SIAM J. NUMER. ANAL.
Vol. 34, No. 5, pp. 1792–1807, October 1997

008

ERROR ANALYSIS OF SYMPLECTIC MULTIPLE TIME STEPPING*

TODD R. LITTELL†, ROBERT D. SKEEL‡, AND MEIQING ZHANG§

Abstract. Most time-dependent problems obtain only limited benefit from variable stepsize, and to exploit more fully the existence of multiple time scales, it is necessary to use a multiplicity of stepsizes simultaneously. For example, many systems of ODEs have right-hand sides which naturally decompose into a sum of terms with different time scales, in which case the various terms can be sampled with different stepsizes—an idea called multiple time stepping (MTS). Addressed here is the question of how to choose these stepsizes in the context of molecular dynamics. This application and others are often modeled by Hamiltonian systems, for which symplectic integrators are often preferred because they introduce discretization error that can be very nearly interpreted as a perturbation to the scalar Hamiltonian function itself. The method analyzed here is a symplectic MTS method that has received recent attention. Analyzed is a problem that models the bonded interactions in molecular dynamics (MD). The method of analysis is quite novel in that it utilizes a slight change of variables that is more favorable to the numerical solution without seriously affecting the quantities of interest. The study concludes with a specific prescription for determining stepsize ratios.

Key words. leapfrog, Störmer, Verlet, multiple time steps, variable stepsize, symplectic integrator, molecular dynamics, method of modified equations

AMS subject classification. 65L20

PII. S0036142995287094

1. Introduction.

Through the years much emphasis has been placed on the importance of varying the stepsize in numerical integrators for ODEs. However, for large systems, and especially N-body systems, it is typical that during most of the integration there are processes somewhere in the system which require small stepsizes; so in practice the use of a universal variable stepsize leads to stepsizes that are usually small and sometimes very small. Much better efficiency seems possible if, in addition, different stepsizes are used for different "interactions." To be more precise, it happens that for many applications the right-hand side has a natural decomposition into a sum of terms, each term depending on only a few dependent variables. In N-body systems these terms represent interactions between only two or a few bodies; in chemical kinetics these terms represent reactions; in time-dependent partial differential equations discretized in space by finite elements, these terms are the elementwise contributions. In multiple time stepping (MTS) each of these terms has its own stepsize, meaning that some terms are sampled more often than others. By integrating more slowly changing terms with larger stepsizes, we can preserve the accuracy of the method and yet reduce the amount of computation. MTS methods should be distinguished from *multirate* methods, in which each dependent variable is integrated with its own stepsize. MTS was proposed first for astronomy [10] and later for molecular dynamics (MD) [20] a number of years ago.

*Received by the editors June 2, 1995; accepted for publication (in revised form) January 8, 1996. This work was performed in part at the Beckman Institute of the University of Illinois and supported in part by DOE/NSF grant DE-FG02-91-ER25099/DMS-9304268, NIH grant P41RR05969, and NSF/ARPA grant BIR-9318159. The work of the first and third authors was completed while at the University of Illinois. The work of the second author was also performed while visiting the Courant Institute and while at the University of Valladolid as Iberdrola Visiting Professor.

http://www.siam.org/journals/sinum/34-5/28709.html

†USACERL-PL-N, P.O. Box 9005, Champaign, IL 61826-9005 (littell@osiris.cso.uiuc.edu).

‡Department of Computer Science, University of Illinois at Urbana-Champaign, 1304 West Springfield Avenue, Urbana, IL 61801-2987 (skeel@cs.uiuc.edu).

§1226 West 31 Place, #2F, Chicago, IL 60608.

N-body problems are typically Hamiltonian systems of differential equations for which the right-hand sides are generated by a *scalar* Hamiltonian function. It is often appropriate to use *symplectic* methods to integrate such problems, but only recently was a symplectic MTS method discovered [2, 7, 22]. The question that we seek to answer here is how to choose the stepsize ratios to minimize computational cost (for a given level of accuracy) for this symplectic MTS method. We answer this question by doing what may be the first detailed error analysis of MTS. This analysis is carried to completion for a simplified model of the springlike bonded interactions in a classical mechanics model of macromolecular dynamics, yielding a relationship between stepsize ratios and corresponding stiffness ratios. The results and their implications for MD are given in the conclusion of this paper.

It is fairly easy to characterize the *efficiency* of integrators in terms of the stepsize (for explicit integrators at least) but not so easy to give a useful expression for the *accuracy* in terms of stepsize, method parameters, and problem parameters. In this work our starting point is the recently popularized idea of modified equations, the idea that the discretization error can be represented as a perturbation to the differential equations. For Hamiltonian systems one can go one step further and represent the error in terms of a perturbation to the Hamiltonian function itself—*if and only if* the integrator is symplectic [16]. One of the main points of this paper is that by *slightly* perturbing the computed solution we can make the perturbation to the Hamiltonian much smaller—a combined forward and backward error analysis.

The perturbations we consider here are theoretical transformations for the purpose of improving the error analysis. The use of computed transformations was first advocated by Butcher [3] for the purpose of improving the order of Runge–Kutta methods. In the context of symplectic integration this idea was first published by Rowlands [15] and mentioned in a more general form by Suzuki [21]. Details are given in the lively paper [24] on solar system dynamics and the followup paper by McLachlan [13].

The results obtained here have the weakness that they are based on an asymptotic theory—the smallest stepsize h times the highest frequency must be small—the very thing we hope to overcome in low-accuracy computations. We hope that future work will remove this limitation. This weakness of symplectic MTS is borne out in numerical experiments [8].

A large class of physical phenomena can be modeled by the Hamiltonian system

$$\begin{aligned} \dot{q} &= H_p(q,p), \\ \dot{p} &= -H_q(q,p), \end{aligned} \tag{1}$$

where the scalar function $H(q,p)$ is the Hamiltonian. Many systems of particles can appropriately be modeled by a separable system of the form

$$\begin{aligned} \dot{q} &= T_p(p), \\ \dot{p} &= -V_q(q), \end{aligned} \tag{2}$$

i.e., $H(q,p) = T(p) + V(q)$, where $q(t)$ and $p(t)$ represent the collection of particle positions and momenta, respectively.

For most applications, MD simulations are performed over long time intervals to gather statistical information about the system. This is in opposition to many short-term simulations where the main concern is calculating particle trajectories. Numerical experiments, which have been confirmed by recent theoretical results,

show that symplectic integrators perform better than nonsymplectic integrators for such long time simulations [5, 17]. Such numerical methods include the popular Störmer/leapfrog/Verlet method. Roughly speaking, symplectic integrators exactly solve a perturbed Hamiltonian system and, hence, conserve energy better than nonsymplectic integrators. A formal definition of the symplectic property and a theoretical discussion can be found in [1]. The simplest characterization of a symplectic map is the following: it is sufficient and necessary for the map χ to satisfy

$$(\nabla\chi)^{\mathrm{t}} J(\nabla\chi) = J,$$

where

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$$

and $\nabla\chi$ is the Jacobian of χ [1].

