

## INSTABILITIES AND INACCURACIES IN THE INTEGRATION OF HIGHLY OSCILLATORY PROBLEMS\*

M. P. CALVO<sup>†</sup> AND J. M. SANZ-SERNA<sup>†</sup>

**Abstract.** By means of a two-frequency test problem, we analyze the instabilities and inaccuracies that may impair the performance of multiple time steps/split operator integrators in highly oscillatory situations, such as those encountered in molecular dynamics, astrophysics, or partial differential equations describing waves. Considered are the impulse (Verlet-I/r-RESPA) method, the mollified impulse method, and the reversible averaging integrator. The analysis covers errors in positions, momenta, and energy.

**Key words.** oscillatory problems, averaging, mollified impulse method

**AMS subject classifications.** 65L20, 65M12, 70F10, 70H05

**DOI.** 10.1137/080727658

**1. Introduction.** This paper is devoted to an analysis of the dangers of applying some numerical integrators to highly oscillatory problems of the form

$$M \frac{d^2}{dt^2} q = f(q),$$

or, equivalently,

$$(1) \quad \frac{d}{dt} p = f(q), \quad \frac{d}{dt} q = M^{-1} p.$$

Here  $q$  is a vector of positions or coordinates,  $p$  the corresponding vector of momenta,  $f$  the vector of forces, and  $M$  a positive-definite, diagonal matrix of masses. This sort of problem appears frequently in many applications, including molecular dynamics, astrophysics, particle accelerators, partial differential equations that describe wave phenomena, etc. We are interested in cases where different time scales coexist in the solutions of the system (1) and multiple-time-step/split step techniques are used for the numerical integration ([9, Chapter VIII] and [11, Chapter 10]). All methods considered in this paper are meant to be used with step lengths chosen to match the slowly varying components of the solution and not restricted by the presence of components of very high frequency. In a sense, we are then dealing with a situation analogous to that encountered in stiff solvers for problems whose solutions have components that decay at widely different rates.

The methods that we investigate follow the track of the fast oscillations because they operate under the assumption (see section 2) that the fast components of the forces may be evaluated efficiently. Therefore the scope of the methods considered here is essentially different from that of alternative techniques such as the heterogeneous multiscale method [3], [4], [1], [14] that attempt only to integrate the “macroscale” of the problem (see also [12]).

---

\*Received by the editors June 18, 2008; accepted for publication (in revised form) October 1, 2008; published electronically February 27, 2009. This research was supported by project MTM 2007-63257, DGI, MEC, Spain.

<http://www.siam.org/journals/sisc/31-3/72765.html>

<sup>†</sup>Departamento de Matemática Aplicada, Universidad de Valladolid, Valladolid, Spain (maripaz@mac.uva.es, sanzsern@mac.uva.es, <http://hermite.mac.cie.uva.es/sanzserna>).

It has been known for a long time [2] that multiple-time-step approaches like those considered in this paper may lead to instabilities that are due to resonances between the frequencies present in the problem and the rate at which the solution is sampled by the numerical method. Furthermore, these schemes, if not used with care, may yield stable numerical runs that completely misrepresent the true solution (see, e.g., [5]). Here we introduce a two-degrees-of-freedom linear test problem to analyze such instabilities and inaccuracies. Due to the relatively simple structure of the test equations, it is possible to obtain (perhaps with some help from a symbolic manipulation program) closed form expressions for all relevant quantities in the numerical schemes and then to compare them to their counterparts in the true theoretical solution. This is not very different from the approach taken in section 3 of [8].<sup>1</sup>

The paper is divided into seven sections with some mathematical proofs relegated to two appendices. Section 2 contains a description of the three methods to be analyzed: the impulse method [7], [15], the mollified impulse method introduced in [5], and the reversible averaging integrator from [10]. The test problem is presented in section 3, and then sections 4, 5, and 6 contain the results of our analysis. The conclusions of our study are listed in section 7.

Although all our results are mathematically rigorous they are not always presented in full detail; that would have resulted in too long of a paper.

**2. Numerical methods.** The three numerical techniques considered in this paper were devised to be applied in possibly nonlinear, realistic problems. We therefore describe them in a nonlinear setting, but note that the corresponding formulations become simpler when the problem at hand is linear. In fact, for linear problems the methods can alternatively be viewed as exponential integrators; see [9, Chapter XIII].

**2.1. The impulse method (IM).** We begin our study by describing the (Verlet-I/r-RESPA) *impulse method* [7], [15], a technique for the numerical integration of (1) in cases where the forces  $f$  can be split into a strong part  $f_s$  and a weak part  $f_w$

$$(2) \quad \frac{d}{dt}p = f_s(q) + f_w(q), \quad \frac{d}{dt}q = M^{-1}p,$$

in such a manner that  $f_w$  does not contribute any fast modes to the solution and the *reduced* problem

$$(3) \quad \frac{d}{dt}p = f_s(q), \quad \frac{d}{dt}q = M^{-1}p$$

possesses highly oscillatory modes (and perhaps slow modes as well). The idea of taking advantage of the strong/weak splitting is appealing in situations where integrating the reduced system is much cheaper than solving numerically the full system. An example is provided by cases where the reduced problem may be solved analytically at a cost that is independent (or almost independent) of the step length. In other cases the cost of evaluating the weak forces is considerably higher than that of evaluating the strong forces; those cases include differential equations from molecular dynamics or astrophysics where the strong forces involve only next-neighbor interactions and the weak forces comprise interactions between all particles. In these situations the aim

---

<sup>1</sup>In the model problem used in [8] the matrix that originates the fast oscillations has been brought to diagonal form through a change of variables. Here we work with the original dependent variables, which are, of course, easier to understand in physical terms. Also note that the RAI is *not* invariant with respect to changes of variables.

is to sample the weak forces as sparingly as possible and, certainly, at a rate that is not determined by the periods of the fast motions, i.e., by the stiffness of the reduced problem.

The IM integrates (2) by interspersing evaluations of the weak forces at intervals of length  $h$  with integrations of the reduced system (3) that may be assumed to be exact either because they are performed analytically or because, in a multiple-time-step approach, they are carried out with a time step much shorter than  $h$ . If we denote by  $p^n$  and  $q^n$  the numerical approximations to the true solution values of the momenta  $p(t^n)$  and coordinates  $q(t^n)$  at the step point  $t^n = nh$ , then a step  $n \rightarrow n + 1$  of the IM may be described in the following composition or split-step pattern.

*Kick:* Evaluate the weak force  $\bar{f}_w^n = f_w(q^n)$  and then set

$$(4) \quad p^{n+} = p^n + \frac{h}{2} \bar{f}_w^n.$$

*Oscillation:* Use the  $h$ -flow of the reduced problem (3) to advance from  $(p^{n+}, q^n)$  to  $(p^{n+1-}, q^{n+1})$ . In other words,  $(p^{n+1-}, q^{n+1})$  is obtained by integrating, over a time interval of length  $h$ , the system (3) with initial conditions  $(p^{n+}, q^n)$ .

*Kick:* Set  $p^{n+1} = p^{n+1-} + \frac{h}{2} \bar{f}_w^{n+1}$  with  $\bar{f}_w^{n+1} = f_w(q^{n+1})$ .

Of course, the force  $\bar{f}_w^{n+1}$  at the second kick of the current step coincides with the force at the first kick of the next step, so that the method essentially consists of a sequence of oscillations  $(p^{n+}, q^n) \rightarrow (p^{n+1-}, q^{n+1})$  followed by kicks  $p^{n+1+} = p^{n+1-} + h\bar{f}_w^{n+1}$ ; it is not necessary to compute the vector  $p^{n+1}$  unless output at  $t^{n+1}$  is required.

Clearly, the IM would be exact if  $f_w \equiv 0$  so that hopes may be entertained that it would integrate the full system (2) with errors that are small *uniformly* in the stiffness of the problem. Unfortunately, it turns out [5], [13] that the expected  $O(h^2)$  errors materialize only if  $h$  is small with respect to the smallest period of the reduced system. Hence the IM suffers from an order reduction in the presence of stiffness.

**2.2. Mollified impulse method (MIM).** In an attempt to overcome the shortcomings of the IM, the paper [5] suggested a *mollified* version that, under suitable hypotheses, is able to find the coordinates  $q$  with errors that are  $O(h^2)$ , where the implied error constant depends only on the energy of the solution being found (and on bounds for  $f_w$  and its derivatives).

Indeed there is a *family* of mollified methods rather than a single one; the article [5] introduced a method for each choice of a weight function, and later [13] the family grew to include a method for each choice of a pair  $(\psi, \phi)$  of weight functions. Here a weight function is a bounded, integrable real-valued function  $\chi(t)$  that is assumed to be even  $\chi(-t) \equiv \chi(t)$  and to satisfy

$$\int_{-\infty}^{\infty} \chi(s) ds = 1.$$

Note that we do not require  $\chi \geq 0$ .

The MIM differs from its “plain” counterpart only in that the force  $\bar{f}_w^n$  used in (4) to kick at  $t^n$  has the more complicated format

$$(5) \quad \bar{f}_w^n = \mathcal{M}(q^n, h) f_w(\mathcal{A}(q^n, h)).$$

Here, given  $q^n$  and  $h$ ,  $\mathcal{A}(q^n, h)$  represents an average of values of  $q$  and  $\mathcal{M}(q^n, h)$  is a so-called mollifier matrix. The use of  $\mathcal{A}$  avoids [5] the dangers associated with

sampling at grid points a quickly varying  $f_w(q(t))$ , and the role of  $\mathcal{M}$  is to mimic the way in which extra external forces contribute to build up the momentum in highly oscillatory problems. More precisely, for the method defined by the pair of weight functions  $(\psi, \phi)$ , the averaged value is defined by

$$\mathcal{A}(q^n, h) = \frac{1}{h} \int_{-\infty}^{\infty} q^*(t) \phi\left(\frac{t}{h}\right) dt = \int_{-\infty}^{\infty} q^*(hs) \phi(s) ds,$$

where  $q^*(t)$  is obtained by solving the reduced problem (3) with initial conditions  $q = q^n, p = 0$ . The idea behind the mollification process is not quite as simple. The plain IM can be seen as an exact integration of the approximation ( $\delta$  stands for the Dirac function)

$$\frac{d^2}{dt^2}q = f_s(q) + \sum_n \delta\left(\frac{t - t^n}{h}\right) f_w(q^n),$$

and by the mollification process this “impulsive” formulation is replaced by a less abrupt version where the weak force acts in a time-distributed, continuous fashion:

$$(6) \quad \frac{d^2}{dt^2}q = f_s(q) + \sum_n \psi\left(\frac{t - t_n}{h}\right) f_w^{*n}$$

(here  $f_w^{*n} = f_w(\mathcal{A}(q^n, h))$ ,  $n = 0, 1, \dots$ , are the values of the force to be mollified). The exact solution of (6) can be obtained by an application of the nonlinear variation-of-constants formula to the reduced problem, and this involves solving the variational equation for (3). The mollification matrices  $\mathcal{M}(q^n, h)$  required by the algorithm are found [5], [13] by numerically implementing this variation-of-constants approach.

In a nutshell, a MIM differs from its plain counterpart in that, at each step, additional integrations of the reduced problem and its variational equation are performed in order to build up a mollifying matrix  $\mathcal{M}(q^n, h)$  and an average position  $\mathcal{A}(q^n, h)$  to be used in (5). Full details on how to implement the MIM may be obtained from [5], [13].

