SIAM J. SCI. COMPUT.
Vol. 18, No. 1, pp. 223–238, January 1997

012

EXPLICIT SYMPLECTIC INTEGRATORS USING HESSIAN–VECTOR PRODUCTS*

M. A. LÓPEZ-MARCOS†, J. M. SANZ-SERNA†, AND ROBERT D. SKEEL‡

Abstract. In 1991 Rowlands proposed an effectively fourth-order, effectively two-stage, explicit symplectic integrator based on using a Hessian–vector product to modify the force evaluation in the leapfrog method, and evidence indicates that for modest accuracy this method is highly competitive. Here we explore the possible existence of even more efficient fourth-order explicit symplectic integrators, also based on the use of Hessian–vector products and the concept of effective order. First it is shown that the cost of a force evaluation plus a Hessian–vector product is less than twice the cost of the force alone for a sum of two-body interactions. Then a new method is found that is generally better than both the method of Rowlands and that of Calvo, according to both a theoretical measure of the error and limited numerical experiments. The basic motivation behind the new method is quite simple: do a Hessian–vector computation only every other step, significantly cutting costs while only marginally increasing the error. The idea of effective order means that we allow for both the possibility of preprocessing the initial values before application of the basic method and the possibility of postprocessing the values obtained by the basic method, but at output points only. For some applications processing is unnecessary, but in any case processing has been shown to be possible at low additional cost. The derivation of the new method illustrates how to simplify by the use of H-series the determination of parameters for methods of increased effective accuracy.

Key words. leapfrog, multiderivative method, molecular dynamics, symplectic integration, method of modified equations

AMS subject classification. 65L20

CR subject classification. G.1.3

PII. S104827595288085

1. Introduction.

There is much interest in the use of symplectic integrators for the numerical solution of Hamiltonian systems. Evidence suggests that these integrators are superior to traditional integrators for computations of low accuracy in which it is unnecessary to resolve all the details of the solution, for example, in molecular dynamics and galaxy simulations. For low accuracy computations, numerical methods of fourth-order accuracy are as a general rule the most efficient (although second order is often used because of its simplicity). A particularly efficient fourth-order symplectic integrator has been the method of Calvo and Sanz-Serna [3], which requires four force evaluations per step. However, a fourth-order method with fewer stages developed by Rowlands [14] has been shown recently [10] to be equally efficient. In this paper we explore the possible existence of even more efficient fourth-order explicit symplectic integrators, based on the idea in Rowlands's method (presented in section 2) of using Hessian–vector products and the concept of effective order.

*Received by the editors June 21, 1995; accepted for publication (in revised form) October 16, 1995.

http://www.siam.org/journals/sisc/18-1/28808.html

†Departamento de Matemática Aplicada y Computación, Facultad de Ciencias, Universidad de Valladolid, Valladolid, Spain (lopezmar@cpd.uva.es, sanzserna@cpd.uva.es). The research of these authors was supported in part by grant DGICYT PB92-254.

‡Department of Computer Science (and Beckman Institute University of Illinois at Urbana-Champaign), 1304 West Springfield Avenue, Urbana, IL 61801-2987 (skeel@cs.uiuc.edu). The research of this author was performed while at the University of Valladolid as Iberdrola Visiting Professor, with additional support from DOE/NSF grant DE-FG02-91-ER25099/DMS-9304268, NIH grant P41RR05969, and NSF/ARPA grant ASC-9318159.

Explicit symplectic methods of Runge–Kutta type of order greater than two all have the minor drawback that some of their weights must be negative [18] and, moreover, this must occur for both position and momentum variables [4]; these negative weights tend to be compensated by large positive weights, yielding error terms with somewhat larger coefficients. One possible way to circumvent this limitation is to incorporate higher derivatives of the Hamiltonian (for example, its Hessian) into the numerical method. Unfortunately, for general Hamiltonians and general separable Hamiltonians, it was proved by Murua [13] that explicit methods that exploit higher derivatives cannot be symplectic. This result does not apply, however, to the Hamiltonian

$$H(q,p) = T(p) + V(q), \quad T(p) = \frac{1}{2} p^T M^{-1} p, \tag{1}$$

for which the differential equations are

$$\frac{d}{dt} q = H_p(q,p) = M^{-1} p, \quad \frac{d}{dt} p = -H_q(q,p) = -V_q(q). \tag{2}$$

For example, the paper [20] mentions that Koseleff [7] gives a symplectic higher derivative method of order four that requires two modified force evaluations per step, in which a small correction involving a Hessian–vector product is added to the force $-V_q(q)$. The possibilities broaden further if we relax the condition of being fourth order to that of being *effectively* fourth order [2]. This means that we allow methods, perhaps of order smaller than four, that yield fourth-order accurate results after processing the numerical solution. A notable example is the aforementioned method of Rowlands, in which the force in the Störmer/leapfrog/Verlet method is modified by the addition of a Hessian–vector product to yield an effective order of four.

Section 3 shows that methods that use Hessian–vector products may be seen as the limiting case when abscissas of methods that evaluate the force at two or three distinct but closely spaced abscissas coalesce. Section 4 shows that for a sum of two-body interactions the incremental cost of evaluating a Hessian–vector product once the force has been evaluated is generally less than or equal to the cost of doing the original force evaluation. In our theoretical and experimental comparisons we therefore count a force evaluation together with a Hessian–vector computation as having a cost equivalent to two force evaluations. Section 6 explores a simple generalization of Rowlands's method. A new method is found that, according to a theoretical measure of the error, is significantly better than both the method of Rowlands and that of Calvo. The basic motivation behind the new method is quite simple: perform a Hessian–vector computation only every other time—significantly cutting costs while only marginally increasing the error. (It is important to note here that if the error is proportional to Eh^4 for some coefficient E, then to achieve a given accuracy on a unit interval the number of steps h^{-1} is proportional to $E^{1/4}$ and the work required is proportional to $mE^{1/4}$, where m is the work per step, say, in number of function evaluations.) Section 7 gives experimental results that compare the three methods and indicate the advantages of Hessian methods.

The new method has effective order four. This means that it is fourth-order accurate only if the numerical values produced by the formulas are appropriately interpreted. The initial values need to be preprocessed before application of the time-stepping scheme and the values thus obtained need to be postprocessed at desired output points. A discussion of processing and effective order is given in [10]. Not all errors produced by a given method may be removed by processing. For instance,

phase errors along a periodic orbit cannot be so removed, while errors in the shape of the numerical computed orbit can [16]. In practice, to obtain fourth-order accuracy, the postprocessing requires only the formation of second-order differences of numerical solution values [10, 17]. (The single preprocessing step may be done by extra force evaluations.) In applications for which only statistical or qualitative information is desired one may even dispense with the processing. In particular, in molecular dynamics and galaxy simulations initial values are known only partially, so preprocessing makes little sense. However, postprocessing may still make sense for some purposes, for example, monitoring fluctuations/drift in the energy. In such a case, to attach significance to changes in energy, the postprocessing transformation should be asymptotically equal to a symplectic one. In any case, the *main reason* for considering processing is that it reveals the true accuracy of a scheme (by retaining, for the most part, only error not removable by a change of variables) and thus directs us to schemes that are intrinsically more accurate. In sum, we have in mind that these methods, although derived on the basis that they will be processed, might generally be used with little or no processing.