For the separable system (2) the Störmer/leapfrog/Verlet method is given by

$$\begin{aligned} P^{i+1/2} &= P^i - \tfrac{h}{2} V_q(Q^i), \\ Q^{i+1} &= Q^i + hT_p(P^{i+1/2}), \\ P^{i+1} &= P^{i+1/2} - \tfrac{h}{2} V_q(Q^{i+1}). \end{aligned} \tag{3}$$

It is a second-order numerical integrator which performs better in practice than the typical Runge–Kutta methods for long-term MD simulations. Some MD simulations involve tens of thousands of particles with a force vector which is very expensive to compute. The force terms may include sums of Coulomb and/or van der Waals forces. Due to the computational intensity of these simulations the "generalized" Verlet (MTS) methods were developed [7]. They reduce the required amount of computation by integrating forces that vary more slowly with a larger stepsize than those that vary more rapidly. To this end, a particle pair will fall into a certain distance class, which will in turn determine how frequently the corresponding force will be evaluated.

Although the common Störmer/leapfrog/Verlet method is symplectic, it is not the case that MTS extensions preserve the symplectic property. One extension that does preserve the symplectic property is the method called Verlet-I in [7, 2] and independently proposed in [22]. The practical value of this method for macromolecular dynamics is confirmed in [11] and [23].

This symplectic MTS method is the method of interest in this paper. For ease of exposition we consider this method restricted to two time steps, for which the Hamiltonian must be written as

$$H(q,p) = T(p) + W(q) + U(q),$$

where

$$\begin{aligned} T &= \text{total kinetic energy}, \\ W &= \text{potential energy corresponding to hard forces, and} \\ U &= \text{potential energy corresponding to soft forces}. \end{aligned}$$

The hard forces $-W_q(q)$ are evaluated with the stepsize h and the soft forces $-U_q(q)$ are evaluated with the stepsize Nh, where N is an integer. Thus, the computed

values (Q^{iN}, P^{iN}) represent the state of the system on the macro steps $t = i(Nh)$, and for $1 \le k \le N-1$ the computed values (Q^{iN+k}, P^{iN+k}) represent the state of the system on the intermediate micro steps. The algorithm generating these values is given by

$$
\begin{aligned}
\bar{P}^{iN} &= P^{iN} - \tfrac{1}{2} NhU_q(Q^{iN}) \\
&\text{for } k = 1, 2, \ldots, N, \\
P^{iN+k-1/2} &= \bar{P}^{iN+k-1} - \tfrac{1}{2} hW_q(Q^{iN+k-1}), \\
Q^{iN+k} &= Q^{iN+k-1} + hT_p(P^{iN+k-1/2}), \\
\bar{P}^{iN+k} &= P^{iN+k-1/2} - \tfrac{1}{2} hW_q(Q^{iN+k}), \\
P^{iN+N} &= \bar{P}^{iN+N} - \tfrac{1}{2} NhU_q(Q^{iN+N}).
\end{aligned} \tag{4}
$$

One interpretation of this algorithm is that a burst of energy (corresponding to the soft force) is added to the system at the beginning and end of every macro step.

The idea can be applied recursively: we can split W into a softer part and a harder part and apply the dual time step idea to each substep of the algorithm above. The general MTS leapfrog method partitions the potential according to $V(q) = V_0(q) + V_1(q) + \cdots + V_{L-1}(q)$, where the ith force vector $-\nabla V_i(q)$ is evaluated with time step h_i. In practice the stepsizes have been given by $h_i = 2^i h$, where h is the micro stepsize. In this paper we seek to answer the following question: given stepsize ratios 2^i, what rule of thumb can be used to partition the potential $V(q)$?

In the remainder of the paper we present modified problem analysis in section 2, symplectically equivalent methods in section 3, optimal stepsize ratio in section 4, and conclusions in section 5.

2. Modified problem analysis. Modified problem analysis is a technique for describing the discretization error of a numerical method which solves a differential system, e.g.,

$$\dot{z} = f(z), \quad z(0) = z_0$$

with true solution $z(t)$.

Let the solution calculated by a numerical method (with fixed time step h) be denoted by $\{Z^i\}$. The modified problem analysis seeks to find a modified problem of the form

$$\dot{Z} = \tilde{f}(Z), \quad Z(0) = \tilde{Z}^0,$$

with true solution $Z(t)$ such that $Z(t^i)$ and Z^i differ by terms of order h to a higher power than do $z(t^i)$ and Z^i. (Note that the dependence of Z, $\tilde{f}$, and $\tilde{Z}^0$ on h is suppressed in the notation.)

Examples of calculating modified problems are given in [6], while recent works [4, 17, 9] give a more systematic development of modified equations. These latter papers use the algebra of elementary differential trees for finding the modified equation.

Any symplectic method will yield a modified equation of the form $\dot{Z} = J \cdot \nabla \widetilde{H}$ for some perturbed Hamiltonian $\widetilde{H}$ [17].

2.1. The Störmer/leapfrog/Verlet method. We can rewrite the Hamiltonian system (1) as

$$\dot{z} = \{-H, \mathrm{id}\}(z), \tag{5}$$

where $z = (q, p)$, id is the identity *mapping* given by $\mathrm{id}(z) \equiv z$, and the Poisson bracket is defined as $\{F, G\} = F_q^{\mathrm{t}} G_p - F_p^{\mathrm{t}} G_q$. We define a differential operator by[1]

$$\{F, \cdot\}G = \{F, G\},$$

and then the solution of (5) can be written formally as

$$z(t) = \exp(\{-tH, \cdot\})\mathrm{id}(z(0)).$$

The Störmer/leapfrog/Verlet method (3) can be expressed as

$$\begin{aligned} Z^{i+1} &= \exp(\{-\tfrac{h}{2}V, \cdot\})\mathrm{id} \circ \exp(\{-hT, \cdot\})\mathrm{id} \circ \exp(\{-\tfrac{h}{2}V, \cdot\})\mathrm{id}(Z^i) \\ &= (\exp(\{-\tfrac{h}{2}V, \cdot\}) \exp(\{-hT, \cdot\}) \exp(\{-\tfrac{h}{2}V, \cdot\}))\mathrm{id}(Z^i). \end{aligned}$$

Note that in going from a composition of flows to a composition of operators the order must be reversed [16, p. 159]. From the Baker–Campbell–Hausdorff (BCH) formula [25] or [16, eq. (12.12)][2] we have

$$\begin{aligned} &\exp(\{-\tfrac{h}{2}V, \cdot\}) \exp(\{-hT, \cdot\}) \exp(\{-\tfrac{h}{2}V, \cdot\}) \\ &= \exp(\{-hT - hV - \tfrac{1}{12}h^3\{\{V, T\}, T\} + \tfrac{1}{24}h^3\{\{T, V\}, V\} + O(h^5), \cdot\}). \end{aligned} \tag{6}$$

Thus, we have a modified equation for scheme (6)

$$\dot{Z} = \{-\tilde{H}, \mathrm{id}\}(Z),$$

where

$$\tilde{H} = T + V + h^2(\tfrac{1}{12}T_p^{\mathrm{t}} V_{qq} T_p - \tfrac{1}{24}V_q^{\mathrm{t}} T_{pp} V_q).$$

This result is very much like a result obtained by Yoshida in [26] for another similar scheme, where $\tilde{H} = T + V + h^2\left(\frac{1}{12}V_q^{\mathrm{t}} T_{pp} V_q - \frac{1}{24}T_p^{\mathrm{t}} V_{qq} T_p\right)$.