For a suitably normalized, general class of problems (2) with linear strong forces, it is proved rigorously in [13] (see also [5], [9], [6]) that a MIM produces errors in the variables  $p$  and  $q$  of size  $O(h)$  uniformly with respect to the stiffness of the reduced problem if and only if the mollification weight function satisfies

$$(7) \quad \sum_{j=-\infty}^{\infty} \psi(t - j) \equiv 1.$$

This amounts to demanding that, if  $\psi$  is seen as a basis function to interpolate in time the values  $f_w^{*n}$  in (6), then the interpolation is exact when the function being interpolated is constant. Furthermore, when (7) holds, a MIM yields errors in the coordinates  $q$  that are  $O(h^2)$  uniformly in the stiffness if and only if

$$(8) \quad \sum_{j=-\infty}^{\infty} \phi(t - j) \equiv 1.$$

From Fourier analysis, it is well known that the conditions (7) and (8) are equivalent to the requirements

$$(9) \quad \widehat{\psi}(2\pi n) = 0, \quad n = \pm 1, \pm 2, \dots,$$

and

$$(10) \quad \widehat{\phi}(2\pi n) = 0, \quad n = \pm 1, \pm 2, \dots,$$

for the corresponding Fourier transforms  $\widehat{\psi}$  and  $\widehat{\phi}$ .

A variety of weight functions have been suggested in the literature; see [5], [8], [9]. The simplest (in a sense made precise in [13]) is the so-called *short* weight with  $\chi(t) = 1$  for  $-1/2 < t < 1/2$  and  $\chi(t) = 0$  for  $|t| \geq 1/2$ . The corresponding Fourier transform is the sinc function  $\widehat{\chi}(\omega) = (2/\omega) \sin(\omega/2)$ .

**2.3. A reversible averaging integrator (RAI).** Another extension of the Verlet algorithm to a multiple-time-scale setting was put forward in [10] (see also [11, section 10.3]). The vector of forces in (1) is not decomposed as in (2); rather, it is assumed that, due either to the structure of the potential energy or to the relative size of the different masses involved, the coordinates  $q$  may be partitioned into a set of “slow” variables  $Q$  and “fast” variables  $\theta$ . Then, after denoting by  $P$  and  $\mu$  the momenta corresponding to the  $Q$  and  $\theta$  variables and introducing the obvious partitions of the vector  $f$  and the diagonal mass matrix  $M$ , the system (1) becomes

$$(11) \quad \frac{d}{dt}P = f_Q(Q, \theta), \quad \frac{d}{dt}Q = M_Q^{-1}P,$$

$$(12) \quad \frac{d}{dt}\mu = f_\theta(Q, \theta), \quad \frac{d}{dt}\theta = M_\theta^{-1}\mu.$$

A step of the algorithm may be described in the following kick/oscillate/kick pattern:

*Kick:* Hold the variables  $Q$  frozen at the value  $Q^n$  and integrate the fast system (12) in a time interval of length  $h$  using the initial data  $(\mu^n, \theta^n)$ . Compute an averaged<sup>2</sup> force

$$f_Q^{n+} = h^{-1} \int_{t^n}^{t^{n+1}} f_Q(Q^n, \theta(t)) dt$$

and use it to kick the slow momentum variable:

$$P^{n+1/2} = P^n + \frac{h}{2} f_Q^{n+}.$$

*Oscillate:* Now the variables  $Q, \mu, \theta$  are advanced from  $t^n$  to  $t^{n+1}$ . One holds  $P(t)$  frozen at its most recent value  $P^{n+1/2}$ , which, in view of the second equality in (11), leads to a linearly growing  $Q(t) = Q^n + (t - t^n)M_Q^{-1}P^{n+1/2}$ . This formula for  $Q(t)$  is used to compute  $Q^{n+1}$  and also taken to (12), which can then be integrated over the interval  $t^n \leq t \leq t^{n+1}$  to advance the variables  $\mu$  and  $\theta$  from  $(\mu^n, \theta^n)$  to  $(\mu^{n+1}, \theta^{n+1})$ .

*Kick:* The step is completed by updating  $P$  from  $P^{n+1/2}$  to  $P^{n+1}$  with the help of a new kick. To ensure the reversibility<sup>3</sup> of the overall algorithm one proceeds in a

<sup>2</sup>As suggested in [10], more general averages may be considered, but this possibility has not been explored by us.

<sup>3</sup>The IM and MIM are also reversible. The reversibility of the IM is an almost obvious consequence of the kick/oscillation/kick pattern. The situation for the MIM is more subtle; reversibility is obtained because the required auxiliary integrations of the reduced problem use an “artificial” initial condition  $p = 0$ , and therefore  $\mathcal{A}$  and  $\mathcal{M}$  are not changed by a time reversal. The RAI builds on the partitioning of the variables to achieve reversibility without artificially setting to zero the initial value of the fast momenta in the auxiliary integration used to average the force.

symmetric fashion: i.e., (12) is integrated, with  $Q$  frozen at  $Q^{n+1}$  backwards in time from  $t^{n+1}$  to  $t^n$  using the “final” data  $\mu^{n+1}$ ,  $\theta^{n+1}$ , and the values of  $\theta$  obtained in this way are used to average the slow forces  $f_Q$  before kicking.

The error behavior of this reversible method does not appear to have been investigated in the literature.

**3. A test problem.** If  $\omega > 0$  denotes a parameter and  $\alpha$  is a constant,  $0 < \alpha < 2$ , then we study the equations

$$\begin{aligned} \frac{d}{dt}p_1 &= -q_1 + \omega^\alpha(q_2 - q_1), \\ \frac{d}{dt}q_1 &= p_1, \\ \frac{d}{dt}p_2 &= -\omega^\alpha(q_2 - q_1), \\ \frac{d}{dt}q_2 &= \omega^{2-\alpha}p_2, \end{aligned}$$

i.e., the Hamiltonian system corresponding to the Hamiltonian function

$$(13) \quad H = T + V, \quad T = \frac{1}{2}p_1^2 + \frac{\omega^{2-\alpha}}{2}p_2^2, \quad V = \frac{1}{2}q_1^2 + \frac{\omega^\alpha}{2}(q_2 - q_1)^2.$$

These equations may describe the oscillations of a mechanical system involving two material points; the first, with unit mass, is attached to a rigid wall through a massless spring with unit elastic constant and the second, with mass  $\omega^{-2+\alpha}$ , is linked to the first through a massless spring with elastic constant  $\omega^\alpha$ . The system is constrained to move along a straight line; the coordinates  $q_i$  measure the displacements of the masses from their equilibrium positions and  $p_i$  are the corresponding momenta. The value of  $H$  gives the total mechanical energy and is, of course, a conserved quantity along solutions of the system.

We are primarily interested in the case where  $\omega \gg 1$  so that the spring joining both masses is strong and the second mass is light; as will be shown below, the system then has fast oscillations whose frequency is  $\approx \omega$ . The constant  $\alpha$  determines the relative contributions of the strong spring and the light mass to the existence of fast oscillations. When  $\alpha$  is near 0, the fast oscillations are due to the smallness of the second mass; for  $\alpha$  near 2 the system is rapidly oscillatory due to the stiffness of the spring. The intermediate case  $\alpha = 1$  was used in the original paper [10] and again in [11] as a model problem for the behavior of the RAI.

Before undertaking the study of the numerical integrators presented in section 2, we investigate in some detail the nature of the true solutions of the model problem. These mathematical details play an important role later in *analyzing* the integrators, but are, of course, not needed at all to obtain the numerical solutions themselves.

By looking for normal modes where both masses oscillate with a common frequency, we find the four particular solutions given by (rows correspond to the variables  $p_1, q_1, p_2, q_2$  in this order)

$$(14) \quad \begin{bmatrix} \cos(\Omega_- t) \\ \Omega_-^{-1} \sin(\Omega_- t) \\ \xi_- \cos(\Omega_- t) \\ (1 + \xi_- \omega^{-\alpha} \Omega_-^2) \Omega_-^{-1} \sin(\Omega_- t) \end{bmatrix}, \quad \begin{bmatrix} -\sin(\Omega_- t) \\ \Omega_-^{-1} \cos(\Omega_- t) \\ -\xi_- \sin(\Omega_- t) \\ (1 + \xi_- \omega^{-\alpha} \Omega_-^2) \Omega_-^{-1} \cos(\Omega_- t) \end{bmatrix},$$

TABLE 1

Upper part: Oscillation amplitudes for the variables  $p_i$  and  $q_i$  in the normal modes. Lower part: Energy values in the normal modes.

	Slow modes	Fast modes
$p_1$	$O(1)$	$O(\omega^{-1+\alpha/2})$
$q_1$	$O(1)$	$O(\omega^{-2+\alpha/2})$
$p_2$	$O(\omega^{-2+\alpha})$	$O(\omega^{-1+\alpha/2})$
$q_2$	$O(1)$	$O(\omega^{-\alpha/2})$
$p_1 + p_2$	$O(1)$	$O(\omega^{-3+\alpha/2})$
$q_2 - q_1$	$O(\omega^{-2})$	$O(\omega^{-\alpha/2})$
$H$	$O(1)$	$O(1)$
$H_{weak}$	$O(1)$	$O(\omega^{-2+\alpha})$
$H_{strong}$	$O(\omega^{-2+\alpha})$	$O(1)$

and

$$(15) \quad \omega^{-1+\alpha/2} \begin{bmatrix} (-1 + \xi_+) \cos(\Omega_+ t) \\ \xi_+ \Omega_+ \sin(\Omega_+ t) \\ \cos(\Omega_+ t) \\ (\omega^{-\alpha} + \xi_+) \Omega_+ \sin(\Omega_+ t) \end{bmatrix}, \quad \omega^{-1+\alpha/2} \begin{bmatrix} -(-1 + \xi_+) \sin(\Omega_+ t) \\ \xi_+ \Omega_+ \cos(\Omega_+ t) \\ -\sin(\Omega_+ t) \\ (\omega^{-\alpha} + \xi_+) \Omega_+ \cos(\Omega_+ t) \end{bmatrix},$$

where  $\Omega_-, \Omega_+$  ( $\Omega_- < \Omega_+$ ) are the two nonnegative roots of the characteristic equation

$$\Omega^4 - (\omega^2 + \omega^\alpha + 1)\Omega^2 + \omega^2 = 0$$

and

$$\xi_- = \Omega_-^{-2} - 1, \quad \xi_+ = (1 - \Omega_+^2)^{-1}.$$

The solutions (14) and (15) have been scaled to ensure that for each of them the energy  $H$  in (13) remains bounded as  $\omega \rightarrow \infty$ .