The idea of effective order appeared first in [2] and later in the symplectic context in [14, 19]. For the symplectic case the ideas are explored to greater depth and more systematically in [20, 12], which use the term "corrector" for the preprocessing transformation. The latter of these papers uses an exact, computable representation of the corrector in the derivation of formulas. However, this is cumbersome for multiderivative methods and unnecessary if the details of processing are regarded to be of secondary importance. In section 5, we instead represent a sufficiently general symplectic transformation as the flow of an H-series, and we give a methodology for determining parameters for explicit symplectic generalized Runge–Kutta–Nyström (RKN) methods of increased effective accuracy.

2. Rowlands's method. As motivation, the method of Rowlands [14] is derived first. This method is also given by [20, Eq. (78)].

We begin with a method for (1) that is known by the names Störmer, leapfrog, and Verlet, and under the last of these names is currently the method of choice in molecular dynamics. In its "velocity form" [1] this method can be expressed as half a kick [20]

$$P^{n+1/2} = P^n - \frac{h}{2} V_q(Q^n)$$

followed by a drift

$$Q^{n+1} = Q^n + hM^{-1}P^{n+1/2}$$

and another half a kick

$$P^{n+1} = P^{n+1/2} - \frac{h}{2} V_q(Q^{n+1}).$$

More abstractly, we can write a Verlet step as

$$\phi_{hV/2} \circ \phi_{hT} \circ \phi_{hV/2}, \tag{3}$$

where ϕ_{hH} denotes the 1-flow of the Hamiltonian system with Hamiltonian hH, or equivalently the h-flow for Hamiltonian H.

Notation. We are using uppercase for raw "misinterpreted" numerical approximations and will use lowercase for processed numerical approximations.

The numerical solution is known to be "very nearly" the analytical solution of a nearby Hamiltonian system whose Hamiltonian is [16]

$$\tilde{H}(Q,P) = H + h^2 \frac{1}{12} T_p V_{qq} T_p - h^2 \frac{1}{24} V_q T_{pp} V_q + O(h^4), \tag{4}$$

with H as in (1). Here we use the convention (used for example by *Mathematica*) that when a vector, matrix, or higher-order object is followed by another such object, there is an implied summation over the last index of the first object and the first index of the second object. For several objects juxtaposed, a sensible grouping is implied.

One can interpret the first error term in (4) as a perturbation to the kinetic energy T and the second as a perturbation to the potential energy V. The perturbation to the kinetic energy changes the constant mass matrix M into a different and possibly position-dependent matrix, which seems unphysical. This has motivated [14, 16] a canonical change of variables that makes the perturbation more meaningful. By definition a canonical transformation yields a new Hamiltonian system whose Hamiltonian is simply the result of substituting the change of variables into the original Hamiltonian. The desired canonical change of variables $(Q,P) = \chi_h(q,p)$ should satisfy $\chi_h = \mathrm{id} + O(h^2)$, where id is the identity mapping $\mathrm{id}(q,p) = (q,p)$, so it is enough to consider the h-flow of a Hamiltonian. Moreover, for methods based on the composition of flows of T and V, it is enough [13] to consider Hamiltonians that are linear combinations of *elementary Hamiltonians* [5]. Only elementary Hamiltonians of order two can modify the dominant h^2 terms of the modified Hamiltonian and there is only one such elementary Hamiltonian of order two $T_p V_q$, so we try

$$H_\chi = h\lambda T_p V_q.$$

The change of variables is thus

$$\begin{aligned} Q &= q + h^2 \lambda T_{pp} V_q + O(h^4), \\ P &= p - h^2 \lambda V_{qq} T_p + O(h^4). \end{aligned} \tag{5}$$

The expressions $T_{pp}V_q$ and $V_{qq}T_p$ are the *elementary differentials* of order two. After substitution of (5) into (4) the perturbed Hamiltonian becomes

$$\tilde{H}(\chi(q,p)) = H + h^2 \left(-\lambda + \frac{1}{12}\right) T_p V_{qq} T_p + h^2 \left(\lambda - \frac{1}{24}\right) V_q T_{pp} V_q + O(h^4).$$

The choice $\lambda = \frac{1}{12}$ gives

$$\tilde{H}(\chi(q,p)) = H + \frac{h^2}{24} V_q T_{pp} V_q + O(h^4).$$

This new interpretation of the numerical solution of the Verlet method defines a "processed" Verlet method.

We can compensate for the error of the processed Verlet method by applying the Verlet time-stepping (3) to the Hamiltonian with potential energy

$$V_h := V - \frac{h^2}{24} V_q M^{-1} V_q$$

instead of V. The result is the Rowlands time-stepping

$$\phi_{hV_h/2} \circ \phi_{hT} \circ \phi_{hV_h/2}, \tag{6}$$

for which the error in the Hamiltonian of the transformed numerical variables becomes merely $O(h^4)$. Moreover, the h^2 discrepancy between the transformed and the

original variables is simply due to the transformation and therefore remains bounded if more timesteps are taken. We can actually perform the transformation on the initial conditions, apply (6), and, for output values only, do postprocessing by means of the reverse transformation. Then the method would be truly fourth order. The transformation need be performed only up to fourth-order accuracy and need not be symplectic since its effects are not propagated by the numerical integrator. To do this transformation computationally, approximate the preprocessing by

$$Q^0 = q^0 + \frac{h^2}{24} M^{-1}(V_q^{1/2} + V_q^{-1/2}), \quad P^0 = p^0 - \frac{h}{12}(V_q^{1/2} - V_q^{-1/2}),$$

where $V_q^{\pm 1/2} = V_q(q^0 \pm \frac{h}{2} M^{-1} p^0)$, and the postprocessing by formulas

$$q^n = Q^n + \frac{1}{12}(Q^{n+1} - 2Q^n + Q^{n-1}), \quad p^n = P^n - \frac{1}{12}(P^{n+1} - 2P^n + P^{n-1}),$$

whose derivation is given in [10]. However, as mentioned in the introduction, initial conditions are often unknown and there may be no need to do the transformations.