2.2. The symplectic MTS method. We can write one macro step of the symplectic MTS method (4) in the following form:

$$\begin{aligned} &Z^{iN+N} = \\ &(\exp(\{-\tfrac{Nh}{2}U, \cdot\})(\exp(\{-\tfrac{h}{2}W, \cdot\}) \exp(\{-hT, \cdot\}) \exp(\{-\tfrac{h}{2}W, \cdot\}))^N \\ &\cdot \exp(\{-\tfrac{Nh}{2}U, \cdot\}))\mathrm{id}(Z^{iN}). \end{aligned}$$

This is a simple way to demonstrate that the method is symplectic. From the BCH formula

$$\begin{aligned} &\exp(\{-\tfrac{Nh}{2}U, \cdot\})(\exp(\{-\tfrac{h}{2}W, \cdot\}) \exp(\{-hT, \cdot\}) \exp(\{-\tfrac{h}{2}W, \cdot\}))^N \exp(\{-\tfrac{Nh}{2}U, \cdot\}) \\ &= \exp(\{-\tfrac{Nh}{2}U, \cdot\}) \exp(\{-Nh\tilde{H}, \cdot\}) \exp(\{-\tfrac{Nh}{2}U, \cdot\}) + O(h^4) \\ &= \exp(\{-Nh\hat{H}, \cdot\}) + O(h^4), \end{aligned}$$

[1] Following [16, p. 156], we use $\{F, \cdot\}$ rather than $\{\cdot, F\}$ as our Lie operator because the former admits a homomorphism [16, p. 157] from functions to Lie operators that permits the use of the same Baker–Campbell–Hausdorff formulas [16, eqs. (12.11), (12.12)] for both the Lie operator and the function that generates it. Criticism of this definition in [18] was ill considered.

[2] In the formula that is three lines from the bottom of page 163 in [16] there should be a factor of two multiplying the first term.

where

$$\begin{aligned}\hat{H} &= \tilde{H} + U + (Nh)^2\left(\tfrac{1}{12}\{\{U,\tilde{H}\},\tilde{H}\} - \tfrac{1}{24}\{\{\tilde{H},U\},U\}\right)\\ &= T + W + U + h^2\left(\tfrac{1}{12}T_p^{\mathrm{t}}W_{qq}T_p - \tfrac{1}{24}W_q^{\mathrm{t}}T_{pp}W_q\right)\\ &\quad + N^2h^2\left(\tfrac{1}{12}T_p^{\mathrm{t}}U_{qq}T_p - \tfrac{1}{12}U_q^{\mathrm{t}}T_{pp}W_q - \tfrac{1}{24}U_q^{\mathrm{t}}T_{pp}U_q\right).\end{aligned} \tag{7}$$

Equation (7) shows that the method is an $O(h^2)$ method.

The term $\frac{1}{12}h^2T_p^{\mathrm{t}}(W_{qq}+U_{qq})T_p$ can be viewed as a perturbation to the kinetic energy, the term $-\frac{1}{24}h^2W_q^{\mathrm{t}}T_{pp}W_q$ as a perturbation to W, the term $-\frac{1}{24}N^2h^2U_q^{\mathrm{t}}T_{pp}U_q$ as a perturbation to U, and the term $-\frac{1}{12}N^2h^2U_q^{\mathrm{t}}T_{pp}W_q$ as a perturbation to the potential energy in general. This last term is disconcerting. It is a factor of N more than one might expect from, say, the geometric average of the other two potential energy error terms.

3. Symplectically equivalent methods. Here we combine forward and backward error analysis to show that MTS is really more accurate than suggested. We show that the numerical solution never differs by more than some *modest multiple* of h^2 from the exact solution of some nearby problem that has more nearly the same structure as H than does

$$\begin{aligned}\hat{H}(Q,P) = T + W + U &+ \tfrac{h^2}{12}T_p^{\mathrm{t}}\left(W_{qq}+N^2U_{qq}\right)T_p\\ &- \tfrac{h^2}{24}\left(W_q^{\mathrm{t}}T_{pp}W_q + 2N^2U_q^{\mathrm{t}}T_{pp}W_q + N^2U_q^{\mathrm{t}}T_{pp}U_q\right).\end{aligned}$$

The forward error is modest in the sense that it is a function only of phase space and does not depend on the history of the trajectory.

Our approach was motivated by the question of how to measure the dominant h^2 error term in $\hat{H}$. It did not seem enough to measure, say, the size of this term. The tiniest change in the *value* of the Hamiltonian can produce enormous changes in the dynamics. For example, we can find an arbitrarily small perturbation to the Hamiltonian such that its gradient vanishes whenever $[q^{\mathrm{t}},p^{\mathrm{t}}]^{\mathrm{t}}$ are all exactly representable as floating-point numbers. (The set of all such values $[q^{\mathrm{t}},p^{\mathrm{t}}]^{\mathrm{t}}$ form a lattice in phase space.) Such a system would be perpetually motionless for any set of initial values that were exactly representable. In other words, an algorithm that returns a constant solution gives the exact solution of a Hamiltonian system that deviates an arbitrarily small amount from the given system.

So what we seem to need is some measure of the error in the gradient of the Hamiltonian. However, it is not meaningful to compare the q-gradient with the p-gradient, to compare force errors with velocity errors. Moreover, there may be cancellation of the error in the two gradients. If it were possible to express the error only as a perturbation to the force and not to T_p, the situation might be more satisfactory. Perhaps there exists a (canonical) transformation such that in the transformed coordinates only the force is perturbed. We show below by direct construction that this is the case up to $O(h^4)$. This same question was asked and answered for Störmer/leapfrog/Verlet in [15].

We shall follow the popular practice of using a generating function of the second kind $S(Q,p)$ to create a canonical transformation:

$$q = S_p(Q,p), \quad P = S_q(Q,p),$$

which constitutes implicit equations expressing the original variables (Q,P) as a function of the new variables (q,p). Others [21, 24, 13, 12] use Hamiltonian flows to

generate canonical transformations. We need to change things by only $O(h^2)$, so we choose

$$S(Q,p) = p^{\mathrm{t}}Q + h^2 s(Q,p) + O(h^4). \tag{8}$$

Hence,

$$\begin{aligned} q &= Q + h^2 s_p(Q,p) + O(h^4), \\ P &= p + h^2 s_q(Q,p) + O(h^4) \end{aligned}$$

which implies

$$\begin{aligned} Q &= q - h^2 s_p(q,p) + O(h^4), \\ P &= p + h^2 s_q(q,p) + O(h^4). \end{aligned}$$