After solving the characteristic equation and expanding in negative powers of the parameter  $\omega$ , we find the asymptotic behavior of the frequencies,

$$(16) \quad \Omega_- = 1 - \frac{1}{2}\omega^{-2+\alpha} + O(\omega^{-4+2\alpha}), \quad \omega \rightarrow \infty,$$

$$(17) \quad \Omega_+ = \omega + \frac{1}{2}\omega^{-1+\alpha} + O(\omega^{-3+2\alpha}), \quad \omega \rightarrow \infty,$$

and of the *small* coefficients  $\xi_\pm$ :

$$(18) \quad \xi_- = +\omega^{-2+\alpha} + O(\omega^{-4+2\alpha}), \quad \omega \rightarrow \infty,$$

$$(19) \quad \xi_+ = -\omega^{-2} + O(\omega^{-4+\alpha}), \quad \omega \rightarrow \infty.$$

In view of (16) and (17), for large  $\omega$  the solutions in (14) represent slow modes with frequency  $\approx 1$  and the solutions in (15), with frequency  $\approx \omega$ , are fast modes. The estimates (18) and (19) may be used to find the oscillation amplitudes of the dynamic variables  $p_i$  and  $q_i$  (see Table 1). In a slow mode, the elongation  $q_2 - q_1$  of the strong spring is negligible and the system almost behaves as if both masses were joined by a rigid, massless rod driven by the soft spring. For the fast modes,  $q_1$  is much smaller than  $q_2$ : the small mass moves driven by the strong spring (which stores an  $O(1)$  potential energy  $V$ ) while the weak spring plays little role (the smallness of

its elongation limits its potential energy to  $O(\omega^{-4+\alpha})$ . Also note that in a fast mode  $p_1$  and  $p_2$  almost cancel each other; the masses move in opposition of phase and the lightness of the second mass, relative to the first, is compensated by the fact that it undergoes much wider oscillations.

The lower part of Table 1 displays the values of the total energy  $H$  in the slow and fast modes, as well as those of the energy

$$(20) \quad H_{weak} = \frac{1}{2}p_1^2 + \frac{1}{2}q_1^2$$

in the subsystem consisting of the first mass and the weak spring and the energy

$$(21) \quad H_{strong} = \frac{\omega^{2-\alpha}}{2}p_2^2 + \frac{\omega^\alpha}{2}(q_2 - q_1)^2$$

in the subsystem consisting of the light mass and the strong spring. We emphasize that  $H_{weak}$  and  $H_{strong}$  are *not* conserved quantities of our system.

The most general solution with  $H = O(1)$  is, of course, obtained by forming linear combinations of the four normal modes with coefficients of size  $O(1)$ , i.e., is of the form

$$(22) \quad \mathbf{u}(t) = \mathcal{P}\mathcal{G}\mathbf{c},$$

where  $\mathbf{u}(t)$  is the solution vector, the entries of  $\mathbf{c}$  are the weights in the linear combination,  $\mathcal{G}$  represents the rotation matrix

$$\mathcal{G} = \mathcal{G}(t; \omega) = \begin{bmatrix} \cos(\Omega_- t) & -\sin(\Omega_- t) & 0 & 0 \\ \sin(\Omega_- t) & \cos(\Omega_- t) & 0 & 0 \\ 0 & 0 & \cos(\Omega_+ t) & -\sin(\Omega_+ t) \\ 0 & 0 & \sin(\Omega_+ t) & \cos(\Omega_+ t) \end{bmatrix},$$

and  $\mathcal{P} = \mathcal{P}(\omega)$  is given by

$$\mathcal{P} = \begin{bmatrix} 1 & 0 & (-1 + \xi_+)\omega^{-1+\alpha/2} & 0 \\ 0 & \Omega_-^{-1} & 0 & \xi_+\Omega_+\omega^{-1+\alpha/2} \\ \xi_- & 0 & \omega^{-1+\alpha/2} & 0 \\ 0 & 1 + \xi_-\omega^{-\alpha}\Omega_- & 0 & (\omega^{-\alpha} + \xi_+)\Omega_+\omega^{-1+\alpha/2} \end{bmatrix}.$$

Finally, the flow of the system is obtained by replacing in (22) the vector of coefficients  $\mathbf{c}$  by its expression in terms of the initial condition:

$$(23) \quad \mathbf{u}(t) = \Phi(t; \omega)\mathbf{u}(0) = \mathcal{P}\mathcal{G}(t; \omega)\mathcal{P}^{-1}\mathbf{u}(0).$$

Thus the eigenvalues of the matrix  $\Phi(t; \omega)$  are  $\lambda = \exp(\pm i\Omega_\pm t)$  and the corresponding eigenspaces are determined by  $\mathcal{P}$ .

The numerical methods described in section 2 replace the true evolution (23) by a time-discrete counterpart

$$(24) \quad \mathbf{u}^n = \tilde{\Phi}(h; \omega)^n \mathbf{u}^0,$$

where the (method-dependent) propagator matrix  $\tilde{\Phi}$  advances the solution over a single time step. When the propagator possesses eigenvalues of unit modulus, it can

be factorized in the form  $\tilde{\Phi}(h; \omega) = \tilde{\mathcal{P}}\tilde{\mathcal{G}}(h; \omega)\tilde{\mathcal{P}}^{-1}$ , where  $\tilde{\mathcal{G}}(h, \omega)$  is a rotation matrix that approximates  $\mathcal{G}(h, \omega)$  and  $\tilde{\mathcal{P}}$  is an approximation to  $\mathcal{P}$ . Therefore, in that case, the numerical process (24) contains, in an implicit way, a discrete version of (23) of the form

$$(25) \quad \mathbf{u}^n = \left(\tilde{\mathcal{P}}\tilde{\mathcal{G}}(h; \omega)\tilde{\mathcal{P}}^{-1}\right)^n \mathbf{u}^0 = \tilde{\mathcal{P}}\tilde{\mathcal{G}}(h; \omega)^n\tilde{\mathcal{P}}^{-1}\mathbf{u}^0 = \tilde{\mathcal{P}}\tilde{\mathcal{G}}(nh; \omega)\tilde{\mathcal{P}}^{-1}\mathbf{u}^0.$$

**4. An analysis of the RAI.** As pointed out before, the instance  $\alpha = 1$  of the test problem in section 3 was considered in the paper [10] (where the RAI was introduced) and, later, also in [11]. As in those publications, we apply the RAI with the obvious choice  $Q = q_1$  and  $\theta = q_2$ .

**4.1. Stability.** The expression of the RAI propagator as a function of  $h$  and  $\omega$  is forbiddingly complex and need not be reproduced here (for  $\alpha = 1$  it may be seen in [10]). The corresponding characteristic equation turns out to be friendlier and is given by

$$(26) \quad \lambda^4 - a\lambda^3 + b\lambda^2 - a\lambda + 1 = 0,$$

where

$$a = 2 + 2c - h^2 - s^2\omega^{\alpha-2}, \quad b = 2 + 4c - 2ch^2 - 2s^2\omega^{\alpha-2},$$

with

$$c = \cos(\omega h), \quad s = \sin(\omega h).$$

The symmetry of the coefficients in (26) ensures that its roots appear in pairs  $\lambda, \lambda^{-1}$ , as befits a reversible integrator.

For a stable integration, where  $\mathbf{u}^n$  remains bounded as  $n$  increases, we require that the eigenvalues of the propagator  $\tilde{\Phi}$  have modulus = 1 and (if they are multiple) possess nondefective eigenvectors. The proof of the following result is given in Appendix A.

PROPOSITION 1. *For the test problem, the condition*

$$(27) \quad h < 2\sqrt{1 - \omega^{\alpha-2}}$$

*guarantees the stability of the RAI. More precisely, if this condition holds, then we have the following:*

- *For  $\omega h \neq k\pi, k = 1, 2, \dots$ , the RAI propagator has four distinct nonreal eigenvalues of unit modulus.*
- *For  $\omega h = (2k - 1)\pi, k = 1, 2, \dots$ , the propagator has a nondefective double eigenvalue  $\lambda = -1$  and a pair of (nonreal) complex-conjugate eigenvalues of unit modulus.*
- *For  $\omega h = 2k\pi, k = 1, 2, \dots$ , the propagator has a nondefective double eigenvalue  $\lambda = 1$  and a pair of (nonreal) complex-conjugate eigenvalues of unit modulus.*

In the limit  $\omega \rightarrow \infty$ , the light mass and the stiff spring disappear from our mechanical system and it is good news that, correspondingly, the condition (27) converges to the stability requirement  $h < 2$  of Verlet method applied to the “slow” harmonic oscillator with Hamiltonian (20). Thus the stability limit of the RAI method is determined only by the slow motion, a favorable property discovered experimentally in [10] for the case  $\alpha = 1$  and not shared by the IM or the MIM.

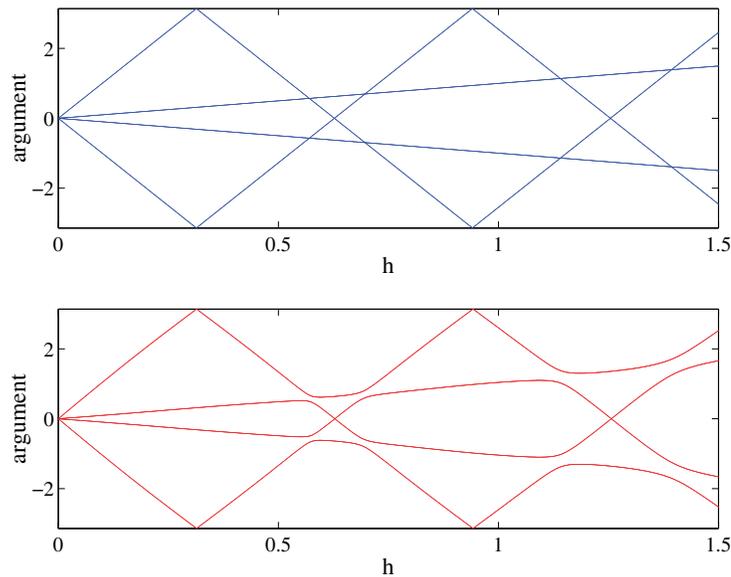


FIG. 1. Behavior of the argument of the eigenvalues of the true flow (top) and RAI propagator (bottom) as functions of the step length  $h$  with  $\omega$  fixed.

In Figure 1, where<sup>4</sup>  $\alpha = 1$ ,  $\omega = 10$ , we have depicted, as functions of  $h$ , the arguments of the four eigenvalues of the true flow  $\Phi(h; \omega)$ , and their counterparts in the RAI propagator  $\tilde{\Phi}(h; \omega)$ . Angles have been reduced to the interval  $[-\pi, \pi]$ . For  $h = 0$  the propagator and the flow are the identity, the four eigenvalues coalesce at  $\lambda = 1$ , and the corresponding arguments coalesce at 0. As  $h$  increases, the pair of fast eigenvalues move rapidly away from 1 and for  $h \approx \pi/\omega$  they cross at  $\lambda = -1$  (argument  $\pm\pi$ ). As  $h$  approaches the neighborhood of  $2\pi/\omega$ , the fast eigenvalues of the true flow return to the starting point  $\lambda = 1$  after having completed a full turn around the unit circle  $|\lambda| = 1$ ; in such a turn each of the two fast eigenvalues has met and crossed a slow eigenvalue. For the RAI propagator (as noted in [10]) such eigenvalue crossings do not take place.

**4.2. Order reduction.** The analysis of numerical methods usually involves the study of their behavior when they are applied with  $h \rightarrow 0$  to a given problem. Here the study of the limit  $h \rightarrow 0$  with  $\omega$  fixed would yield little information on the different methods because they are meant to be used when  $h$  is *not small* relative to the period  $2\pi/\omega$  of the fast oscillation. We rather consider the limit process

$$(28) \quad h \rightarrow 0, \quad \omega \rightarrow \infty, \quad h\omega = \eta > 0,$$

where  $\eta$  is a *constant*. Thus the problem to which the methods are applied changes as  $h$  is decreased. We also assume that the initial conditions are chosen in such a way that, in the limit (28), the energy  $H$  of the solutions remains bounded by a constant independent of  $h$  and  $\omega$ .

