \frac{\operatorname{sinc}^2(\xi)}{\operatorname{sinc}^2(\frac{1}{2}\xi)} = \frac{\operatorname{sinc}^2(\frac{1}{2}\xi)\cos^2(\frac{1}{2}\xi)}{\operatorname{sinc}^2(\frac{1}{2}\xi)} = \cos^2(\frac{1}{2}\xi),$$

which is clearly visible, in Figure 1, as period 2π oscillations in the magnitude of energy deviation. On the other hand, for Method G, we have the filter $\psi(\xi) = \operatorname{sinc}^3(\xi)$, so

$$\frac{\gamma(\xi)}{\phi(\xi)} = \frac{\operatorname{sinc}^3(\xi)}{\operatorname{sinc}^2(\frac{1}{2}\xi)} = \frac{\operatorname{sinc}^3(\frac{1}{2}\xi)\cos^3(\frac{1}{2}\xi)}{\operatorname{sinc}^2(\frac{1}{2}\xi)} = \operatorname{sinc}(\frac{1}{2}\xi)\cos^3(\frac{1}{2}\xi),$$

which leads to rapid decay in the magnitude of energy deviation.

5. Conclusion. This paper was motivated by the fact that, while conventional trigonometric integrators have many desirable properties—especially with respect to stability, accuracy, and energy behavior—there are “no-go theorems” making it impossible for any single integrator to have these good properties simultaneously. Other work, particularly on multiforce methods, showed a way around these obstacles, but at the cost of several nonlinear force evaluations per time step. On the other hand, the observations of Stern and Grinspun [15] regarding the IMEX method suggested that, by modifying the fast frequency, one might find another way around these obstacles, without suffering the greater computational cost required by multiforce methods.

By extending the modulated Fourier expansion techniques of Hairer and Lubich [9], Hairer, Lubich, and Wanner [10], we have shown that, for *modified trigonometric integrators*, it is indeed possible to get around these no-go theorems—and that the IMEX method is in fact the unique modified trigonometric integrator which correctly models the multiscale phenomenon of slow energy exchange. Moreover, the IMEX method maintains desirable properties with respect to resonance stability and preservation of adiabatic invariants, while also being uniformly of second-order accuracy in global error, and does so without any additional computational cost relative to conventional trigonometric integrators. Finally, we have shown that while all of these integrators exhibit near conservation of oscillatory energy, only some of them—in particular, the Gautschi-type trigonometric integrators and the IMEX method—consistently model the magnitude of deviations in this adiabatic invariant.

Acknowledgments. The authors wish to thank the organizers and participants in the 2011 Oberwolfach Workshop on Geometric Numerical Integration—especially Ernst Hairer, Marlis Hochbruck, and Christian Lubich—for their valuable feedback on this work. We also thank the anonymous referees for their helpful comments and suggestions.

REFERENCES

- [1] J. J. BIESIADECKI AND R. D. SKEEL, *Dangers of multiple time step methods*, J. Comput. Phys., 109 (1993), pp. 318–328.
- [2] A. DEBUSSCHE AND E. FAOU, *Modified energy for split-step methods applied to the linear Schrödinger equation*, SIAM J. Numer. Anal., 47 (2009), pp. 3705–3719.
- [3] P. DEUFLHARD, *A study of extrapolation methods based on multistep schemes without parasitic solutions*, Z. Angew. Math. Phys., 30 (1979), pp. 177–189.
- [4] E. FERMI, J. PASTA, AND S. ULAM, *Studies of Nonlinear Problems*, Report LA-1940, Los Alamos National Laboratory, Los Alamos, NM, 1955.
- [5] 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.
- [6] B. GARCÍA-ARCHILLA, J. M. SANZ-SERNA, AND R. D. SKEEL, *Long-time-step methods for oscillatory differential equations*, SIAM J. Sci. Comput., 20 (1999), pp. 930–963.
- [7] W. GAUTSCHI, *Numerical integration of ordinary differential equations based on trigonometric polynomials*, Numer. Math., 3 (1961), pp. 381–397.
- [8] V. GRIMM AND M. HOCHBRUCK, *Error analysis of exponential integrators for oscillatory second-order differential equations*, J. Phys. A, 39 (2006), pp. 5495–5507.
- [9] E. HAIRER AND C. LUBICH, *Long-time energy conservation of numerical methods for oscillatory differential equations*, SIAM J. Numer. Anal., 38 (2000), pp. 414–441.
- [10] E. HAIRER, C. LUBICH, AND G. WANNER, *Geometric Numerical Integration*, 2nd ed., Springer Ser. Comput. Math. 31, Springer-Verlag, Berlin, 2006.
- [11] M. HOCHBRUCK AND C. LUBICH, *A Gautschi-type method for oscillatory second-order differential equations*, Numer. Math., 83 (1999), pp. 403–426.
- [12] B. LEIMKUHLER AND S. REICH, *Simulating Hamiltonian Dynamics*, Cambridge Monogr. Appl. Comput. Math. 14, Cambridge University Press, Cambridge, 2004.
- [13] R. I. MCLACHLAN AND G. R. W. QUISP, *Splitting methods*, Acta Numer., 11 (2002), pp. 341–434.
- [14] D. R. J. O'NEALE AND R. I. MCLACHLAN, *Reconsidering trigonometric integrators*, ANZIAM J., 50 (2009), pp. 320–332.
- [15] A. STERN AND E. GRINSFUD, *Implicit-explicit variational integration of highly oscillatory problems*, Multiscale Model. Simul., 7 (2009), pp. 1779–1794.
- [16] M. ZHANG AND R. D. SKEEL, *Cheap implicit symplectic integrators*, Appl. Numer. Math., 25 (1997), pp. 297–302.