Substituting into the perturbed Hamiltonian, we have

$$\begin{aligned} \hat{H} = T + W + U + h^2 T_p^{\mathrm{t}} &(\tfrac{1}{12}\left(W_{qq} + N^2 U_{qq}\right) T_p + s_q) \\ - h^2 (W_q + U_q)^{\mathrm{t}} s_p &- \tfrac{h^2}{24}\left(W_q^{\mathrm{t}} T_{pp} W_q + 2N^2 U_q^{\mathrm{t}} T_{pp} W_q + N^2 U_q^{\mathrm{t}} T_{pp} U_q\right) + O(h^4). \end{aligned} \tag{9}$$

The kinetic energy error term disappears if we choose

$$s = -\tfrac{1}{12} T_p^{\mathrm{t}} \left(W_q + N^2 U_q\right).$$

With this choice we have

$$\hat{H} = T + W + U + \tfrac{h^2}{24}\left(W_q^{\mathrm{t}} T_{pp} W_q + 2U_q^{\mathrm{t}} T_{pp} W_q + N^2 U_q^{\mathrm{t}} T_{pp} U_q\right) + O(h^4). \tag{10}$$

No longer is there a significant perturbation to the kinetic energy term; moreover, the uncomfortably large term $2N^2 U_q^{\mathrm{t}} T_{pp} W_q$ in the potential energy error has become a remarkably small term $2U_q^{\mathrm{t}} T_{pp} W_q$.

One might ask whether we might do even better with such a transformation. For example, can we transform away all of the dominant error term? We show that this is not possible in any extended region of phase space, in particular, not at a "nonsingular" stationary point of the Hamiltonian. To keep the algebra simple, we consider the case $W = V$ and $U \equiv 0$.

PROPOSITION 1. *For the Störmer/leapfrog/Verlet method it is not possible to transform away all of the dominant error term in any region of phase space that includes a stationary point of the Hamiltonian where its Hessian is nonsingular.*

Proof. For the generating function (8) the transformed modified Hamiltonian (9) becomes

$$\begin{aligned} \hat{H} = T + V + h^2 T_p^{\mathrm{t}} &(\tfrac{1}{12} V_{qq} T_p + s_q) \\ &- h^2 V_q^{\mathrm{t}} s_p - \tfrac{h^2}{24} V_q^{\mathrm{t}} T_{pp} V_q + O(h^4). \end{aligned}$$

We wish to choose s so that

$$T_p^{\mathrm{t}} s_q - V_q^{\mathrm{t}} s_p \equiv -\tfrac{1}{12} T_p^{\mathrm{t}} V_{qq} T_p + \tfrac{1}{24} V_q^{\mathrm{t}} T_{pp} V_q.$$

Differentiating with respect to q and p yields

$$\begin{aligned} s_{qq} T_p - V_{qq} s_p - s_{qp} V_q &\equiv -\tfrac{1}{12} V_{qqq} T_p T_p + \tfrac{1}{12} V_{qq} T_{pp} V_q, \\ T_{pp} s_q + s_{pq} T_p - s_{pp} V_q &\equiv -\tfrac{1}{6} T_{pp} V_{qq} T_p + \tfrac{1}{24} T_{ppp} V_q V_q. \end{aligned} \tag{11}$$

At a point (q^*, p^*) at which $V_q = 0$, $T_p = 0$, (11) implies

$$-V_{qq}s_p = 0, \quad T_{pp}s_q = 0.$$

Assuming that V_{qq} and T_{pp} are both nonsingular at (q^*, p^*), we must have

$$s_p = 0, \ s_q = 0 \quad \text{at } (q^*, p^*).$$

Differentiate the first equation of (11) with respect to q and the second with respect to p and evaluate at (q^*, p^*), and we get

$$\begin{aligned} -V_{qq}s_{pq} - s_{qp}V_{qq} &= \tfrac{1}{12}V_{qq}T_{pp}V_{qq}, \\ T_{pp}s_{qp} + s_{pq}T_{pp} &= -\tfrac{1}{6}T_{pp}V_{qq}T_{pp}. \end{aligned}$$

For brevity let $\boldsymbol{V} := V_{qq}$, $\boldsymbol{T} := T_{pp}$, and $\boldsymbol{S} := s_{qp}$. The equations above become

$$\begin{aligned} \boldsymbol{V}\boldsymbol{S}^{\mathrm{t}} + \boldsymbol{S}\boldsymbol{V} &= -\tfrac{1}{12}\boldsymbol{V}\boldsymbol{T}\boldsymbol{V}, \\ \boldsymbol{T}\boldsymbol{S} + \boldsymbol{S}^{\mathrm{t}}\boldsymbol{T} &= -\tfrac{1}{6}\boldsymbol{T}\boldsymbol{V}\boldsymbol{T}. \end{aligned} \tag{12}$$

Let (λ, v) be a generalized eigenpair so that

$$\boldsymbol{V}v = \lambda \boldsymbol{T}^{-1}v.$$

Premultiply the first of (12) by v^{t} and postmultiply by v to get

$$\lambda(v^{\mathrm{t}}\boldsymbol{T}^{-1}\boldsymbol{S}^{\mathrm{t}}v + v^{\mathrm{t}}\boldsymbol{S}\boldsymbol{T}^{-1}v) = -\tfrac{1}{12}\lambda^2(v^{\mathrm{t}}\boldsymbol{T}^{-1}v). \tag{13}$$

Premultiply the second of (12) by $v^{\mathrm{t}}\boldsymbol{T}^{-1}$ and postmultiply by $\boldsymbol{T}^{-1}v$ to get

$$(v^{\mathrm{t}}\boldsymbol{T}^{-1}\boldsymbol{S}^{\mathrm{t}}v + v^{\mathrm{t}}\boldsymbol{S}\boldsymbol{T}^{-1}v) = -\tfrac{1}{6}\lambda(v^{\mathrm{t}}\boldsymbol{T}^{-1}v). \tag{14}$$

Subtracting (13) from λ times (14) we then conclude that $\lambda = 0$, which is a contradiction. ∎

Another example of this phenomenon of symplectic equivalence, noted earlier by Suzuki [21], is the symplectic Euler method:

$$\begin{aligned} Q_{n+1} &= Q_n + hT_p(P_n), \\ P_{n+1} &= P_n - hV_q(Q_{n+1}). \end{aligned}$$

A transformation based on the second-kind generating function

$$S(Q, p) = p^{\mathrm{t}}Q - \tfrac{h}{2}V(Q)$$

yields endpoint Verlet, and

$$S(Q, p) = p^{\mathrm{t}}Q + \tfrac{h}{2}T(p)$$

yields midpoint Verlet.