We have integrated the test problem with  $\alpha = 1/2, 1, 3/2$ , initial conditions corresponding to each of the four normal modes in (14) and (15) and several values of

<sup>4</sup>We have chosen a not very high value of  $\omega$  to obtain a clean figure where different branches appear clearly apart from each other. Higher values of  $\omega$  do not alter the topology of the eigenvalue crossings.

TABLE 2  
*Experimental orders of convergence for the RAI.*

	$\omega h \neq 2k\pi$			$\omega h = 2k\pi$		
	$\alpha = 1/2$	$\alpha = 1$	$\alpha = 3/2$	$\alpha = 1/2$	$\alpha = 1$	$\alpha = 3/2$
$p_1$	5/4	1/2	1/4	3/4	1/2	1/4
$q_1$	3/2	1	1/2	3/4	1/2	1/4
$p_2$	5/4	1/2	1/4	5/4	1/2	1/4
$q_2$	3/4	1/2	1/2	3/4	1/2	1/4
$H$	3/2	1	1/2	2	2	2
$H_{weak}$	3/2	1	1/2	3/2	1	1/2
$H_{strong}$	2	1	1/2	3/2	1	1/2

$\eta$  (including  $\eta = \pi$  and  $\eta = 2\pi$ ), and have measured the rate of convergence, in the limit (28), for each of the variables  $p_i$ ,  $q_i$ , the total energy  $H$  in (13), and the energies  $H_{weak}$  and  $H_{strong}$  defined in (20) and (21). For given values of  $\alpha$ , and  $\eta$ , the rates of convergence depend on the particular normal mode being integrated, and, for each variable, we have listed in Table 2 the *minimum rate* over all four modes. The rates of convergence found in this way turn out not to depend on  $\eta$  as long as  $\eta \neq 2k\pi$ ,  $k = 1, 2, \dots$  (nonresonance). From the table, it is clear that the second order of convergence one would naively expect from the reversibility of the method is far from manifesting itself and that an *order reduction* is present in the limit (28). For  $\alpha = 3/2$  (very strong spring and not very light second mass) the order reduction is very marked; the situation improves when  $\alpha$  is smaller. Note that a fortiori, for the various variables, the orders of convergence of the RAI *uniformly in  $\omega$*  cannot exceed the values in Table 2.

An illustration of the performance of the integrator is given in Figure 2, which corresponds to the intermediate case  $\alpha = 1$  used in [10], [11]. The initial condition is the second fast mode in (15) and  $h = 0.125$ ,  $\omega = 400$  (a nonresonant combination with  $\eta = 50$ ). Note that, even though the numerical method behaves in a stable way, it produces a misleading solution. In the plots, the components of the theoretical solution, whose period is  $\approx 0.01$ , have been represented only at grid points in order to allow a better comparison with their numerical counterparts. For the variable  $q_1$  and in agreement with our earlier discussion of the behavior of the normal modes, the amplitude of the true solution  $q_1(t)$  is negligible. For  $p_1(t)$ ,  $p_2(t)$ ,  $q_2(t)$ , the representation in Figure 2 introduces a stroboscopic effect because the fast zigzags of these functions in the intervals between grid points cannot be “seen.” In an alternative representation using a fine time grid, the graphs of these functions appear as almost solid horizontal bands whose vertical amplitudes coincide with those shown in the current figure.

The remainder of this section is devoted to an analysis of the mechanisms that give rise to the complicated pattern of orders of convergence shown in Table 2. Subsections 4.3 and 4.4 deal with errors in  $p_i$ ,  $q_i$  and subsection 4.5 with errors in energy.

**4.3. Errors in the numerical frequencies.** According to (23), from time  $t^0 = 0$  to time  $t^n = nh$ , the slow and fast phases of the true solution increase by amounts  $\Omega_- t^n$  and  $\Omega_+ t^n$ . For the numerical solution in (25), the increases in phase of the numerical solution are of the form  $\tilde{\Omega}_- t^n$  and  $\tilde{\Omega}_+ t^n$ , where  $\tilde{\Omega}_- h$  and  $\tilde{\Omega}_+ h$  denote the rotation angles of the numerical scheme in a single step, i.e., the arguments of the eigenvalues of the propagator  $\tilde{\Phi}(h, \omega)$ . Therefore the success of any numerical method very much hinges on the approximations that the frequencies  $\tilde{\Omega}_\pm$  of the propagator provide to the true frequencies  $\Omega_\pm$ .

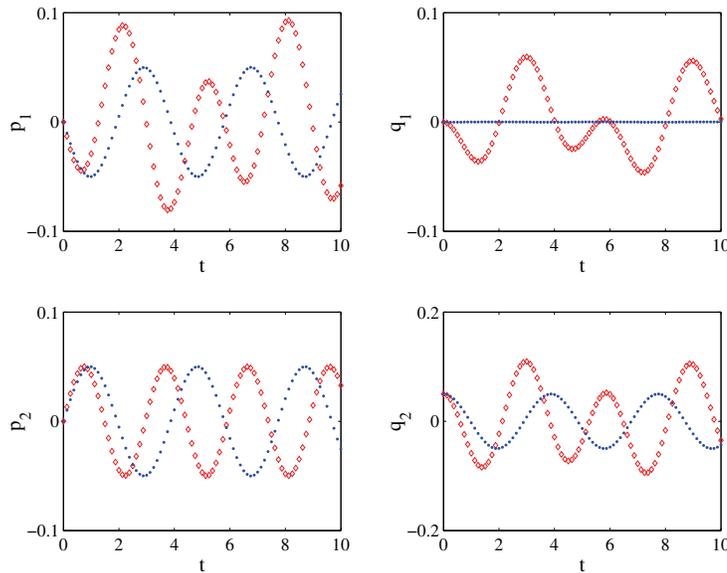


FIG. 2. True (dots) and RAI (diamonds) solutions.

Our next result, proved in Appendix A, shows the behavior of the numerical slow frequency.

PROPOSITION 2. For the RAI, in the limit (28) with  $\eta \neq 2k\pi$ ,  $k = 1, 2, \dots$ ,

$$(29) \quad \tilde{\Omega}_- = 1 - \frac{1 + \cos(\eta)}{4\eta^{2-\alpha}} h^{2-\alpha} + o(h^{2-\alpha}), \quad h \rightarrow 0,$$

and therefore, in view of (16),

$$\tilde{\Omega}_- - \Omega_- = \frac{1 - \cos(\eta)}{4\eta^{2-\alpha}} h^{2-\alpha} + o(h^{2-\alpha}), \quad h \rightarrow 0.$$

Since, under the hypothesis of this proposition,  $1 - \cos(\eta) \neq 0$ , we see that the numerical slow frequency is  $O(h^{2-\alpha})$  away from its true counterpart and this introduces an upper limit of  $2 - \alpha$  for the rates of convergence in the dynamic variables  $p_1$ ,  $q_1$ ,  $q_2$ . The effect on  $p_2$  is somewhat more subtle. In view of Table 1, the amplitude of the slow oscillations of  $p_2$  in solutions of bounded energy is only  $O(\omega^{-2+\alpha})$ , i.e.,  $O(h^{2-\alpha})$ . Hence, for  $1 < \alpha < 2$ , there is an order reduction in  $p_2$  from the expected 2 to  $2(2 - \alpha)$  (here  $2 - \alpha$  units come from approximating the true frequency and another  $2 - \alpha$  units from the small amplitude). For  $\alpha \leq 1$ , the quantity  $2(2 - \alpha)$  is not smaller than the nominal rate of convergence, and therefore the inaccuracy of the slow frequency does not entail an order reduction in this variable (see the second column of Table 3 below).

Figure 3, where again  $\alpha = 1$ , illustrates the behavior of  $\tilde{\Omega}_-$  as a function of  $\eta = \omega h$  for  $h = 1, 1/2, 1/4, 1/8$ . For fixed nonresonant  $\eta$ , i.e., for  $\eta \neq 2k\pi$ ,  $k = 1, 2, \dots$ , the numerical slow frequency approaches 1 as  $h \downarrow 0$ , in agreement with (29). However, for fixed  $h$  and  $\eta \rightarrow 2k\pi$ , we have  $\tilde{\Omega}_- \rightarrow 0$ . This could have been foreseen from Figure 1: the eigenvalue branches that for  $\eta$  away from a resonance  $\omega h = 2k\pi$  approximate the slow true eigenvalues come together at a resonance where they give rise to the double

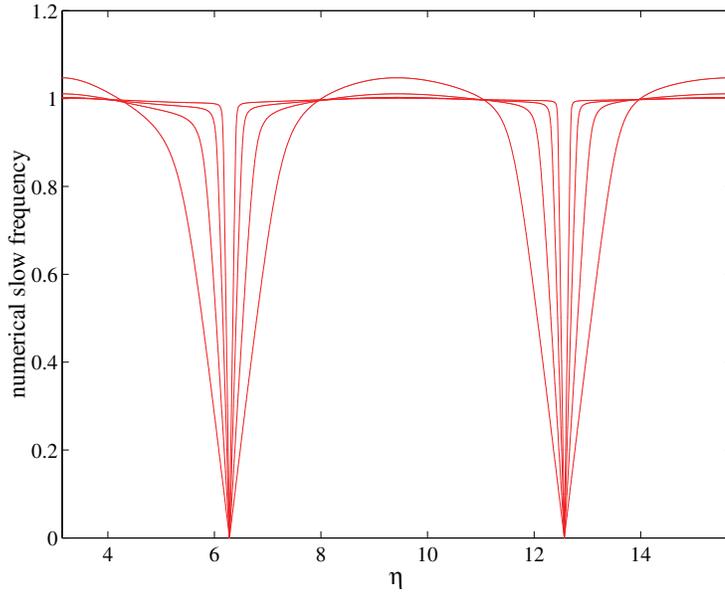


FIG. 3. Slow numerical frequency of the RAI as a function of  $\eta = \omega h$  for  $h = 1, 1/2, 1/4, 1/8$ .

eigenvalue  $\lambda = 1$ , whose argument equals 0 (see the next subsection). Figure 3 clearly bears out that the estimate (29) cannot be uniform in  $\eta$ .

We now turn to the analysis of the fast frequency  $\tilde{\Omega}_+$  where, for reasons that will become clear presently, it is simpler to word the result in terms of the eigenvalue argument  $\tilde{\Omega}_+ h$  rather than in terms of the frequency itself.

PROPOSITION 3. *In the limit (28),*

$$(30) \quad \tilde{\Omega}_+ h = \eta + \frac{\sin(\eta)}{2\eta^{2-\alpha}} h^{2-\alpha} + o(h^{2-\alpha}), \quad h \rightarrow 0,$$

and therefore, in view of (17),

$$\tilde{\Omega}_+ h - \Omega_+ h = \frac{\sin(\eta) - \eta}{2\eta^{2-\alpha}} h^{2-\alpha} + o(h^{2-\alpha}), \quad h \rightarrow 0.$$

From this result, for  $\alpha < 1$  the fast frequency is approximated by the RAI only to order  $1 - \alpha$ , while for  $\alpha \geq 1$  the numerical frequency does not even converge to the true value  $\tilde{\Omega}_+$ . As was the case for the variable  $p_2$  in the slow modes, this order reduction in the frequency is not automatically transferred to the variables  $p_i, q_i$ , all of which possess small amplitudes in a fast mode. Listed in the third column of Table 3 are the order reductions in the dynamic variables resulting from the inaccuracy of  $\tilde{\Omega}_+$ ; as before, the values displayed are found by combining the estimates of the error in the frequency with those of the sizes of the variables.