3. Hessian methods as a limiting case of gradient methods. In [11] it is shown that the Rowlands method is the limiting case of the 1-parameter family of the symmetric four-point first same as last (FSAL) effectively fourth-order RKN methods if we choose the parameter so that two points coalesce at each end. The accuracy decreases but not dramatically as the limit is approached, but the cost drops from three down to two function evaluations. Thus it would appear that we are getting nearly the accuracy of three function evaluations for the price of two. And, in fact, this is borne out by both the theoretical and experimental results in [11]. It is shown here how to obtain Hessian methods as limiting cases of RKN methods.

Let us use a superscript s, $s+$, or $s-$ to denote an approximation at time $t = sh$. Consider the following four intermediate substeps of a symplectic RKN method, where it is assumed that we have already obtained values P^{s-a} and Q^{s-}:

$$\begin{aligned} Q^{s-c\varepsilon} &= Q^{s-} + h(-c\varepsilon) M^{-1} P^{s-a}, \\ P^{s-1/\sqrt{\varepsilon}} &= P^{s-a} - h\left(a - \frac{1}{\sqrt{\varepsilon}}\right) V_q(Q^{s-c\varepsilon}), \\ Q^s &= Q^{s-c\varepsilon} + hc\varepsilon M^{-1} P^{s-1/\sqrt{\varepsilon}}, \\ P^{s+b} &= P^{s-1/\sqrt{\varepsilon}} - h\left(b + \frac{1}{\sqrt{\varepsilon}}\right) V_q(Q^s). \end{aligned} \tag{7}$$

Here a, b, c, and $\varepsilon > 0$ are method parameters. Thus these substeps advance from approximations Q^{s-}, P^{s-a} to approximations Q^s, P^{s+b}. The net advance in the P variable is $(a+b)h$ units of time. For the Q variable, the substeps produce a value Q^s which approximates the solution at the same value of time $t = sh$ as the initially given approximation Q^{s-}. As $\varepsilon \to 0$ in (7),

$$\begin{aligned} Q^s &= Q^{s-} + O(h^2\sqrt{\varepsilon}), \\ P^{s+b} &= P^{s-a} - h(a+b) V_q(Q^s) - h^3 c V_{qq}(Q^s) M^{-1} V_q(Q^s) + O(h^2\sqrt{\varepsilon}), \end{aligned}$$

which in the limit is a single substep that advances the P variable from P^{s-a} to P^{s+b} by using the potential gradient and potential Hessian evaluated at Q^s.

The foregoing construction cannot represent the midpoint of a family of symmetric methods; rather one has to form the composition of method fragment (7), where $b = 0$

together with its adjoint:

$$\begin{aligned}
Q^{s-c\varepsilon} &= Q^{s-} + h(-c\varepsilon)M^{-1}P^{s-a},\\
P^{s-1/\sqrt{\varepsilon}} &= P^{s-a} - h\left(a - \frac{1}{\sqrt{\varepsilon}}\right)V_q(Q^{s-c\varepsilon}),\\
Q^{s} &= Q^{s-c\varepsilon} + hc\varepsilon M^{-1}P^{s-1/\sqrt{\varepsilon}},\\
P^{s+1/\sqrt{\varepsilon}} &= P^{s-1/\sqrt{\varepsilon}} - h\frac{2}{\sqrt{\varepsilon}}V_q(Q^{s}),\\
Q^{s+c\varepsilon} &= Q^{s} + hc\varepsilon M^{-1}P^{s+1/\sqrt{\varepsilon}},\\
P^{s+a} &= P^{s+1/\sqrt{\varepsilon}} - h\left(a - \frac{1}{\sqrt{\varepsilon}}\right)V_q(Q^{s+c\varepsilon}),\\
Q^{s+} &= Q^{s+c\varepsilon} + h(-c\varepsilon)M^{-1}P^{s+a}.
\end{aligned}$$

As $\varepsilon \to 0$,

$$\begin{aligned}
Q^{s+} &= Q^{s-} + O(h^2\sqrt{\varepsilon}),\\
P^{s+a} &= P^{s-a} - 2haV_q(Q^s) - 2h^3cV_{qq}(Q^s)M^{-1}V_q(Q^s) + O(h^3\sqrt{\varepsilon} + h^2\varepsilon),
\end{aligned}$$

which in the limit is a single substep.

The paper [20] gives a fourth-order approximation to Rowlands's method that requires two force evaluations per step and replaces the Hessian–vector product by a difference of gradients, but this approximate method is not symplectic. By undoing the limiting process described in this section, Rowlands' method and other Hessian methods may be approximated by symplectic RKN methods.

4. Efficient implementation of the Hessian–vector product. Instead of the evaluation of V_q, the Rowlands method requires the more complicated evaluation

$$\nabla_q V_h := V_q - \frac{h^2}{12}V_{qq}M^{-1}V_q.$$

In the case of an N-body problem with 2-body interactions this can be done at the cost of at most two independent force evaluations and typically much less because of reuse of calculations such as square roots. This section (also [11]) describes how to organize the computation to achieve such an economy in computing time.

For the N-body problem the collective position vector q consists of 3-dimensional positions $r_1, r_2, \ldots, r_N$ of the bodies. For 2-body interactions the potential energy is a sum over all interacting pairs of atoms of the form

$$\frac{1}{2}\phi(\|r_j - r_i\|_2^2).$$

Likewise, the gradient $V_q(q)$ is a sum of terms like

$$\phi'(r_{ij}^2)\begin{bmatrix} 0\\ -r_{ij}\\ 0\\ r_{ij}\\ 0 \end{bmatrix},$$

where r_{ij} denotes $r_j - r_i$, and the 0's denote columns of zeros appropriate to leaving $-r_{ij}$ and r_{ij} in positions that correspond to the ith and jth forces, respectively.

Similarly, the Hessian V_{qq} is a sum of "element" matrices like

$$\begin{bmatrix} 0 \\ -I \\ 0 \\ I \\ 0 \end{bmatrix} \left(\phi'(r_{ij}^2)I + 2\phi''(r_{ij}^2) r_{ij} r_{ij}^{\mathrm{T}}\right) \begin{bmatrix} 0 & -I & 0 & I & 0 \end{bmatrix},$$

where I denotes the 3-by-3 identity matrix and the 0's denote blocks of zeros appropriate to leaving the identity matrices in positions that correspond to the ith and jth forces. The Hessian–vector product $V_{qq}d$, where d consists of 3-vectors $\delta_1, \delta_2, \ldots, \delta_N$, is computed as a sum of element matrices times vectors:

$$\begin{bmatrix} 0 \\ -I \\ 0 \\ I \\ 0 \end{bmatrix} \left(\phi'_{ij}\delta_{ij} + 2\phi''_{ij} r_{ij}^{\mathrm{T}} \delta_{ij} r_{ij}\right),$$

where δ_{ij} denotes $\delta_j - \delta_i$ and the scalars ϕ'_{ij} and $2\phi''_{ij}$ are to be computed and stored during the force evaluation.