4. Optimal stepsize ratio. We would like to develop a relationship between softness of the force and time step. To keep the analysis both tractable and rigorous, we investigate a model problem that has some resemblance to the bonded interactions of macromolecules. We consider a linear sequence of masses $m_1, m_2, \ldots, m_n, m_{n+1}$, such that masses i and $i+1$ are joined by a spring of stiffness κ_i with natural length l_i. If we let $q = [q_1, q_2, \ldots, q_n, q_{n+1}]^{\mathrm{t}}$ denote positions, and let

$$E = \begin{bmatrix} -1 & & & \\ 1 & -1 & & \\ & 1 & \ddots & \\ & & \ddots & -1 \\ & & & 1 \end{bmatrix}_{(n+1)\times n},$$

we can express the potential energy due to the compression of the ith spring as $\frac{1}{2}\kappa_i((E^{\mathrm{t}}q)_i - l_i)^2$. The potential energy of all terms is thus

$$W(q) + U(q) = \tfrac{1}{2}(E^{\mathrm{t}}q - l)^{\mathrm{t}}K(E^{\mathrm{t}}q - l),$$

where $l = [l_1, l_2, \ldots, l_n]^{\mathrm{t}}$ and $K = \mathrm{diag}(\kappa_1, \kappa_2, \ldots, \kappa_n)$. Introducing the "mask matrix" $J = \mathrm{diag}(j_1, j_2, \ldots, j_n)$, where

$$j_i = \begin{cases} 0, & \text{spring } i \text{ is hard,} \\ 1, & \text{spring } i \text{ is soft,} \end{cases}$$

allows us to express the soft potential energy as

$$U(q) = \tfrac{1}{2}(E^{\mathrm{t}}q - l)^{\mathrm{t}}JK(E^{\mathrm{t}}q - l).$$

The kinetic energy is

$$T(p) = \tfrac{1}{2}p^{\mathrm{t}}M^{-1}p,$$

where $M = \mathrm{diag}(m_1, m_2, \ldots, m_n, m_{n+1})$.

A calculation shows that the dominant error term $h^2u(q,p)$ of the Hamiltonian (10) has

$$\begin{aligned} u &= \tfrac{1}{24}((W+U)_q^{\mathrm{t}}M^{-1}(W+U)_q + (N^2-1)U_q^{\mathrm{t}}M^{-1}U_q) \\ &= \tfrac{1}{24}(E^{\mathrm{t}}q - l)^{\mathrm{t}}K^{1/2}AK^{1/2}(E^{\mathrm{t}}q - l), \end{aligned} \tag{15}$$

where

$$A = K^{1/2}(E^{\mathrm{t}}M^{-1}E + (N^2-1)JE^{\mathrm{t}}M^{-1}EJ)K^{1/2}. \tag{16}$$

The transformed numerical solution has an asymptotic expansion

$$q^{iN} = q(iNh) + h^2\delta(iNh) + O(h^4), \quad p^{iN} = p(iNh) + h^2\varepsilon(iNh) + O(h^4),$$

whose first two terms are the same as the first two terms of the solution of the Hamiltonian system with Hamiltonian $H(q,p) + h^2u(q,p)$. Thus, we have

$$\begin{aligned} &\dot{q} + h^2\dot{\delta} = M^{-1}(p + h^2\varepsilon) + O(h^4), \\ &\dot{p} + h^2\dot{\varepsilon} = -EK(E^{\mathrm{t}}(q + h^2\delta) - l) - \tfrac{h^2}{12}EK^{1/2}AK^{1/2}(E^{\mathrm{t}}q - l) + O(h^4). \end{aligned}$$

Hence,

$$\dot{q} = M^{-1}p, \quad \dot{p} = -EK(E^{\mathrm{t}}q - l), \tag{17}$$

$$\dot{\delta} = M^{-1}\varepsilon, \quad \dot{\varepsilon} = -EKE^{\mathrm{t}}\delta - \tfrac{1}{12}EK^{1/2}AK^{1/2}(E^{\mathrm{t}}q - l). \tag{18}$$

It will be shown that the dominant error term of the position $h^2\delta(t)$ has a part which is bounded for all time t and a part which is growing linearly with t.

We consider two measures of error:

1. $$\max_{H(q,p)=H_0} |h^2 u|.$$

   Note that $u = 0$ at an equilibrium $H_q = H_p = 0$, so there is no additive constant that artificially augments the error h^2u.

2. $$\max_{H(q(0),p(0))=H_0} \|M^{1/2}h^2\delta(t)\|_2,$$

   where the maximum is supposed to be taken over all initial values $q(0)$, $p(0)$ yielding some fixed energy H_0. This is the mass-weighted 2-norm of the dominant term in the position error.

Our choice for optimal N will depend on quantities $\omega_i^2 = (m_i^{-1} + m_{i+1}^{-1})\kappa_i$. The quantity ω_i is the frequency of the ith spring if unattached to the others. Particularly important is the maximum such frequency, which we denote by ω, and the maximum such frequency over soft springs only, which we denote by $\hat{\omega}$.

The spectral radius of the matrix A happens to be the key quantity in determining the size of our error measures, so the following is proved first.

LEMMA 1.

$$\max\{\omega^2, N^2\hat{\omega}^2\} \le \rho(A) \le 2\max\{\omega^2, N^2\hat{\omega}^2\}.$$

Proof. From the definition (16) of A we have

$$A\hat{e}_i = \begin{bmatrix} -m_i^{-1}\sqrt{\kappa_{i-1}\kappa_i}(1 + (N^2-1)j_{i-1}j_i) \\ (m_i^{-1} + m_{i+1}^{-1})\kappa_i(1 + (N^2-1)j_i) \\ -m_{i+1}^{-1}\sqrt{\kappa_i\kappa_{i+1}}(1 + (N^2-1)j_ij_{i+1}) \end{bmatrix} \begin{matrix} i-1 \\ i \\ i+1, \end{matrix}$$

where our notation does not show components known to be zero. Of course, $\rho(A) = \rho(K^{-1/2}AK^{1/2})$, and

$$(K^{-1/2}AK^{1/2})\hat{e}_i = \begin{bmatrix} -m_i^{-1}\kappa_i(1 + (N^2-1)j_{i-1}j_i) \\ (m_i^{-1} + m_{i+1}^{-1})\kappa_i(1 + (N^2-1)j_i) \\ -m_{i+1}^{-1}\kappa_i(1 + (N^2-1)j_ij_{i+1}) \end{bmatrix} \begin{matrix} i-1 \\ i \\ i+1, \end{matrix}$$

where $\hat{e}_i$ is the ith unit vector. Hence,

$$\begin{aligned} \rho(A) &\le \max_i \|(K^{-1/2}AK^{1/2})\hat{e}_i\|_1 \\ &\le \max_i 2(m_i^{-1} + m_{i+1}^{-1})\kappa_i(1 + (N^2-1)j_i) \\ &\le \max_i 2(\omega_i^2 + (N^2-1)\hat{\omega}_i^2), \qquad \text{where } \hat{\omega}_{i^2} = \omega_i^2 j_i, \\ &= 2\max\{\omega^2, N^2\hat{\omega}^2\}. \end{aligned}$$

Also,

$$\rho(A) \geq \max_i \tfrac{\hat{e}_i^{\mathrm{t}} A \hat{e}_i}{\hat{e}_i^{\mathrm{t}} \hat{e}_i} = \max_i(\omega_i^2 + (N^2-1)\hat{\omega}_i^2)$$
$$= \max\{\omega^2, N^2\hat{\omega}^2\}. \qquad \square$$