The table also contains the (rather complicated) combined effect of the inaccuracies of the slow and fast numerical frequencies. For  $\alpha = 1/2, 1, 3/2$  the formulae in the fourth column exactly reproduce the numbers in the left half ( $\eta \neq 2k\pi$ ) of Table 2; we conclude that, for nonresonant  $\eta$ , the orders of convergence in  $p_i, q_i$  found in the numerical experiments are totally explained by the lack of accuracy in the frequencies described in Propositions 2 and 3. In other words, for nonresonant  $\eta$ , the errors that

TABLE 3

Reduced orders of convergence for the RAI arising from the inaccuracy of the numerical frequencies. Stars indicate no order reduction. For  $\omega h = 2k\pi$ ,  $k = 1, 2, \dots$ , and due to eigenvector inaccuracies, there is a further reduction to  $1 - \alpha/2$  in the variables  $p_1, q_1, q_2$ .

	Slow frequency	Fast frequency	Combined effect
$p_1$	$2 - \alpha$	$2 - \frac{3\alpha}{2}, \alpha \leq 1$ $1 - \frac{\alpha}{2}, 1 \leq \alpha$	$2 - \frac{3\alpha}{2}, \alpha \leq 1$ $1 - \frac{\alpha}{2}, 1 \leq \alpha$
$q_1$	$2 - \alpha$	*** $\alpha \leq 2/3$ $3 - \frac{3\alpha}{2}, 2/3 \leq \alpha \leq 1$ $2 - \frac{\alpha}{2}, 1 \leq \alpha$	$2 - \alpha$
$p_2$	*** $\alpha \leq 1$ $4 - 2\alpha, 1 \leq \alpha$	$2 - \frac{3\alpha}{2}, \alpha \leq 1$ $1 - \frac{\alpha}{2}, 1 \leq \alpha$	$2 - \frac{3\alpha}{2}, \alpha \leq 1$ $1 - \frac{\alpha}{2}, 1 \leq \alpha$
$q_2$	$2 - \alpha$	$1 - \frac{\alpha}{2}, \alpha \leq 1$ $\frac{\alpha}{2}, 1 \leq \alpha$	$1 - \frac{\alpha}{2}, \alpha \leq 1$ $\frac{\alpha}{2}, 1 \leq \alpha \leq \frac{4}{3}$ $2 - \alpha, \frac{4}{3} \leq \alpha$

arise from the implied approximation  $\tilde{\mathcal{P}}$  to the exact matrix  $\mathcal{P}$  in (23) do not provide additional bottlenecks for the orders of convergence of the variables  $p_i, q_i$ .

**4.4. Errors in the propagator eigenvectors.** For resonant  $\eta$  Proposition 2 does not apply. The propagator at such values of  $\eta$  is given by

$$(31) \quad \tilde{\Phi} = \begin{bmatrix} 1 - \frac{1}{2}h^2 & -h + \frac{1}{4}h^3 & 0 & 0 \\ h & 1 - \frac{1}{2}h^2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ h & -\frac{1}{2}h^2 & 0 & 1 \end{bmatrix},$$

a matrix whose eigenvalues are those of the upper left  $2 \times 2$  block (i.e., those of the propagator of the Verlet method as applied to the slow Hamiltonian (20)) and the double eigenvalue 1. According to (16) and (17), these eigenvalues are, respectively,  $O(h^{3-\alpha})$  and  $O(h^{2-\alpha})$  away from the true  $\exp(\pm i\Omega_-h) \approx 1 \pm ih$  and  $\exp(\pm i\Omega_+h) \approx \exp(\pm i2k\pi h) = 1$ , so that the order reduction due to inaccuracies in the frequencies discussed above and shown in Table 3 also holds for resonant  $\eta$ .

Furthermore, for resonant  $\eta$ , there is an additional source of order reduction, linked to the inability of the numerical propagator to approximate accurately the eigenvectors of the true flow. From (31), it is clear that the left eigenspace of  $\tilde{\Phi}$  or  $\tilde{\Phi}^n$  associated with the slow eigenvalues coincides with the 2-plane of the variables  $p_1, q_1$  while the corresponding fast left eigenspace is the 2-plane of the variables  $p_2, q_2 - q_1$ . Thus the values of  $p_1^n, q_1^n$  evolve slowly and are not influenced by the initial conditions  $p_2^0, q_2^0$  and the values of  $p_2^n$  and  $q_2^n - q_1^n$  do not vary at all along the integration, as associated with the eigenvalue  $\lambda = 1$ . There is a stroboscopic effect at work: since the rate at which the solution is sampled coincides with the fast numerical frequency, fast modes at grid points  $t^n$  are seen as constant.

For the true flow in (23), it is easily shown that the linear combinations of variables that evolve slowly are of the form

$$(32) \quad (1 + \dots)p_1 + (1 + \dots)p_2, \quad (1 + \dots)q_1 + (\omega^{\alpha-2} + \dots)q_2,$$

while the fast combinations are given by

$$(33) \quad -(\omega^{\alpha-2} + \dots)p_1 + (1 + \dots)p_2, \quad (1 + \dots)q_2 - (1 + \dots)q_1$$

(for each coefficient only the leading term in the expansion in negative powers of  $\omega$  has been shown). We conclude that, at a resonance and in the limit (28), the structure of the numerical left eigenspaces is not consistent with that of the true flow:  $p_1 + p_2$  is a slow variable for the flow but not for the numerical method.

The preceding observation leads readily to the construction of an example that bears out the extra order reduction in  $p_1, q_1, q_2$  that takes place for resonant  $\eta$  (see Table 2, right). In fact, it is enough to choose the initial condition leading to the first *fast* true mode in (15). In the numerical solution, the variables  $p_1$  and  $q_1$  evolve as if the Verlet method were applied to the *slow* Hamiltonian (20); hence, for these two variables, the numerical and true solution will be completely out of phase and the errors will be of the same size as the solution itself, i.e.,  $O(h^{1-\alpha/2})$ . In addition,  $q_2^n = q_1^n$ , and this also implies an  $O(h^{1-\alpha/2})$  error in  $q_2$ .

**4.5. Energy behavior.** We end this section by looking at the errors in the energy  $H = H(p_1, p_2, q_1, q_2)$ . A naive analysis based on the errors in the variables  $p_i$  and  $q_i$  would lead to pessimistic conclusions, because the Lipschitz constants of  $H$  with respect to those variables grow with  $\omega$ . However, all of the numerical methods considered in this article make use of exact solutions of a rapidly oscillatory subsystem, and this implies that those substeps do not change the energy of the subsystem. As a consequence, the energy behavior is better than one may at first fear.

We consider first the resonant case. As discussed in the preceding subsection, in the numerical solution, the variables  $p_2$  and  $q_2 - q_1$  remain constant, and therefore  $H_{strong}$  in (21) will also stay at its initial value. The variables  $p_1$  and  $q_1$  evolve as they would in the application of the Verlet method to the Hamiltonian problem with energy (20), and this implies that  $H_{weak}$  varies along the integration by an amount  $O(h^2)$  (uniformly in  $\omega$ ). The conclusion is that  $H$ , which remains constant in the true solution, changes by an  $O(h^2)$  quantity in the numerical solution and no order reduction is present, as found experimentally in Table 2. However, the RAI does not mimic accurately the physics of the problem, because it achieves small errors in  $H$  essentially through an artificial conservation of  $H_{weak}$  and  $H_{strong}$  separately when these quantities are *not* conserved along the true solution.

Before leaving the resonant case, we point out that the  $O(h^{2-\alpha})$  experimental errors in  $H_{weak}$  reported in Table 2 arise in the integration of the slow modes. For these we have seen in Table 3 that  $p_1$  and  $q_1$  possess errors  $O(h^{2-\alpha})$  and in this way  $H_{weak}$  just inherits the order reduction from momentum and position. The errors in  $H_{strong}$  must necessarily be of size  $O(h^{2-\alpha})$ , since the total  $H = H_{weak} + H_{strong}$  possesses  $O(h^2)$  errors.

In the nonresonant case,  $H_{weak}$  also inherits the  $O(h^{2-\alpha})$  errors in  $p_1$  and  $q_1$ , but now there is no partial cancellation with the errors in  $H_{strong}$  and  $H$  suffers from an order reduction to  $2 - \alpha$ .

**5. Analysis of the IM and MIM.** The following treatment is patterned after that of the preceding sections; many details will be skipped. The application of the

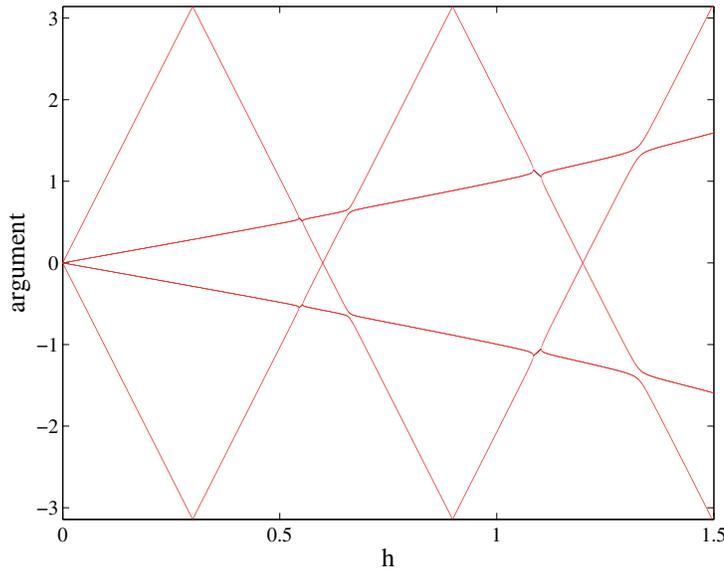


FIG. 4. Behavior of the argument of the eigenvalues of the IM propagator as functions of the step length  $h$  with  $\omega$  fixed.

IM and MIM to the test problem is, of course, based on the splitting of forces where the strong force  $f_s$  and weak force  $f_w$ , respectively, correspond to the strong and weak springs.