As an example, an inverse square law has the form $\phi(s) = 2s^{-1/2}$ for which $\phi'(s) = -s^{-3/2}$ and $\phi''(s) = \frac{3}{2}s^{-5/2}$. Clearly the additional cost of obtaining ϕ'' is only a fraction of the cost of computing ϕ'. However, the formation of the Hessian–vector product requires two passes through the data, each pass involving a collection and a distribution of data. In the context of large-scale computing on a parallel computer, this doubles the communication costs, which are significant for distributed-memory machines. Hence, it is best to count the Hessian-vector product as equal in work to the evaluation of a second collective force vector.

The computation of the Hessian-vector product for an inverse-law force calculated by the fast multipole or related methods is the subject of current research [8].

5. Processing and effective order. In the derivation of symplectic methods it seems useful to consider symmetric (also called self-adjoint, time-reversible, or reflexive) methods. Suppose we are given a symplectic method that it not symmetric. The adjoint (or time reversal) of this method has equally good accuracy, and, for most families of methods including explicit RKN methods, equal computational cost. If we compose the original method with its adjoint, we obtain a new method with apparently equal computational costs and superior accuracy. The accuracy is superior because the asymptotic expansion of the error for the composed method is devoid of terms with odd powers of h.

For second- or fourth-order symmetric methods that use only elementary differentials, the modified Hamiltonian is an H-series [13] in h having as coefficients linear combinations of elementary Hamiltonians of the appropriate order. Hence the modified Hamiltonian must be of the form

$$\begin{aligned}\tilde{H}(Q,P) = H &+ h^2 A T_p V_{qq} T_p + h^2 B V_q T_{pp} V_q \\ &+ h^4 C T_p V_{qq} T_{pp} V_{qq} T_p + h^4 D V_q T_{pp} V_{qq} T_{pp} V_q \\ &+ h^4 F V_q T_{pp} V_{qqq} T_p T_p + h^4 G T_p V_{qqqq} T_p T_p T_p + O(h^6).\end{aligned}$$

We wish to perform a canonical transformation $(Q,P) = \chi_h(q,p)$ with $\chi_h = \mathrm{id} + O(h^2)$ that reduces the error terms of the modified Hamiltonian $\tilde{H}$. Any such

canonical transformation can be expressed as the h-flow of a Hamiltonian H_χ. A useful choice for H_χ will be an H-series consisting only of odd powers of h, so we write

$$H_\chi = h\lambda T_p V_q + h^3 \mu T_p V_{qq} T_{pp} V_q + \nu h^3 T_p V_{qqq} T_p T_p,$$

where the coefficients λ, μ, ν are to be determined.

For the analysis the explicit formation of the transformation $(Q, P) = \chi_h(q, p)$ can be avoided through the use of Lie series and the Baker–Campbell–Hausdorff (BCH) formula. Using Poisson brackets, defined for scalar functions $f(q,p)$, $g(q,p)$ by

$$\{f, g\} = f_q g_p - f_p g_q,$$

we can formally express the h-flow of the Hamiltonian system (2) by

$$\phi_{hH} = \exp(\{-hH, \cdot\})\mathrm{id}.$$

The processed method after n steps is the result of applying to the initial values the transformation

$$\phi_{h\tilde{H}}^{-1} \circ (\phi_{h\tilde{H}})^n \circ \phi_{hH_\chi} = (\phi_{hH_\chi}^{-1} \circ \phi_{h\tilde{H}} \circ \phi_{hH_\chi})^n,$$

for which

$$\begin{aligned}\phi_{h\tilde{H}}^{-1} \circ \phi_{h\tilde{H}} \circ \phi_{hH_\chi} &= \exp(-\{-hH_\chi, \cdot\})\mathrm{id} \circ \exp(\{-h\tilde{H}, \cdot\})\mathrm{id} \circ \exp(\{-hH_\chi, \cdot\})\mathrm{id} \\ &= \exp(\{-hH_\chi, \cdot\}) \exp(\{-h\tilde{H}, \cdot\}) \exp(-\{-hH_\chi, \cdot\})\mathrm{id}.\end{aligned}$$

The application of the BCH formula to the special product $\mathrm{e}^C \mathrm{e}^K \mathrm{e}^{-C}$, where C and K are Lie operators, simplifies considerably, yielding the formula [20]

$$\mathrm{e}^C \mathrm{e}^K \mathrm{e}^{-C} = \exp\left(K + [C, K] + \frac{1}{2}[C, [C, K]] + \cdots\right) = \exp(\mathrm{e}^{[C,\cdot]} K),$$

where the (square) brackets denote the commutator $[A, B] = AB - BA$. Employing the homomorphism from functions with Poisson brackets to Lie operators with commutators gives

$$\exp(\{-hH_\chi, \cdot\}) \exp(\{-h\tilde{H}, \cdot\}) \exp(-\{-hH_\chi, \cdot\}) = \exp(\{-h\tilde{H} \circ \chi_h, \cdot\}),$$

where

$$\begin{aligned}-h\tilde{H} \circ \chi_h &= -h\tilde{H} + \{-hH_\chi, -h\tilde{H}\} + \frac{1}{2}\{-hH_\chi, \{-hH_\chi, -h\tilde{H}\}\} + \cdots \\ &= \exp(\{-hH_\chi, \cdot\})(-h\tilde{H}),\end{aligned}$$

and we have

$$\begin{aligned}\tilde{H}(\chi(q,p)) &= \tilde{H} - h\{H_\chi, \tilde{H}\} + \frac{h^2}{2}\{H_\chi, \{H_\chi, \tilde{H}\}\} + O(h^6) \\ &= \cdots \\ &= H(q,p) + h^2(A - \lambda) T_p V_{qq} T_p + h^2(B + \lambda) V_q T_{pp} V_q \\ &\quad + h^4(C + \lambda^2 - 2A\lambda - \mu) T_p V_{qq} T_{pp} V_{qq} T_p \\ &\quad + h^4(D + \lambda^2 + 2B\lambda + \mu) V_q T_{pp} V_{qq} T_{pp} V_q \\ &\quad + h^4\left(F - \mu + 3\nu + A\lambda - \frac{\lambda^2}{2}\right) V_q T_{pp} V_{qqq} T_p T_p + h^4(G - \nu) T_p V_{qqqq} T_p T_p T_p \\ &\quad + O(h^6).\end{aligned} \tag{8}$$