PROPOSITION 2. *The error in energy satisfies*

$$\max_{H(q,p)=H_0} |h^2 u| = \tfrac{h^2}{12} H_0 \rho(A) \leq \tfrac{h^2}{6} H_0 \max\{\omega^2, N^2\hat{\omega}^2\}.$$

Proof. Using (15), the problem becomes

$$\text{maximize } \tfrac{h^2}{24}(E^{\mathrm{t}}q - l)^{\mathrm{t}} K^{1/2} A K^{1/2}(E^{\mathrm{t}}q - l)$$
$$\text{subject to } \tfrac{1}{2}p^{\mathrm{t}}M^{-1}p + \tfrac{1}{2}(E^{\mathrm{t}}q - l)^{\mathrm{t}} K (E^{\mathrm{t}}q - l) = H_0.$$

Letting $\hat{q} = K^{1/2}(E^{\mathrm{t}}q - l)$, it simplifies to maximizing $\tfrac{h^2}{24}\hat{q}^{\mathrm{t}} A \hat{q}$ subject to $\tfrac{1}{2}p^{\mathrm{t}}M^{-1}p + \tfrac{1}{2}\hat{q}^{\mathrm{t}}\hat{q} = H_0$. Introducing a Lagrange multiplier μ for the constraint, we get the $2n+1$ equations

$$\tfrac{h^2}{12} A\hat{q} + \mu\hat{q} = 0,$$
$$\mu M^{-1} p = 0,$$
$$\tfrac{1}{2}p^{\mathrm{t}}M^{-1}p + \tfrac{1}{2}\hat{q}^{\mathrm{t}}\hat{q} = H_0.$$

Clearly $p = 0$. The solution $\hat{q} = 0$ corresponds to a minimum of the objective function because A is symmetric positive definite. A nonzero solution $\hat{q}$ must be an eigenvector of A for which the objective function has the value

$$\tfrac{h^2}{24}\lambda\hat{q}^{\mathrm{t}}\hat{q} = \tfrac{h^2}{24}\lambda \cdot 2H_0,$$

where λ is the eigenvector corresponding to $\hat{q}$. Hence, the maximum value of the objective function is

$$\tfrac{h^2}{12} H_0 \rho(A).$$

The inequality follows from the lemma. $\square$

PROPOSITION 3. *The mass-weighted* 2*-norm of the dominant term in the position error satisfies*

$$\tfrac{1}{n}\tfrac{h^2}{24}(t+O(1))\sqrt{2H_0}\rho(A) \leq \max_{H(q(0),p(0))=H_0} \|M^{1/2}h^2\delta(t)\| \leq \tfrac{h^2}{24}(t+O(1))\sqrt{2H_0}\rho(A)$$
$$\leq \tfrac{h^2}{24}(t+O(1))\sqrt{2H_0}\max\{\omega^2, N^2\hat{\omega}^2\}.$$

Remark. We believe that the factor $1/n$ in the lower bound can be improved using the structure of A, perhaps to $1/(n_{\mathrm{h}}+1)$.

Proof. To simplify (17) and (18), we do a change of variables

$$\bar{q} = M^{1/2}(q - q_0), \quad \bar{p} = M^{-1/2}p,$$
$$\bar{\delta} = M^{1/2}\delta, \quad \bar{\varepsilon} = M^{-1/2}\varepsilon,$$

where q_0 is chosen to satisfy $E^{\mathrm{t}}q_0 = l$. The energy becomes

$$H = \tfrac{1}{2}\bar{p}^{\mathrm{t}}\bar{p} + \tfrac{1}{2}\bar{q}^{\mathrm{t}}\Omega^2\bar{q},$$

where

$$\Omega^2 = M^{-1/2} E K E^{\mathrm{t}} M^{-1/2}$$

and the differential systems become

$$\dot{\bar{q}} = \bar{p}, \quad \dot{\bar{p}} = -\Omega^2 \bar{q}, \tag{19}$$

$$\dot{\bar{\delta}} = \bar{\varepsilon}, \quad \dot{\bar{\varepsilon}} = -\Omega^2 \bar{\delta} - \tfrac{1}{12} B A B^{\mathrm{t}} \bar{q}, \tag{20}$$

where

$$B = M^{-1/2} E K^{1/2}.$$

The solution of system (19) can be expressed

$$\begin{bmatrix} \bar{q}(s) \\ \bar{p}(s) \end{bmatrix} = \begin{bmatrix} \cos\Omega(s-t) & \Omega^{-1}\sin\Omega(s-t) \\ -\Omega\sin\Omega(s-t) & \cos\Omega(s-t) \end{bmatrix} \begin{bmatrix} \bar{q}(t) \\ \bar{p}(t) \end{bmatrix}$$

and that of system (20) is

$$\begin{bmatrix} \bar{\delta}(t) \\ \bar{\varepsilon}(t) \end{bmatrix} = \int_0^t \begin{bmatrix} \cos\Omega(t-s) & \Omega^{-1}\sin\Omega(t-s) \\ -\Omega\sin\Omega(t-s) & \cos\Omega(t-s) \end{bmatrix} \begin{bmatrix} 0 \\ -\frac{1}{12} B A B^{\mathrm{t}} \bar{q}(s) \end{bmatrix} ds.$$

Therefore,

$$\begin{aligned} \bar{\delta}(t) &= -\tfrac{1}{12} \textstyle\int_0^t \Omega^{-1}\sin\Omega(t-s)\, BAB^{\mathrm{t}} \begin{bmatrix} \cos\Omega(s-t) & \Omega^{-1}\sin\Omega(s-t) \end{bmatrix} ds \begin{bmatrix} \bar{q}(t) \\ \bar{p}(t) \end{bmatrix} \\ &= -\tfrac{1}{12} \textstyle\int_0^t B\hat{\Omega}^{-1}\sin\hat{\Omega}(t-s)\, A \begin{bmatrix} \cos\hat{\Omega}(s-t) & \hat{\Omega}^{-1}\sin\hat{\Omega}(s-t) \end{bmatrix} ds \begin{bmatrix} B^{\mathrm{t}}\bar{q}(t) \\ B^{\mathrm{t}}\bar{p}(t) \end{bmatrix}, \end{aligned}$$

where

$$\hat{\Omega}^2 = B^{\mathrm{t}} B.$$

This matrix is symmetric positive definite and can be diagonalized by a similarity transformation $Q^{\mathrm{t}}\hat{\Omega}Q$ with an orthogonal matrix, which we denote here by Q. Hence, we can write

$$\hat{\Omega} = Q\Lambda Q^{\mathrm{t}}$$

and

$$\bar{\delta}(t) = -\tfrac{1}{12} BQ \textstyle\int_0^t \Lambda^{-1}\sin\Lambda(t-s)\, Q^{\mathrm{t}} AQ \begin{bmatrix} \cos\Lambda(s-t) & \Lambda^{-1}\sin\Lambda(s-t) \end{bmatrix} ds \begin{bmatrix} Q^{\mathrm{t}}B^{\mathrm{t}}\bar{q}(t) \\ Q^{\mathrm{t}}B^{\mathrm{t}}\bar{p}(t) \end{bmatrix}.$$