**5.1. Stability of the IM.** For the IM, the characteristic polynomial is again of the symmetric form (26), but now the coefficients are given by

$$(34) \quad a = 2 + 2c^* - h^2 + \frac{\omega^\alpha}{\Omega^2}h^2 - s^* \frac{\omega^\alpha}{\Omega^3}h,$$

$$(35) \quad b = 2 + 4c^* - 2c^*h^2 + 2c^* \frac{\omega^\alpha}{\Omega^2}h^2 - 2s^* \frac{\omega^\alpha}{\Omega^3}h,$$

where

$$\Omega = \sqrt{\omega^2 + \omega^\alpha} = \omega + \frac{1}{2}\omega^{-1+\alpha} + O(\omega^{-3+2\alpha})$$

is the frequency of the reduced system and

$$c^* = \cos(\Omega h), \quad s^* = \sin(\Omega h).$$

Figure 4 should be compared to Figure 1 and shows the behavior of the arguments of the four eigenvalues of the IM propagator as functions of  $h$  (once more,  $\omega = 10$  and  $\alpha = 1$ ). For  $k = 1, 2, \dots$ , as  $h \uparrow 2k\pi/\omega$  with fixed  $\omega$ , the fast eigenvalues meet the slow eigenvalues at a point  $h = h_k(\omega)$ . At this value of  $h$ , the numerical propagator possesses a pair of nonreal double eigenvalues  $\lambda, \lambda^{-1}$ , thus mimicking the behavior of the eigenvalues of the flow. However, for  $h$  just above the critical  $h_k(\omega)$ ,  $k = 1, 2, \dots$ , there is an instability interval  $h_k(\omega) < h < h_k^*(\omega)$  for which the numerical eigenvalues do not have unit modulus but are of the form  $\rho \exp(i\gamma), \rho \exp(-i\gamma), \rho^{-1} \exp(i\gamma), \rho^{-1} \exp(-i\gamma), \rho > 1$ ; in these intervals Figure 4 displays only two values  $\pm\gamma$  of the

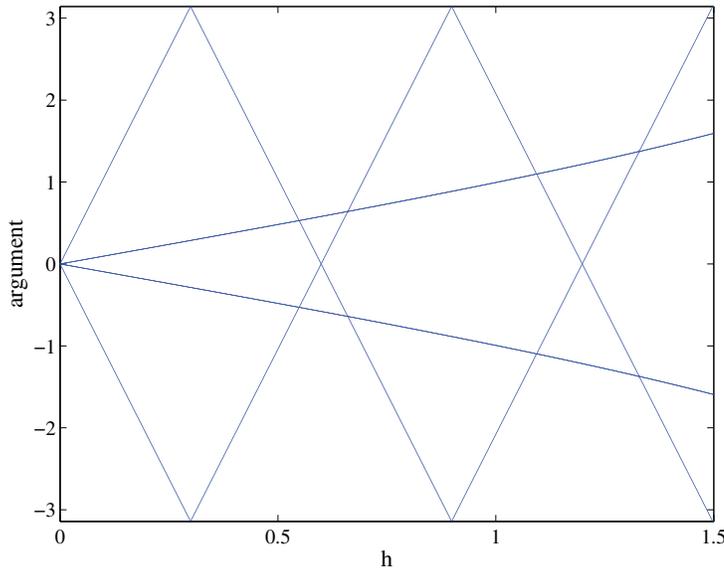


FIG. 5. Behavior of the argument of the eigenvalues of the MIM propagator (short filter) as functions of the step length  $h$  with  $\omega$  fixed.

argument rather than the four that exist for “most” values of  $h$ . On the other hand, and just as was the case for the RAI, the IM has near crossings just above  $h = 2k\pi$ ,  $k = 1, 2, \dots$ , instead of the crossings of the true flow.

Since the instability intervals  $h_k(\omega) < h < h_k^*(\omega)$  are present for arbitrarily large  $\omega$ , it follows that a stability condition similar to (27) does not exist for the IM. In other words, in order to obtain stable runs with the IM, it is necessary either to use small time steps  $h = O(\omega^{-1})$  or to be careful with the choice of  $h$  so as to avoid the (narrow) instability intervals.

**5.2. Stability of the MIM.** For the MIM, the characteristic equation is of the symmetric form (26) and very similar to that of the IM. In lieu of (34) and (35), the mollified version has

$$(36) \quad a = 2 + 2c^* - h^2 + \frac{\omega^\alpha}{\Omega^2} h^2 - \widehat{\psi}\widehat{\phi}s^* \frac{\omega^\alpha}{\Omega^3} h,$$

$$(37) \quad b = 2 + 4c^* - 2c^* h^2 + 2c^* \frac{\omega^\alpha}{\Omega^2} h^2 - 2\widehat{\psi}\widehat{\phi}s^* \frac{\omega^\alpha}{\Omega^3} h,$$

where  $\widehat{\psi}$  and  $\widehat{\phi}$  denote the Fourier transforms in (9) and (10) evaluated at  $\Omega h$ .

Figure 5 is similar to Figures 1 and 4 and uses the short weight function for  $\psi$  and  $\phi$ . At first sight it appears that the MIM is very successful in approximating the behavior of the true flow in Figure 1. However, it should be pointed out that a more careful analysis of the characteristic equation shows that the topology of the MIM eigenvalue branches is exactly the same as that of the IM, with instability intervals close to  $2k\pi/\omega$ ,  $k = 1, 2, \dots$ . The difference with the IM is that, for the MIM propagator, the stability intervals are extremely narrow and the near crossings are extremely tight. For the lowest ( $k = 1$ ) crossing in Figure 4, the instability interval of the IM is  $0.54403 < h < 0.55284$ , while the corresponding interval in Figure 5 is  $0.54821 < h < 0.54901$  with a width  $< 0.001$  (see Appendix B for further discussion).

**5.3. Order reduction.** As was the case for the RAI, the order reductions in the IM and MIM may be borne out by considering the limit (28). However, the details of the analysis become somewhat simpler by using instead the limit

$$(38) \quad h \rightarrow 0, \quad \Omega \rightarrow \infty, \quad h\Omega = \eta^* > 0,$$

with  $\eta^*$  constant. It turns out (see Appendix B) that for the IM and MIM there is no order reduction in the approximation of the slow frequency  $\Omega_-$  and an order reduction to  $3 - \alpha$  ( $\alpha > 1$ ) in the approximation of the fast frequency  $\Omega_+$ . The latter does not imply any order reduction in the positions  $q_1$  and  $q_2$  but, for  $\alpha > 4/3$ , leads to an order reduction to  $4 - 3\alpha/2$  in the momenta  $p_1$  and  $p_2$ . The remaining subsections are devoted to exploring the order reductions that result from inaccuracies in the eigenvectors.

**5.4. Eigenvectors in the nonresonant case.** For the IM, when  $\eta^* \neq 2k\pi$ ,  $k = 1, 2, \dots$ , the order reduction in the variables  $p_1$  and  $p_2$  is more severe than that implied by the lack of accuracy in the fast frequency we have just described. In fact, assume that the initial condition is taken from the second slow mode in (14) and consider the first time step  $t^0 \rightarrow t^1$ . For the variable  $p_2$  the true value is given by  $p_2(h) = -\xi_- \sin(\Omega_- h)$  and is therefore of size  $O(\omega^{-2+\alpha} h) = O(h^{3-\alpha})$ . For the numerical solution, the variable  $p_2$  does not change at kicks because the soft force acts only on  $p_1$ ; this alters the physics of the problem, where the soft spring acts on both  $p_1$  and  $p_2$  (note that  $d(p_1 + p_2)/dt = q_1$ ). The behavior of  $p_2^1$  in the numerical solution is then determined by the oscillation substep, i.e., essentially by the fast reduced system, and a simple computation, not reproduced here, reveals that, as a consequence of this inconsistency, the error  $p_2^1 - p_2(h)$  is of the same size as  $p_2(h)$  itself and  $p_2$  suffers from an order reduction to  $3 - \alpha$  if  $\alpha > 1$ . The variable  $p_1$  possesses the same reduction; however, the error in total momentum  $p_1 + p_2$  is still  $O(h^2)$  because, at leading order, the errors in  $p_1$  and  $p_2$  cancel each other (the numerical method pumps in at kicks the “right” amount of additional momentum, but applies it entirely to  $p_1$  instead of distributing between  $p_1$  and  $p_2$ , as would have been correct).

Interestingly enough this degraded performance manifests itself at the very first time step rather than being built up as the integration proceeds. However, this should not come as a surprise: while inaccuracies in the frequencies lead to a growing error phase as the number of steps increases, inaccuracies in the eigenvectors are independent of the length of the integration interval (the matrices  $\mathcal{P}$ ,  $\tilde{\mathcal{P}}$  in (23), (25) are independent of  $t$ ). For this reason and as discussed in [13] the analysis of this kind of algorithm cannot be performed in a step by step fashion: the errors after one step may be of the same size as the global errors.

The MIM also suffers from the order reduction described in this subsection: the requirements (9) and (10) operate only for  $\Omega h = 2k\pi$  and cannot be of help away from resonances.

**5.5. Eigenvectors at a resonance.** At a resonance, with  $\eta^* = 2k\pi$ ,  $k = 1, 2, \dots$ , the IM propagator is given by

$$\begin{bmatrix} 1 - \frac{\omega^2}{2\Omega^2} h^2 & -h + \frac{\omega^2}{4\Omega^2} h^3 & -\frac{\omega^2}{2\Omega^2} h^2 & 0 \\ \frac{\omega^2}{\Omega^2} h & 1 - \frac{\omega^2}{2\Omega^2} h^2 & \frac{\omega^2}{\Omega^2} h & 0 \\ 0 & 0 & 1 & 0 \\ \frac{\omega^2}{\Omega^2} h & -\frac{\omega^2}{2\Omega^2} h^2 & \frac{\omega^2}{\Omega^2} h & 1 \end{bmatrix};$$

TABLE 4  
 Reduced orders of convergence for the IM and MIM (stars denote no reduction).

	$\Omega h \neq 2k\pi$	$\Omega h = 2k\pi$	
	IM & MIM	IM	MIM
$p_1, p_2$	*** $\alpha \leq 1$ $3 - \alpha, \alpha > 1$	$2 - \alpha$	*** $\alpha \leq 4/3$ $4 - 3\alpha/2, \alpha > 4/3$
$q_1, q_2$	***	$2 - \alpha/2$	***

the corresponding eigenvalues are then 1 (double) and those of the upper left  $2 \times 2$  block, i.e., those of the Verlet scheme as applied to a harmonic oscillator with frequency  $\omega/\Omega = \Omega_- + O(h^{4-\alpha})$  (cf. Appendix B). For the double fast eigenvalue, the left eigenspace is the 2-plane of the variables  $p_2$  and  $q_2 - q_1$ . This is consistent with the situation for the true flow, where (33) shows that in the limit (38) the fast variables are also  $p_2$  and  $q_2 - q_1$ . On the other hand, after adding the first and third rows of the propagator, we see that

$$(39) \quad \begin{bmatrix} p_1 + p_2 \\ q_1 \end{bmatrix}^{n+1} = \begin{bmatrix} 1 - \frac{\omega^2}{2\Omega^2}h^2 & -h + \frac{\omega^2}{4\Omega^2}h^3 \\ \frac{\omega^2}{\Omega^2}h & 1 - \frac{\omega^2}{2\Omega^2}h^2 \end{bmatrix} \begin{bmatrix} p_1 + p_2 \\ q_1 \end{bmatrix}^n$$

so that the numerical slow left eigenspace is the plane of the variables  $p_1 + p_2, q_1$ . Again this is consistent in the limit (38) with the situation of the true flow (see (32)).

However, (32) and (33) also show that the discrepancy between the numerical and true combinations of slow and fast variables is  $O(\omega^{-2+\alpha})$  and this lack of accuracy gives rise to an order reduction. In fact, when the initial condition is taken from a slow mode, the values of  $p_2^n$  remain constant along the integration, while  $p_2(t^n)$  varies slowly. Hence the error in  $p_2$  will be of the same size as the variable  $p_2$  itself, i.e.,  $2 - \alpha$  (see Table 1). Since the combination  $p_1^n + p_2^n$  evolves accurately, the errors in  $p_1$  will also be of order  $2 - \alpha$ . Furthermore, when the initial condition corresponds to the second fast mode,  $q_1^n$  oscillates at the slow frequency, while the true  $q_1(t^n)$  does so at the fast frequency. Thus the errors in  $q_1$  will be of order  $2 - \alpha/2$  (Table 1). Since  $q_2^n - q_1^n$  evolves accurately, the variable  $q_2$  also suffers from an order reduction to  $2 - \alpha/2$ . A summary of these findings is given in the third column of Table 4.