The parameter λ can be chosen to obtain fourth-order accuracy in (8) if and only if

$$B = -A, \tag{9}$$

in which case

$$\lambda = A \tag{10}$$

is the appropriate choice. Henceforth, let us assume that both (9) and (10) are satisfied, in which case we have

$$\begin{aligned}\tilde{H}(\chi(q,p)) = H(q,p)&\\ &+h^4(C - A^2 - \mu)T_pV_{qq}T_{pp}V_{qq}T_p + h^4(D - A^2 + \mu)V_qT_{pp}V_{qq}T_{pp}V_q\\ &+h^4\left(F - \mu + 3\nu + \frac{1}{2}A^2\right)V_qT_{pp}V_{qqq}T_pT_p\\ &+h^4(G - \nu)T_pV_{qqqq}T_pT_pT_p + O(h^6).\end{aligned}$$

The choice of μ and ν is not so obvious. As a heuristic it is common to choose parameters to minimize the sum of squares of the error coefficients. Minimizing

$$E^2 = (C - A^2 - \mu)^2 + (D - A^2 + \mu)^2 + \left(F - \mu + 3\nu + \frac{1}{2}A^2\right)^2 + (G - \nu)^2$$

gives

$$\begin{aligned}\mu &= \frac{1}{42}A^2 + \frac{10}{21}C - \frac{10}{21}D + \frac{1}{21}F + \frac{1}{7}G,\\ \nu &= -\frac{1}{7}A^2 + \frac{1}{7}C - \frac{1}{7}D - \frac{2}{7}F + \frac{1}{7}G,\end{aligned}$$

$$\begin{bmatrix}\hat{C}\\ \hat{D}\\ \hat{F}\\ \hat{G}\end{bmatrix} = \begin{bmatrix}\frac{11}{21} & \frac{10}{21} & -\frac{1}{21} & -\frac{1}{7}\\ \frac{10}{21} & \frac{11}{21} & \frac{1}{21} & \frac{1}{7}\\ -\frac{1}{21} & \frac{1}{21} & \frac{2}{21} & \frac{2}{7}\\ -\frac{1}{7} & \frac{1}{7} & \frac{2}{7} & \frac{6}{7}\end{bmatrix}\begin{bmatrix}C\\ D\\ F\\ G\end{bmatrix} + \begin{bmatrix}-\frac{43}{42}\\ -\frac{41}{42}\\ \frac{1}{21}\\ \frac{1}{7}\end{bmatrix}A^2.$$

As an example let us look at the error $\tilde{H} - H$ of the symmetrized Calvo method, which is defined to be the original Calvo method [15, p. 113] followed by its adjoint. The symmetrized method is a nine-stage FSAL method and for one step of stepsize $2h$ it is given by

$$\begin{aligned}&\phi_{b_1hV} \circ \phi_{c_1hT} \circ \phi_{b_2hV} \circ \phi_{c_2hT} \circ \phi_{b_3hV} \circ \phi_{c_3hT} \circ \phi_{b_4hV} \circ \phi_{c_4hT} \circ \phi_{2b_5hV}\\ &\circ\phi_{c_4hT} \circ \phi_{b_4hV} \circ \phi_{c_3hT} \circ \phi_{b_3hV} \circ \phi_{c_2hT} \circ \phi_{b_2hV} \circ \phi_{c_1hT} \circ \phi_{b_1hV},\end{aligned}$$

where

$$\begin{aligned}b_1 &= \ \ 0.0617588581356263250,\\ b_2 &= \ \ 0.3389780265536433551,\\ b_3 &= \ \ 0.6147913071755775662,\\ b_4 &= -0.1405480146593733802,\\ b_5 &= \ \ 0.1250198227945261338,\\ c_1 &= \ \ 0.2051776615422863869,\\ c_2 &= \ \ 0.4030212816042145870,\\ c_3 &= -0.1209208763389140082,\\ c_4 &= \ \ 0.5127219331924130343.\end{aligned}$$

To determine the modified Hamiltonian, it is convenient to develop a formula of the form

$$\exp(\{-hf,\cdot\})\exp(\{-hg,\cdot\})\exp(\{-hf,\cdot\}) = \exp(\{-hF,\cdot\})$$

using the BCH-derived formula in [15, eq. (12.12)]. If we exploit the homomorphism from functions and Poisson brackets to Lie operators and commutators and simplify a little, we get

$$F = 2f + g + h^2\left(\frac{1}{6}\{g,g,f\} - \frac{1}{6}\{f,f,g\}\right) + h^4\left(\frac{7}{360}\{f,f,f,f,g\} - \frac{1}{360}\{g,g,g,g,f\}\right.$$
$$\left. + \frac{1}{90}\{f,g,g,g,f\} + \frac{1}{45}\{g,f,f,f,g\} - \frac{1}{60}\{f,f,g,g,f\} + \frac{1}{30}\{g,g,f,f,g\}\right) + O(h^6),$$

where the Poisson bracket is recursively extended to multiple arguments by the rule

$$\{f_1, f_2, \ldots, f_m\} = \{f_1, \{f_2, \ldots, f_m\}\}.$$

With the use of these rules in *Mathematica*, the perturbed Hamiltonian of the symmetrized Calvo method with *stepsize* $8h$ is found to be (coefficients have been truncated)

$$\tilde{H}(Q,P) = T + V + 0.034\,h^4 T_p V_{qq} T_{pp} V_{qq} T_p - 0.053\,h^4 V_q T_{pp} V_{qq} T_{pp} V_q$$
$$- 0.075\,h^4 V_q T_{pp} V_{qqq} T_p T_p + 0.019\,h^4 T_p V_{qqqq} T_p T_p T_p + O(h^6).$$

A stepsize of $8h$ is chosen because each step requires the evaluation of eight forces. The square root of the sum of the squares of error coefficients for stepsize chosen to be $8h$ is 0.0995. Its perturbed Hamiltonian after processing is (again after truncation of the coefficients)

$$\tilde{H}(Q,P) = T + V - 0.007\,h^4 T_p V_{qq} T_{pp} V_{qq} T_p - 0.012\,h^4 V_q T_{pp} V_{qq} T_{pp} V_q$$
$$- 0.006\,h^4 V_q T_{pp} V_{qqq} T_p T_p - 0.017\,h^4 T_p V_{qqqq} T_p T_p T_p + O(h^6).$$

The square root of the sum of the squares of error coefficients for stepsize chosen to be $8h$ is 0.0228. If we take the fourth root of the ratio of this to its previous value, we conclude that the computational effort to achieve a given accuracy is reduced by a factor of 0.69 if we process Calvo's method. This supposes that practical postprocessors can be developed, say, using linear combinations of force values available from intermediate stages of the method.