The integral involves n^2 quadratures of the form $\int_0^t \sin\lambda_i(t-s)\cos\lambda_j(s-t)\,ds$ and n^2 of the form $\int_0^t \sin\lambda_i(t-s)\sin\lambda_j(s-t)\,ds$ where, because of the strict interlacing zeros of Sturm sequences, we know that the λ_i are distinct (as well as being positive). All of these integrals happen to be bounded as t increases with the exception of

$$\int_0^t \sin\lambda_i(t-s)\sin\lambda_i(s-t)\,ds = \tfrac{t}{2} + O(1).$$

Therefore, the mass-scaled position

$$\bar{\delta}(t) = -\tfrac{t}{24} BQ\Lambda^{-1} D\Lambda^{-1} Q^{\mathrm{t}} B^{\mathrm{t}} \bar{p}(t) + O(1),$$

where D denotes the diagonal matrix formed by setting the off-diagonal elements of $Q^{\mathrm{t}}AQ$ to zero, and

$$\sqrt{\bar{\delta}(t)^{\mathrm{t}}\bar{\delta}(t)} = \tfrac{t}{24}\sqrt{2}(\tfrac{1}{2}\bar{p}(t)^{\mathrm{t}} BQ\Lambda^{-1} D^2\Lambda^{-1} Q^{\mathrm{t}} B^{\mathrm{t}} \bar{p}(t))^{1/2} + O(1).$$

Maximizing a function of $q(t)$, $p(t)$ subject to $H(q(0), p(0)) = H_0$ is equivalent to maximizing the same function of q, p subject to $H(q, p) = H_0$ because of the invariance of $H(q(t), p(t))$. We wish to

$$\begin{aligned} &\text{maximize} \quad \tfrac{1}{2}\bar{p}^{\mathrm{t}} BQ\Lambda^{-1} D^2\Lambda^{-1} Q^{\mathrm{t}} B^{\mathrm{t}} \bar{p} \\ &\text{subject to} \quad \tfrac{1}{2}\bar{p}^{\mathrm{t}}\bar{p} + \bar{q}^{\mathrm{t}}\Omega^2\bar{q} = H_0. \end{aligned}$$

From the Lagrangian

$$\tfrac{1}{2}\bar{p}^{\mathrm{t}} BQ\Lambda^{-1} D^2\Lambda^{-1} Q^{\mathrm{t}} B^{\mathrm{t}} \bar{p} + \mu(H_0 - \tfrac{1}{2}\bar{p}^{\mathrm{t}}\bar{p} - \tfrac{1}{2}\bar{q}^{\mathrm{t}}\Omega^2\bar{q}),$$

we obtain the $2n + 1$ equations

$$\begin{aligned} BQ\Lambda^{-1} D^2\Lambda^{-1} Q^{\mathrm{t}} B^{\mathrm{t}} \bar{p} &= \mu\bar{p}, \\ \Omega^2\bar{q} &= 0, \\ \tfrac{1}{2}\bar{p}^{\mathrm{t}}\bar{p} + \tfrac{1}{2}\bar{q}^{\mathrm{t}}\Omega^2\bar{q} &= H_0. \end{aligned}$$

The maximum value, hence, satisfies

$$\tfrac{1}{2}\bar{p}^{\mathrm{t}} BQ\Lambda^{-1} D^2\Lambda^{-1} Q^{\mathrm{t}} B^{\mathrm{t}} \bar{p} = \tfrac{1}{2}\mu\bar{p}^{\mathrm{t}}\bar{p} = \mu H_0,$$

where

$$D^2(\Lambda^{-1} Q^{\mathrm{t}} B^{\mathrm{t}} \bar{p}) = \mu(\Lambda^{-1} Q^{\mathrm{t}} B^{\mathrm{t}} \bar{p}).$$

Therefore, $\mu = \rho(D^2) = \rho(D)^2$, and

$$\max \sqrt{\bar{\delta}(t)^{\mathrm{t}}\bar{\delta}(t)} = \tfrac{t}{24}\sqrt{2H_0}\rho(D) + O(1).$$

Also, because $Q^{\mathrm{t}}AQ$ is symmetric positive definite

$$\rho(D) = \max_i D_{ii} = \max_i (Q^{\mathrm{t}}AQ)_{ii} \le \rho(Q^{\mathrm{t}}AQ) = \rho(A).$$

To prove the lower bound, we note that

$$\rho(D) = \max_i D_{ii} \ge \tfrac{1}{n}\mathrm{trace}(D) = \tfrac{1}{n}\mathrm{trace}(A) \ge \tfrac{1}{n}\rho(A). \qquad \square$$

Now we determine the optimal choice of N. In both cases

$$\text{error} = \theta h^2 \text{const} \cdot \max\{\omega^2, N^2\hat{\omega}^2\},$$

where $\theta \le 1$ and const $= H_0/6$ or $t\sqrt{2H_0}/24$. The stepsize h should be chosen so that

$$\text{error} = tol,$$

where *tol* is the error tolerance. Solving for the reciprocal of the stepsize yields

$$\tfrac{1}{h} = \sqrt{\theta \tfrac{\text{const}}{tol}} \cdot \max\{\omega, N\hat{\omega}\}.$$

Let n_{h} be the number of hard springs and n_{s} the number of soft springs. The computational cost per unit time is reasonably represented as

$$\tfrac{n_{\text{s}}}{Nh} + \tfrac{n_{\text{h}}}{h} = \sqrt{\theta \tfrac{\text{const}}{tol}} \cdot \max\{(\tfrac{n_{\text{s}}}{N} + n_{\text{h}})\omega, (n_{\text{s}} + n_{\text{h}}N)\hat{\omega}\}.$$

The maximum is that of two functions of N, the first decreasing and the second increasing as N increases. This is minimized when the two functions are equal, yielding the choice

$$N = \omega/\hat{\omega}.$$

However, we have neglected the dependence of θ on N. For error measure one (error in Hamiltonian) our objective function may be off by a factor of $2^{1/2}$. So, overlooking the fact that N must be an integer, it is possible that the optimal value of N yields a cost which is smaller by a factor of $2^{1/2}$ than the cost with $N = \omega/\hat{\omega}$.

It is of interest to examine the savings that are achieved with near-optimal MTS. We want to compare our cost

$$\tfrac{n_{\text{s}}}{Nh} + \tfrac{n_{\text{h}}}{h} = ((\hat{\omega}/\omega)n_{\text{s}} + n_{\text{h}})/h$$

with the cost where $N = 1$

$$(n_{\text{s}} + n_{\text{h}})/h.$$

We have a savings in computational cost by a factor

$$\tfrac{(\hat{\omega}/\omega)n_{\text{s}}+n_{\text{h}}}{n_{\text{h}}+n_{\text{s}}} = \tfrac{n_{\text{h}}}{n_{\text{h}}+n_{\text{s}}} + \tfrac{\hat{\omega}}{\omega}\tfrac{n_{\text{s}}}{n_{\text{h}}+n_{\text{s}}},$$

so there is a great saving if both $n_{\text{h}} \ll n_{\text{s}}$ and $\hat{\omega} \ll \omega$. This conclusion is what is to be expected.