Turning now to the MIM, under the assumptions (9) and (10) the propagator at a resonance is

$$\begin{bmatrix} 1 - \frac{\omega^4}{2\Omega^4}h^2 & -\frac{\omega^4}{\Omega^4}h + \frac{\omega^6}{4\Omega^6}h^3 & -\frac{\omega^4}{2\Omega^4}h^2 & -\frac{\omega^{2+\alpha}}{\Omega^4}h + \frac{\omega^{4+\alpha}}{4\Omega^6}h^3 \\ \frac{\omega^2}{\Omega^2}h & 1 - \frac{\omega^4}{2\Omega^4}h^2 & \frac{\omega^2}{\Omega^2}h & -\frac{\omega^{2+\alpha}}{2\Omega^4}h^2 \\ -\frac{\omega^{2+\alpha}}{2\Omega^4}h^2 & -\frac{\omega^{2+\alpha}}{\Omega^4}h + \frac{\omega^{4+\alpha}}{4\Omega^6}h^3 & 1 - \frac{\omega^{2+\alpha}}{2\Omega^4}h^2 & -\frac{\omega^{2\alpha}}{\Omega^4}h + \frac{\omega^{2+2\alpha}}{4\Omega^6}h^3 \\ \frac{\omega^2}{\Omega^2}h & -\frac{\omega^4}{2\Omega^4}h^2 & \frac{\omega^2}{\Omega^2}h & 1 - \frac{\omega^{2+\alpha}}{2\Omega^4}h^2 \end{bmatrix}.$$

By taking linear combinations of rows we find the two implied recursions for the combinations of variables associated with the left eigenspaces:

$$\begin{bmatrix} p_2 - \omega^{-2+\alpha}p_1 \\ q_2 - q_1 \end{bmatrix}^{n+1} = \begin{bmatrix} p_2 - \omega^{-2+\alpha}p_1 \\ q_2 - q_1 \end{bmatrix}^n,$$

TABLE 5  
*Experimental orders of convergence in energy for the IM (top) and MIM (bottom).*

	$\Omega h \neq 2k\pi$			$\Omega h = 2k\pi$		
	$\alpha = 1/2$	$\alpha = 1$	$\alpha = 3/2$	$\alpha = 1/2$	$\alpha = 1$	$\alpha = 3/2$
$H$	2	2	2	3/2	1	1/2
$H_{weak}$	2	2	3/2	3/2	1	1/2
$H_{strong}$	5/2	2	3/2	3/2	1	1/2

	$\Omega h \neq 2k\pi$			$\Omega h = 2k\pi$		
	$\alpha = 1/2$	$\alpha = 1$	$\alpha = 3/2$	$\alpha = 1/2$	$\alpha = 1$	$\alpha = 3/2$
$H$	2	2	2	2	2	2
$H_{weak}$	2	2	3/2	2	2	2
$H_{strong}$	5/2	2	3/2	7/2	3	5/2

$$\begin{bmatrix} p_1 + p_2 \\ \frac{\omega^2}{\Omega^2}(q_1 + \omega^{-2+\alpha}q_2) \end{bmatrix}^{n+1} = \begin{bmatrix} 1 - \frac{\omega^2}{2\Omega^2}h^2 & -h + \frac{\omega^2}{4\Omega^2}h^3 \\ \frac{\omega^2}{\Omega^2}h & 1 - \frac{\omega^2}{2\Omega^2}h^2 \end{bmatrix} \begin{bmatrix} p_1 + p_2 \\ \frac{\omega^2}{\Omega^2}(q_1 + \omega^{-2+\alpha}q_2) \end{bmatrix}^n.$$

In the second recursion note that  $\omega^2/\Omega^2 = (1 + \omega^{-2+\alpha})^{-1} \approx 1$  and that the  $2 \times 2$  matrix is the same one we encountered in (39). Therefore (and as proved in Appendix B in a different way), at a resonance, the IM and MIM share the same frequencies, but differ in the fast and slow combinations of dynamic variables; the MIM does a better job in that they are able to identify the leading term of each coefficient in the linear combinations in (32) and (33). The difference between the behavior of the methods is perhaps best seen in physical terms. For instance, the coordinate that evolves slowly is taken by the IM to be the position  $q_1$  of the first mass and by the MIM to be  $(1 + \omega^{-2+\alpha})^{-1}(q_1 + \omega^{-2+\alpha}q_2)$ , i.e., the abscissa of the center of mass of the whole system, which makes more physical sense. Since  $\alpha < 2$ , the abscissa of the center of mass approaches  $q_1$  as  $\omega$  increases; but, for  $\alpha$  close to 2 the term  $\omega^{-2+\alpha}q_2$  may be substantial and the performance of the IM is impaired by not taking it into account.

A more detailed analysis leads to the results displayed in the last column of Table 4.

**5.6. Energy behavior.** Table 5, which is similar to the lower part of Table 2, contains the experimental order of convergence of the energy errors for the IM and MIM. At a resonance, the IM exactly conserves  $p_2$  and  $q_2 - q_1$  and evolves the pair  $(p_1 + p_2, q_1)$  as if driven by the Verlet integration (39). By using the same argument we employed for the RAI integrator, we conclude that the errors in  $H, H_{weak}, H_{strong}$  are then  $O(h^{2-\alpha})$ . The MIM has no order reduction in  $H$  and  $H_{weak}$  and possesses extra accuracy in  $H_{strong}$ , facts that can be rigorously proved through an analysis of the fast and slow combinations of variables discussed in the preceding subsection.

In the nonresonant case, the IM and MIM share once more the same behavior. From Table 4, we know that, for  $\alpha > 1$ ,  $p_1$  suffers an order reduction to  $3 - \alpha$  and, of course,  $H_{weak}$  inherits the same reduction. Nevertheless the reduction is not transferred to the total energy  $H$ : a simple analysis, not reproduced here, reveals that the leading terms  $O(h^{3-\alpha})$  of the errors in the kinetic energies of the first and second mass cancel each other, as a consequence of the corresponding cancellation of the errors in the momenta  $p_1$  and  $p_2$  discussed in subsection 5.4.

**6. The case of a nonlight second mass.** So far, the constant  $\alpha$  in the test problem has been taken in the range  $0 < \alpha < 2$ . When  $\alpha = 2$ , we encounter a

physically different situation in that the second mass does not vanish in the limit  $\omega \rightarrow 0$ . However, all of the preceding analysis carries through with a number of minor adjustments (for instance, (16) has to be replaced by  $\Omega_- = 1/2 + O(\omega^{-2})$ , the variables  $p_1$  and  $p_2$  are no longer small in a fast mode, the IM does not provide a consistent approximation to the eigenvectors at a resonance, etc.), and the order reductions listed in Tables 3 and 4 also apply. For the IM, this means convergence uniform in  $\omega$  is restricted to first order in positions and zero order in momenta. For the MIM satisfying (9), (10) there is no order reduction in positions, but the momenta are approximated only to first order. In this way we recover the results in [5], [13].

**7. Conclusions.** We have considered a linear test problem for the analysis of numerical methods for highly oscillatory differential equations. While the success of a given method on the test problem does not necessarily imply similar successes in more general situations, it is clear that methods that have difficulties in integrating this simple model cannot be recommended for realistic applications.

Three integrators have been applied to the test problem and the following conclusions have emerged:

- The IM and MIM are not necessarily stable for  $h$  moderate and  $\omega$  large. The  $h$ -intervals of instability of the MIM have been found to be very small and, in any case, much smaller than those of the IM. The RAI does not suffer from any instabilities; its stability limit is dictated only by the period of the slow oscillations.
- We have shown that it is possible for an algorithm to suffer from severe order reductions when applied to the test problem. Numerical integrations in the presence of order reduction may be misleading: the output of the algorithm may show the variables evolving with a “reasonable” superposition of fast and slow oscillations, but it is well possible that the simulation has no resemblance to the true dynamics. In practice it is usually trivial to detect when a numerical integration of an oscillatory problem is unstable; it is enough to monitor the growth in the energy or in the dependent variables themselves. For this reason, we believe that the possibility of order reduction is more dangerous than the threat posed by instabilities. On the other hand, some authors point out that stable, inaccurate integrations may still be usable in some applications (for instance in sampling molecular conformations) while an unstable run is certainly of no value in any circumstances.
- For the model problem, errors are of two kinds. The algorithms may, of course, misrepresent the values of the various frequencies, but it is also possible for a numerical method to identify incorrectly the structure of the normal modes or, equivalently, the combinations of variables that evolve at the different frequencies present in the problem.
- All methods cope better with cases where high frequencies originate from small masses than with cases of stiff forces.
- The pattern of order reductions for the RAI has been shown to be extremely complicated; furthermore, the order reductions for this method are very marked. For the case  $\alpha = 1$  considered in [10] and [11] the errors in  $p_i$ ,  $q_i$  cannot be better than  $O(h^{1/2})$ , while the error in  $H$  cannot be better than  $O(h)$ .
- The IM suffers from an order reduction in the coordinates  $q$ ; nevertheless, for each value of  $\alpha$ , the errors in  $q$  are  $O(h)$  or better. The order reduction in the momenta  $p$  and in the energy  $H$  is more severe, and for  $\alpha \approx 2$  there is no

accuracy left in these variables.

- The MIM does not have any order reduction in  $q$  or  $H$ , and the reduction in  $p$  is, at worst, to  $O(h)$ , regardless of the value of  $\alpha$ . This difference in behavior between the IM and MIM is not due to a more accurate representation in the frequencies (indeed the IM and all members of the MIM class share the same frequencies at a resonance), but to a better identification of the normal mode structure.

Finally we point out that it is always possible to enhance the performance of the methods by incorporating into them analytic knowledge of the solution. For instance, the test problem is more easily integrated if  $p_1$  is replaced by  $p_1 + p_2$  as a dependent variable (cf. (32)), something that may or may not be obvious at first sight, and/or if  $q_2$  is replaced by the elongation  $q_2 - q_1$ . However, it is not clear to us to what extent such an analytic knowledge may be available before numerically integrating a given realistic problem.

**Appendix A: Proofs.** This appendix contains the proofs of the propositions.

*Proof of Proposition 1.* By introducing the auxiliary unknown

$$(40) \quad z = \lambda + \lambda^{-1},$$

the 4th degree equation (26) reduces to

$$z^2 - az + b - 2 = 0,$$

an equation whose discriminant  $\Delta = a^2 - 4(b - 2)$  can be rewritten in the form

$$\Delta = (-2(1 - c) + h^2)^2 + 2s^2\omega^{\alpha-2}h^2 + s^4\omega^{2\alpha-4} + 4s^2(1 - c)\omega^{\alpha-2}.$$

For  $c \neq 1$  the last three terms are  $> 0$ , and for  $c = 1$  and  $h > 0$  the first term is  $> 0$ . Hence, for  $h > 0$ , the equation for  $z$  has two real roots  $z_1 \neq z_2$ .