6. A family of 3-point FSAL methods. Inspired by Rowlands's method, we examine a family of 3-point FSAL Hessian methods

$$\phi_{(1/4+b)hV+dh^3 V_q T_{pp} V_q} \circ \phi_{(1/2)hT} \circ \phi_{(1/2-2b)hV+ch^3 V_q T_{pp} V_q} \circ \phi_{(1/2)hT}$$
$$\circ \phi_{(1/4+b)hV+dh^3 V_q T_{pp} V_q},$$

whose effective cost per step is four function evaluations. The case $b = 0$, $c = -1/192$, $d = -1/384$ is the 2-fold Rowlands method (two concatenated steps of the method with stepsize $h/2$).

The perturbed Hamiltonian of this 3-parameter method after processing has error coefficients

$$\hat{A} = \frac{1}{48} + \frac{b}{4} - \lambda,$$
$$\hat{B} = -\frac{1}{96} - \frac{b}{4} + \frac{b^2}{2} + c + 2d + \lambda,$$
$$\hat{C} = \frac{1}{960} + \frac{b}{24} + \frac{b^2}{12} - \frac{c}{12} + \frac{d}{3} - \frac{\lambda}{24} - \frac{b\lambda}{2} + \lambda^2 - \mu,$$
$$\hat{D} = -\frac{1}{3840} - \frac{5b}{192} + \frac{5b^2}{48} - \frac{b^3}{4} + \frac{c}{12} - bc - \frac{d}{3} - \frac{\lambda}{48} - \frac{b\lambda}{2} + b^2\lambda + 2c\lambda + 4d\lambda + \lambda^2 + \mu,$$
$$\hat{F} = \frac{1}{1920} + \frac{b}{24} - \frac{b^2}{24} - \frac{c}{12} + \frac{d}{3} + \frac{\lambda}{48} + \frac{b\lambda}{4} - \frac{\lambda^2}{2} - \mu + 3\nu,$$
$$\hat{G} = -\frac{1}{11520} - \frac{b}{192} - \nu.$$

With

$$c = -\frac{1}{96} - \frac{b^2}{2} - 2d,$$
$$\lambda = \frac{1}{48} + \frac{b}{4},$$

chosen to make $\hat{A} = 0$ and $\hat{B} = 0$, the other error coefficients become

$$\hat{C} = \frac{17}{11520} + \frac{b}{32} + \frac{b^2}{16} + \frac{d}{2} - \mu,$$
$$\hat{D} = -\frac{1}{640} - \frac{5b}{192} + \frac{b^3}{4} - \frac{d}{2} + 2bd + \mu,$$
$$\hat{F} = \frac{37}{23040} + \frac{3b}{64} + \frac{b^2}{32} + \frac{d}{2} - \mu + 3\nu,$$
$$\hat{G} = -\frac{1}{11520} - \frac{b}{192} - \nu.$$

We further assume that μ and ν are chosen to minimize the sum of the squares of the error coefficients $E^2 = \hat{C}^2 + \hat{D}^2 + \hat{F}^2 + \hat{G}^2$.

The two-fold Rowlands method is obtained by choosing $b = 0$ and $d = \frac{1}{2}c = -\frac{1}{384}$. The processed coefficients are

$$\hat{C} = -0.009, \quad \hat{D} = -0.013, \quad \hat{F} = -0.004, \quad \hat{G} = -0.013.$$

An appropriate measure of error for a fourth-order method is $E^{1/4} = 0.0950$. To take into account the cost of one step we multiply by the number of force evaluations $m = 4$, obtaining 0.380. These values together with those for the unprocessed and processed symmetrized Calvo methods and other methods yet to be discussed are shown in Table 1. In all cases these values are for the processed versions of the methods where λ is chosen to yield fourth-order accuracy and μ and ν are chosen to minimize E. We note here a slight theoretical superiority of the Rowlands method over that of the processed symmetrized Calvo.

Consider now the selection of b and d to minimize E. Keeping b free and choosing d to minimize E gives

$$d = \frac{5}{101376}b^{-1} - \frac{1}{384} - \frac{21}{704}b - \frac{1}{8}b^2$$

TABLE 1
Theoretical comparison of methods.

Method	Error $E^{1/4}$	Scaled error $mE^{1/4}$
unproc. symm. Calvo	0.0702	**0.562**
proc. symm. Calvo	0.0486	**0.389**
Rowlands	0.0950	**0.380**
best b, best d	0.0792	0.317
best b, $d = 0$	0.0803	**0.241**
$b = 0$, $d = 0$	0.0950	0.285

and

$$E = \frac{1}{648806400} + \frac{1}{1351680}b^2 + \frac{1}{11264}b^4.$$

The error measure E approaches a lower bound as $b \to 0$, although this causes the method coefficients c and d to approach plus and minus infinity. The value of the error and work measures for the limiting case are given in Table 1 as "best b, best d."

The method just obtained shows the limitations of our error measure E. This difficulty is less likely to occur if we work with fewer degrees of freedom. An effective way to do this is to use one degree of freedom to reduce the amount of work per step. The obvious choice is $d = 0$ (or $c = 0$, which is essentially equivalent). Then choosing b to minimize E requires solving

$$\frac{11}{56}b^5 + \frac{5}{64}b^4 + \frac{107}{8064}b^3 + \frac{29}{32256}b^2 + \frac{17}{967680}b - \frac{5}{9289728} = 0.$$

There is one real root,

$$b = 0.015425721644647824439.$$

The other coefficients are

$$\begin{aligned} c &= -\frac{1}{96} - \frac{1}{2}b^2, \\ \lambda &= \frac{1}{48} + \frac{1}{4}b, \\ \mu &= \frac{731}{483840} + \frac{29}{1008}b + \frac{1}{32}b^2 - \frac{5}{42}b^3, \\ \nu &= -\frac{1}{26880} - \frac{1}{168}b - \frac{1}{28}b^3. \end{aligned}$$

The processed coefficients are

$$\hat{C} = 0.0008, \quad \hat{D} = -0.0001, \quad \hat{F} = -0.0010, \quad \hat{G} = -0.0031.$$

Table 1 confirms that E for this method is only a little more than before and that the work is dramatically less.

If we like our coefficients to be simple rational numbers, then we might be interested in the choice $b = 0$ (and $c = -\frac{1}{96}$). Table 1 shows that this method is not far from optimal.

This family of methods was defined at the beginning of this section in terms of flows of H-series. It might seem useful to incorporate other elementary Hamiltonians into a method. There is one other elementary Hamiltonian of order three $V_q T_{pp} V_q$, but it would not lead to an explicit method. Among elementary Hamiltonians of order five, only $V_q T_{pp} V_{qq} T_{pp} V_q$ would be useful.