5. Conclusion. We analyzed a symplectic dual time stepping method applied to a one-dimensional model problem of varying masses connected by springs of varying stiffnesses. The springs are partitioned into slow and fast springs and the stepsize for the slow springs is taken to be N times greater than that for the fast springs. We considered as measures of accuracy the error in energy and the error in the 2-norm of the mass-weighted positions. We obtained a prescription for a near optimal N that depends on the frequencies of each mass–spring–mass system considered in isolation from the others. The prescription is

$$N = \omega/\hat{\omega},$$

where ω is the maximum such frequency over all springs and $\hat{\omega}$ is the maximum over all slow springs.

This choice of N is also suggested by the numerical experiments with linear forces in [2]. Similarly, experiments in [14] support the theoretical choice of N for minimal error in energy and in position.

We also have done analysis (not published) showing that if we instead use the errors in the mass-weighted accelerations or the error in the kinetic energy as our measure of accuracy, then the near optimal choice becomes $N = (\omega/\hat{\omega})^{3/2}$.

The formula we recommend, $N = \omega/\hat{\omega}$, is consistent with restrictions expected to be necessary for numerical stability, namely, that $h^2\omega^2$ and $Nh^2\hat{\omega}^2$ be equally restricted—certainly if slow and hard springs were only very weakly coupled.

The covalent bond structure of molecules is not far removed from the one-dimensional model analyzed. Actual molecules have branches but usually only a few loops. The atom chains that constitute molecules are normally full of bends, which reduces the impact of coupling of bondlengths. Hence, we can with some confidence apply the result for the model problem to individual bonded interactions in MD by associating ω with the maximum value of the square root of the spectral radius of the mass-weighted Hessian of the interaction.

For nonbonded interactions the Hessian is unbonded, and one should apply MTS only after creating an artificial splitting of each nonbonded interaction into a (conceptually) infinite sum of interactions with increasing bounds on the Hessian and correspondingly smaller stepsizes [19]. (This partitioning is tantamount to using a variable stepsize for the interaction and at the same time preserving symplecticness.) One might then apply the formula for N to such a situation but with less confidence because even with bounded Hessians these forces are not nearly linear as are most of the bonded forces. Nonetheless, this rule is a plausible one to use until more is known.

REFERENCES

[1] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd ed., Springer-Verlag, New York, 1989.

[2] J. J. Biesiadecki and R. D. Skeel, *Dangers of multiple-time-step methods*, J. Comput. Phys., 109 (1993), pp. 318–328.

[3] J. C. Butcher, *The effective order of Runge-Kutta methods*, in Conference on the Numerical Solution of Differential Equations, A. Dold, Z. Heidelberg, and B. Eckmann, eds., Lecture Notes in Mathematics 109, Springer-Verlag, New York, 1969, pp. 133–139.

[4] M. Calvo, A. Murua, and J. Sanz-Serna, *Modified equations for ODEs*, Contemp. Math., 172 (1994), pp. 63–74.

[5] J. Candy and W. Rozmus, *A symplectic integration algorithm for separable Hamiltonian functions*, J. Comput. Phys., 92 (1991), pp. 230–256.

[6] D. Griffiths and J. Sanz-Serna, *On the scope of the method of modified equations*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 994–1008.

[7] H. Grubmüller, H. Heller, A. Windemuth, and K. Schulten, *Generalized Verlet algorithm for efficient molecular dynamics simulations with long-range interactions*, Molecular Simulation, 6 (1991), pp. 121–142.

[8] H. Grubmüller and P. Tavan, *Efficient algorithms for molecular dynamics simulations of proteins: How good are they?* 1997, submitted.

[9] E. Hairer, *Backward error analysis of numerical integrators and symplectic methods*, Ann. of Numer. Math., 1 (1994), pp. 107–132.

[10] A. Hayli, *Le problème des N corps dans un champ extérieur application a l'évolution dynamique des amas ouverts-I*, Bull. Astronomique, 2 (1967), pp. 67–89.

[11] D. D. Humphreys, R. A. Friesner, and B. J. Berne, *A multiple-time-step molecular dynamics algorithm for macromolecules*, J. Phys. Chem., 98 (1994), pp. 6885–6892.

[12] M. López-Marcos, J. M. Sanz-Serna, and R. D. Skeel, *Cheap enhancement of symplectic integrators*, in Numerical Analysis 1995, D. F. Griffiths and G. A. Watson, eds., Longman Group, London, 1996, pp. 107–122.

[13] R. I. McLachlan, *More on symplectic correctors*, in Integration Algorithms and Classical Mechanics, J. E. Marsden, G. W. Patrick, and W. F. Shadwick, eds., Fields Institute Communications, Vol. 10, Fields Institute, AMS, Providence, RI, July 1996, to appear.

[14] D. J. Hardy and D. I. Okunbor, *Symplectic multiple-time step integration*, submitted; also D. J. Hardy, *Multiple Time Step Symplectic Integration*, M. S. thesis, Department of Computer Science, University of Missouri, Rolla, MO, 1997.
[15] G. Rowlands, *A numerical algorithm for Hamiltonian systems*, J. Comput. Phys., 97 (1991), pp. 235–239.
[16] J. Sanz-Serna and M. Calvo, *Numerical Hamiltonian Problems*, Chapman and Hall, London, 1994.
[17] J. M. Sanz-Serna, *Symplectic integrators for Hamiltonian problem: An overview*, Acta Numerica, 1 (1992), pp. 243–286.
[18] R. D. Skeel, review: *Numerical Hamiltonian Problems* by J. M. Sanz-Serna and M. P. Calvo, SIAM Rev., 37 (1995), pp. 277–279.
[19] R. D. Skeel and J. J. Biesiadecki, *Symplectic integration with variable stepsize*, Ann. Numer. Math., 1 (1994), pp. 191–198.
[20] W. B. Streett, D. Tildesley, and G. Saville, *Multiple time step methods in molecular dynamics*, Mol. Phys., 35 (1978), pp. 639–648.
[21] M. Suzuki, *Improved Trotter-like formula*, Phys. Lett. A, 180 (1993), pp. 232–234.
[22] M. Tuckerman, B. J. Berne, and G. J. Martyna, *Reversible multiple time scale molecular dynamics*, J. Chem. Phys, 97 (1992), pp. 1990–2001.
[23] M. Watanabe and M. Karplus, *Simulation of macromolecules by multiple-time-step methods*, J. Phys. Chem., 99 (1995), pp. 5680–5697.
[24] J. Wisdom, M. Holman, and J. Touma, *Symplectic correctors*, in Integration Algorithms and Classical Mechanics, J. E. Marsden, G. W. Patrick, and W. F. Shadwick, eds., Fields Institute Communications, Vol. 10, Fields Institute, AMS, Providence, RI, July 1996, to appear.
[25] H. Yoshida, *Construction of higher order symplectic integrators*, Phys. Lett. A, 150 (1990), pp. 262–268.
[26] H. Yoshida, *Recent progress in the theory and application of symplectic integrators*, Celest. Mech. Dynam. Astron., 56 (1993), pp. 27–43.