Each value of  $z$  gives rise, via (40), to two (mutually inverse) values of  $\lambda$ , and, with  $z$  real, these will have modulus 1 if and only if  $-2 \leq z \leq 2$ , a requirement on  $z$  that, in terms of the coefficients  $a, b$ , becomes

$$b + 2a + 2 \geq 0, \quad b - 2a + 2 \geq 0.$$

(Note that  $b \pm 2a + 2 = 0$  is equivalent to  $z = \mp 2$  or to a double root  $\lambda = \mp 1$ .)

Now  $b - 2a + 2 = 2h^2(1 - c)$ , and therefore a double root  $\lambda = 1$  occurs if and only if  $c = 1$ , i.e.,  $\omega h = 2k\pi$ ,  $k = 1, 2, \dots$ . On the other hand,

$$b - 2a + 2 = 2(1 + c)[4 - h^2 - 2(1 - c)\omega^{\alpha-2}];$$

the stability condition (27) ensures that the last factor is  $> 0$ . Then, a double root  $\lambda = -1$  may occur only for  $c = -1$ , i.e.,  $\omega h = (2k - 1)\pi$ ,  $k = 1, 2, \dots$ . When  $\lambda = 1$  or  $\lambda = -1$  are double eigenvalues, an inspection of the corresponding propagator matrix shows that they are not defective. (In fact, for  $\lambda = 1$ , this was pointed out in the discussion of (31).) This concludes the proof.

*Proof of Propositions 2 and 3.* The characteristic equation (26) may be rewritten as

$$(41) \quad P_0(\lambda) + Q(\lambda) = 0,$$

where

$$P_0(\lambda) = (\lambda^2 - 2c\lambda + 1)(\lambda^2 - (2 - h^2)\lambda + 1)$$

and  $Q$  denotes the perturbation

$$Q(\lambda) = \lambda^2(\lambda - 2 + \lambda^{-1})s^2\eta^{\alpha-2}h^{2-\alpha}.$$

For the unperturbed equation,  $P_0(\lambda) = 0$ , the roots are  $c \pm is$  and the solutions of

$$\lambda^2 - (2 - h^2)\lambda + 1 = 0,$$

i.e., the eigenvalues  $\lambda \approx 1 \pm ih - (1/2)h^2$  of the Verlet propagator when applied to the slow Hamiltonian (20).

We fix  $\eta$ , see  $h$  as a variable that approaches 0, and look for solutions  $\lambda = \lambda(h)$  of (26) that are perturbations of those of  $P_0 = 0$ . For the fast motions, a perturbed eigenvalue will be of the form

$$\lambda \approx (c + is)(1 + i\beta h^{2-\alpha}),$$

with  $\beta = \beta(\lambda)$  to be determined, and will lead to an angle of rotation

$$\tilde{\Omega}_+ h = \eta + \beta h^{2-\alpha} + o(h^{2-\alpha}).$$

By taking the ansatz for  $\lambda$  to the characteristic equation, we find the following equation for  $\beta$ :

$$\begin{aligned} & -4(c - 1)(-2ic^3 + 2c^2s + 2ic - s)\beta \\ & - 2\eta^{\alpha-2}(c + 1)(c - 1)^2(2c^2 + 2ics - 1) = 0; \end{aligned}$$

for  $c \neq \pm 1$  the coefficient of  $\beta$  does not vanish and (30) follows after simplification of the trigonometric functions.

When  $c = 1$  or  $c = -1$ , the preceding perturbation fails and this does not come as a surprise: the presence of double roots implies the inapplicability of the implicit function theorem. However, for these exceptional values of  $\eta$ , the perturbation  $Q$  vanishes due to the presence of the factor  $s$ , so that  $c + is$  is an exact eigenvalue of (26) and therefore (30) also holds. This completes the proof of Proposition 3.

For Proposition 2, we note that, for the unperturbed eigenvalue  $\lambda = 1 + ih + O(h^2)$ , the factor  $\lambda - 2 + \lambda^{-1}$  that features in the definition of  $Q$  is of size  $O(h^2)$  and therefore the perturbation  $Q(1 + ih + O(h^2))$  is  $O(h^{4-\alpha})$ . The ansatz to be used is then

$$\lambda \approx 1 + ih - \frac{1}{2}h^2 + i\beta h^{3-\alpha};$$

the exponent of  $h$  in the eigenvalue is one unit smaller than that in the perturbation because  $\lambda = 1$  is a double eigenvalue for  $h = 0$ .

The rest of the proof is similar to that of Proposition 3.

**Appendix B: The IM and MIM characteristic equation.** This appendix provides additional insights into the characteristic equation for the IM and MIM propagators and the corresponding eigenvalues.

We begin by noting that the IM coefficients (34) and (35) may be accommodated into the MIM format (36) and (37) by setting  $\hat{\psi} = \hat{\phi} = 1$ . This corresponds to the fact that the IM can be seen as the result of choosing, in the formulation of the family of mollified methods, the mollifier and averaging weight functions to be the Dirac delta function whose Fourier transform is  $\equiv 1$ ; see [13]. (Similarly a method with mollification but without averaging would have  $\hat{\phi} \equiv 1$ .)

As in the preceding appendix, the characteristic equation may be written in the form (41), where now

$$P_0(\lambda) = (\lambda^2 - 2c^*\lambda + 1) \left( \lambda^2 - \left( 2 - h^2 + \frac{\omega^\alpha}{\Omega^2} h^2 \right) \lambda + 1 \right)$$

and the perturbation is given by

$$Q(\lambda) = \lambda^2(\lambda - 2 + \lambda^{-1})\widehat{\psi}\widehat{\phi}s^*\frac{\omega^\alpha}{\Omega^3}h.$$

For the unperturbed equation  $P_0 = 0$  (an equation that is independent of the values of  $\widehat{\psi}$  and  $\widehat{\phi}$  and therefore is shared by all of the methods of the family including the IM), the roots are  $c^* \pm is^*$  and the zeros of

$$(42) \quad \lambda^2 - \left( 2 - h^2 + \frac{\omega^\alpha}{\Omega^2} h^2 \right) \lambda + 1,$$

i.e., the eigenvalues of the Verlet propagator as applied to a harmonic oscillator with a frequency  $(1 - \omega^\alpha/\Omega^2)^{1/2}$  that, as  $\omega \rightarrow \infty$ , differs from the true  $\Omega_-$  in terms  $O(\omega^{4-\alpha})$ . It is perhaps of some interest to note that here the effect of the strong spring on the slow frequency  $\Omega_-$  is taken care of in the unperturbed  $P_0 = 0$ , while for the RAI the unperturbed equation just reflects the slow Hamiltonian (20).

Since, for small  $h$  (uniformly in  $\omega$ ), the unperturbed equation  $P_0 = 0$  has roots of unit modulus, the instabilities of the IM and MIM are introduced only by the perturbation  $Q$ . For a MIM that satisfies the requirements (9) and (10),  $Q$ , as a function of  $h$  with fixed  $\omega$ , possesses at  $h = 2k\pi/\Omega$ ,  $k = 1, 2, \dots$ , a zero of multiplicity  $\geq 3$  ( $s^*$  and both Fourier transforms vanish). Therefore such a MIM may be expected to possess better stability properties than the plain IM, in agreement with the behavior in Figures 4 and 5. Furthermore, experiments not reported here show that weight functions whose Fourier transforms have multiple zeros at  $2k\pi$ ,  $k = 1, 2, \dots$ , yield a MIM whose instability intervals are extremely narrow.

We now turn to an investigation of the numerical slow and fast frequencies in the limit (38). We first consider the cases where  $s^* = 0$ . Then,  $Q = 0$  and, according to the discussion above, the propagator eigenvalues, regardless of the weight functions chosen, are  $c^* \pm is^*$  (which are  $O(h^{4-\alpha})$  away from the true  $\exp(\pm i\Omega_+ h)$ ) and the zeros of (42) (which are  $o(h^3)$  away from the true  $\exp(\pm i\Omega_- h)$ ). Hence, for  $s^* = 0$  or  $\eta^* = k\pi$ ,  $k = 1, 2, \dots$ , there is no order reduction in the slow frequency and an order reduction to  $3 - \alpha$  in the fast frequency, provided that  $\alpha > 1$ . When the latter estimate is combined with the solution sizes in Table 1, it turns out that there is no implied order reduction in  $q_1$ ,  $q_2$  and, for  $\alpha > 4/3$ , a reduction to  $4 - 3\alpha/2$  in  $p_1$  and  $p_2$ .

When  $\eta^* \neq k\pi$ ,  $k = 1, 2, \dots$ , a perturbation/implicit function analysis, similar to that in Appendix A, reveals that the order reduction is not worse than the one we have just described.

#### REFERENCES

- [1] G. ARIEL, B. ENQUIST, AND R. TSAI, *A multiscale method for highly oscillatory ordinary differential equations with resonance*, Math. Comp., to appear.
- [2] J. J. BIESIADECKI AND R. D. SKEEL, *Dangers of multiple-time-step methods*, J. Comput. Phys., 109 (1993), pp. 318–328.

- [3] W. E AND B. ENGQUIST, *The heterogeneous multiscale methods*, Commun. Math. Sci., 1 (2003), pp. 87–132.
- [4] B. ENGQUIST AND R. TSAI, *Heterogeneous multiscale methods for stiff ordinary differential equations*, Math. Comp., 74 (2005), pp. 1707–1742.
- [5] 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.
- [6] V. GRIMM AND M. HOCHBRUCK, *Error analysis of exponential integrators for oscillatory second-order differential equations*, J. Phys. A, 39 (2006), pp. 5495–5507.
- [7] H. GRUBMÜLLER, H. HELLER, A. WINDEMUTH, AND K. SCHULTEN, *Generalized Verlet algorithm for efficient molecular dynamics simulations with long-range interactions*, Mol. Sim., 6 (1991), pp. 121–142.
- [8] E. HAIRER AND CH. LUBICH, *Long-time energy conservation of numerical methods for oscillatory differential equations*, SIAM J. Numer. Anal., 38 (2000), pp. 414–441.
- [9] E. HAIRER, CH. LUBICH, AND G. WANNER, *Geometric Numerical Integration*, 2nd ed., Springer, Berlin, 2006.
- [10] B. LEIMKUHLER AND S. REICH, *A reversible averaging integrator for multiple time-scale dynamics*, J. Comput. Phys., 171 (2001), pp. 95–114.
- [11] B. LEIMKUHLER AND S. REICH, *Simulating Hamiltonian Dynamics*, Cambridge University Press, Cambridge, UK, 2005.
- [12] J. LI, P. G. KEVREKIDIS, C. W. GEAR, AND I. G. KEVREKIDIS, *Deciding the nature of the coarse equation through microscopic simulations: The baby-bathwater scheme*, SIAM Rev., 49 (2007), pp. 469–487.
- [13] J. M. SANZ-SERNA, *Mollified impulse methods for highly oscillatory differential equations*, SIAM J. Numer. Anal., 46 (2008), pp. 1040–1059.
- [14] J. M. SANZ-SERNA, *Modulated Fourier expansions and heterogeneous multiscale methods*, IMA J. Numer. Anal., to appear.
- [15] M. TUCKERMAN, B. J. BERNE, AND G. J. MARTYNA, *Reversible multiple time scale molecular dynamics*, J. Phys. Chem., 99 (1995), pp. 1990–2001.