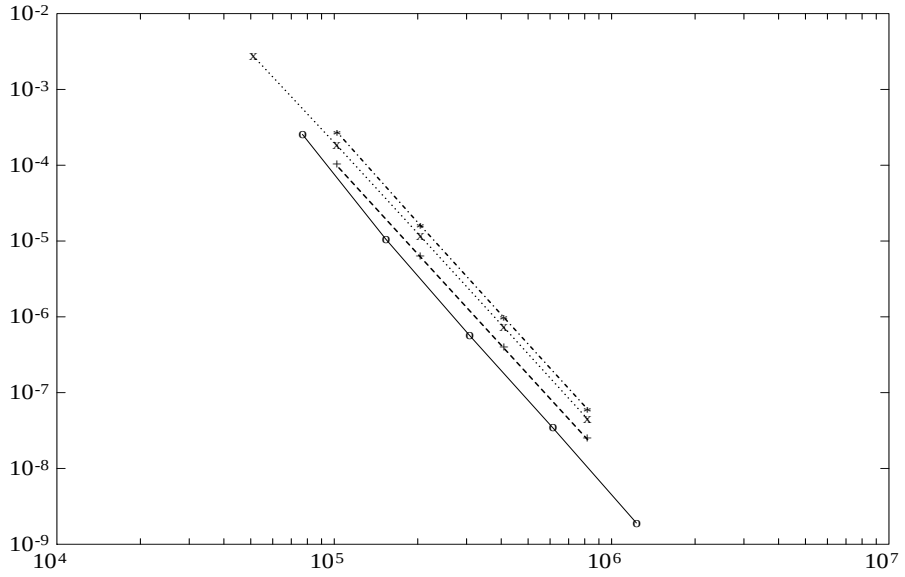

FIG. 1. *Average error norm over one period vs. number of force evaluations for unprocessed Calvo (dashed line and +), processed Calvo (dash-dot line and *), Rowlands (dotted line and ×), and new (solid line and circles) methods.*

7. Numerical experiments. We performed experiments to compare the symmetrized method of Calvo, unprocessed and processed, the method of Rowlands, and the new method (with $d = 0$ and optimal b). In order to compare accuracy we developed a "black box" for pre- and post-processing. Little attention was given to the efficiency of this black box. It obtains asymptotically correct values for elementary differentials required in the processors (see, e.g., (5)) by differencing numerical values obtained by taking a couple of steps of Verlet and Euler in the forward and backward direction.

Two test problems were used. The first is a Kepler problem in two dimensions, $H(q,p) = \frac{1}{2}(p_1^2 + p_2^2) - (q_1^2 + q_2^2)^{-1/2}$, with initial conditions $q_1 = \frac{1}{2}$, $q_2 = 0$, $p_1 = 0$, $p_2 = \sqrt{3}$ chosen to yield an orbit of eccentricity $\frac{1}{2}$ and period 2π [15, section 1.2.4]. The problem was run for 100 periods and the Euclidean norm in $\mathcal{R}^4$ of the error was measured during the last revolution at times $(99 + j/8) \times 2\pi$, $j = 1, \ldots, 8$ and then averaged. The stepsizes were $h = 2\pi/128$, $2\pi/256$, $2\pi/512$, and $2\pi/1024$ for the Calvo method and $2\pi/256$, ..., $2\pi/4096$ for the other two methods. Figure 1 shows the error versus the number of force + Hessian evaluations for each of the methods. We observe that the ranking of the four methods agrees with their theoretical work measure given by Table 1 except that the unprocessed Calvo performs significantly better than predicted and, in particular, better than the processed version. This anomaly could be explained by a fortuitous cancellation of error terms for this particular problem, and indeed we have checked that for other problems the performance of the unprocessed Calvo method is worse than that of the processed Calvo method. Note in this connection that, even though the processed Calvo method has a value of E significantly lower than the unprocessed version, the coefficient in $\tilde{H}$ of the elementary Hamiltonian involving V_{qqqq} does not decrease by processing, so that the processing of the Calvo method cannot be expected to be uniformly beneficial

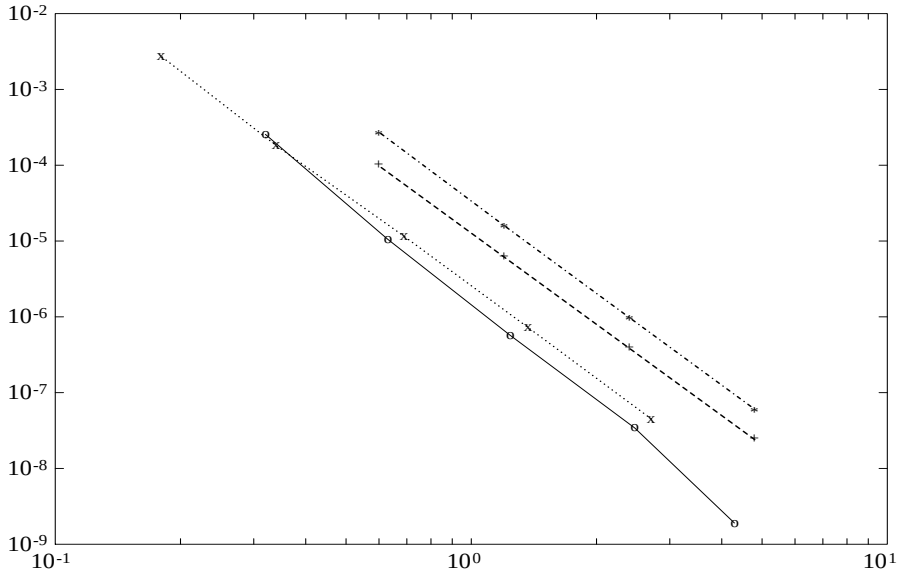

FIG. 2. *Average error norm over one period vs. CPU time for unprocessed Calvo (dashed line and +), processed Calvo (dash-dot line and *), Rowlands (dotted line and ×), and new (solid line and circles) methods.*

across all problems. On the other hand, for the Kepler problem the incremental cost of evaluating the Hessian (which in Figure 1 was assumed to be equal to that of the force evaluation) is expected to be small because of the reuse of the relatively costly square root. This expectation is confirmed by measurements of CPU time shown in Figure 2, which shows a sizable advantage for the Hessian methods.

The second test problem is a simulation of the $N = 256$ atoms of liquid argon as in [6, 9]. The potential energy is a sum of terms over all pairs of atoms of the form

$$\frac{1}{2}\phi(\|r_j - r_i\|_2^2),$$

where $r_1, r_2, \ldots, r_N$ are atomic positions and

$$\phi(s) = 8\epsilon\left(\frac{\sigma^{12}}{s^6} - \frac{\sigma^6}{s^3}\right).$$

The values used are $\epsilon = 1.65324 \times 10^{-21}$ joules, $\sigma = 3.405$ Å, and mass $m = 6.64 \times 10^{-26}$ kilograms. Periodicity of period $L = 22.984$ Å is imposed in each of the three directions, effectively creating for each atom a 3-dimensional lattice of images. This is efficiently implemented with very little error by neglecting all pairwise interactions between atoms (or images) whose x-, y-, or z-separation equals or exceeds $L/2$. Initial conditions were chosen as in [9] to achieve a temperature of 86.5°K. The computations were programmed in nondimensional units where a typical stepsize $h = 0.1$ corresponds to 0.0311 picoseconds of physical time. Simulations were run for time 60 (18.68 picoseconds) and sampled at intervals of 1.2. As our measure of accuracy we choose the root-mean-square deviation of the energy from its initial value. For symplectic integrators the energy deviation represents the value of the perturbation to the Hamiltonian and thus is a meaningful measure of accuracy. The results are

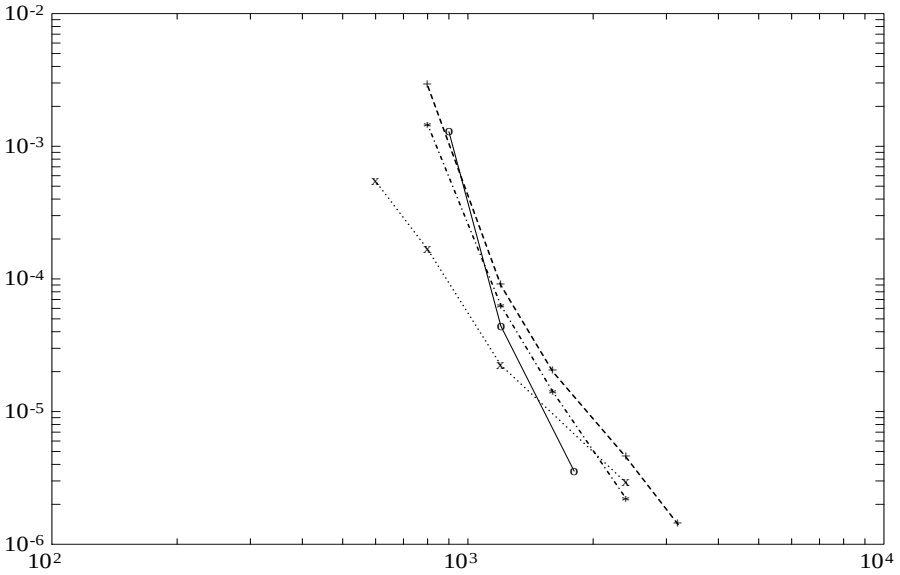

FIG. 3. *Relative energy error vs. number of function evaluations for unprocessed Calvo (dashed line and +), processed Calvo (dash-dot line and *), Rowlands (dotted line and ×), and new (solid line and circles) methods.*

shown in Figure 3. For small stepsize the results are fairly consistent with the theoretical ranking of methods; for larger stepsizes Rowlands's method does remarkably well.

REFERENCES

[1] M. P. ALLEN AND D. J. TILDESLEY, *Computer Simulation of Liquids*, Clarendon Press, Oxford, 1987.

[2] J. C. BUTCHER, *The effective order of Runge-Kutta methods*, in Conference on the Numerical Solution of Differential Equations, J. Ll. Morris, ed., Lecture Notes in Mathematics 109, Springer-Verlag, New York, 1969, pp. 133–139.

[3] M. P. CALVO AND J. M. SANZ-SERNA, *The development of variable-step symplectic integrators, with applications to the two-body problem*, SIAM J. Sci. Statist. Comp., 14 (1993), pp. 936–952.

[4] D. GOLDMAN AND T. J. KAPER, *Nth-order operator splitting schemes and nonreversible systems*, SIAM J. Numer. Anal., 33 (1996), pp. 349–367.

[5] E. HAIRER, *Backward error analysis of numerical integrators and symplectic methods*, Ann. Numer. Math., 1 (1994), pp. 103–132.

[6] D. JANEŽIČ AND B. OREL, *Implicit Runge-Kutta method for molecular dynamics integration*, J. Chem. Inf. Comput. Sci., 33 (1993), pp. 252–257.

[7] P.-V. KOSELEFF, in *Applied Algebra, Algebraic Algorithms, and Error-correcting Codes*, G. Cohen, T. Mora, and O. Moreno, eds., AAECC-10, Springer-Verlag, New York, 1993.

[8] CH. G. LAMBERT, *Efficient dense Hessian computations in molecular minimizations*, Presentation at Multigrid Tutorial, with Applications to Molecular Dynamics, October 10–12, 1995, Weizmann Institute, Rehovot, Israel.

[9] M. LÓPEZ-MARCOS, J. M. SANZ-SERNA, AND J. C. DÍAZ, *Are Gauss-Legendre methods useful in molecular dynamics?*, J. Comput. Appl. Math., 67 (1996), pp. 173–179.

[10] 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, Harlow (Essex), 1996, pp. 107–122.

[11] M. LÓPEZ-MARCOS, J. M. SANZ-SERNA, AND R. D. SKEEL, *Explicit symplectic integrators with maximal stability intervals*, in Numerical Analysis, D. F. Griffiths and G. A. Watson, eds., World Scientific, Singapore, 1996, pp. 163–175.

[12] R. I. McLachlan, *More on symplectic correctors*, in Integration Algorithms and Classical Mechanics, J. E. Marsden, G. W. Patrick, and W. F. Shadwick, eds., American Mathematical Society, Providence, RI, 1996.
[13] A. Murua, *Métodos Simplécticos Desarrollables en P-series*, Ph.D. thesis, Universidad de Valladolid, 1995.
[14] G. Rowlands, *A numerical algorithm for Hamiltonian systems*, J. Comput. Phys., 97 (1991), pp. 235–239.
[15] J. M. Sanz-Serna and M. P. Calvo, *Numerical Hamiltonian Problems*, Chapman and Hall, London, 1994.
[16] R. D. Skeel, T. R. Littell, and M. Zhang, *Error analysis of symplectic multiple time stepping*, SIAM J. Numer. Anal., 34 (1997), to appear.
[17] R. D. Skeel, G. Zhang, and T. Schlick, *A family of symplectic integrators: Stability, accuracy and molecular dynamics applications*, SIAM J. Sci. Comput., 18 (1997), pp. 203–222.
[18] M. Suzuki, *General theory of fractal path integrals with applications to many-body theories and statistical physics*, J. Math. Phys., 32 (1991), pp. 400–407.
[19] M. Suzuki, *Improved Trotter-like formula*, Physics Letters A, 180 (1993), pp. 232–234.
[20] 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., American Mathematical Society, Providence, RI, 1996.