with μ and σ given in the above table. Since $2\mu \neq p$ in general, y' will appear under the integral sign in the exact difference formula.

4. References

1. M. Abramowitz and I. A. Stegun (eds). *Handbook of Mathematical Functions.* National Bureau of Standards, Washington, D.C., 1964.
2. Uri M. Ascher, Robert M. M. Mattheij, and Robert D. Russell. *Numerical Solution of Boundary Value Problems.* SIAM, Philadelphia, 1995.
3. J. C. Butcher, J. R. Cash, G. Moore, and R. D. Russell. Defect correction for two-point boundary value problems on non-equidistant meshes. *Mathematics of Computation*, 64:629–648, 1995.
4. I. Christie and A. R. Mitchell. Upwinding of high order Galerkin methods in conduction-convection problems. *Int. J. Num. Meth. Engng.*, 12:1764–1771, 1978.
5. L. Collatz. *Numerical Treatment of Differential Equations.* Springer, Berlin, 1960.
6. R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume 1. Interscience, New York, 1953.
7. Philip J. Davis. *Interpolation and Approximation.* Dover, New York, 1975.
8. Carl de Boor. *A Practical Guide to Splines.* Springer, New York, 1978.
9. A. Ghizzetti and A. Ossicini. *Quadrature Formulas.* Academic Press, New York, 1970.

In "Numerical Analysis, A.R. Mitchell 75th Birthday Volume", D.F. Griffiths & G.A. Watson editors, World Scientific, Singapore, 1996, pages163-175.

AN EXPLICIT SYMPLECTIC INTEGRATOR WITH MAXIMAL STABILITY INTERVAL

M. A. LOPEZ-MARCOS, J. M. SANZ-SERNA
Departamento de Matemática Aplicada y Computación, Facultad de Ciencias, Universidad de Valladolid, Valladolid, Spain
E-mail: lopezmar@cpd.uva.es, sanzserna@cpd.uva.es

and

ROBERT D. SKEEL
Dept. of Computer Science and Beckman Inst., Univ. of Illinois at Urbana-Champaign 1304 West Springfield Avenue, Urbana, Illinois 61801-2987, USA
E-mail: skeel@cs.uiuc.edu

ABSTRACT

We derive and test a Runge-Kutta-Nyström method that possesses a stability interval that is maximal over all methods that are explicit, symplectic, effectively fourth order and use three force evaluations per step. Effective order four means that the output of the given method is going to be processed so as to enhance its accuracy and that the numerical solution after processing possesses $O(h^4)$ error bounds. By comparing the new method with a similar method of conventional order four, we show that, when designing schemes, the use of the notion of effective order leads to more efficient integrators than the notion of conventional order. The new method is less efficient than a related method, introduced by Rowlands, that uses the Hessian of the potential. This shows the interest of investigating further methods that use the Hessian of the potential.

1. Introduction

In this paper we derive and test a Runge-Kutta-Nyström (RKN) method that possesses a stability interval that is maximal over all methods that are explicit, symplectic, effectively fourth order and use three force evaluations per step. While the recent literature on symplectic integration of Hamiltonian systems is very large, see e.g. Hairer, Nørsett and Wanner[6], Sanz-Serna and Calvo[11], the stability intervals of symplectic methods have received little attention. This is surprising since symplectic integrators are primarily useful in situations where highly accurate solutions are not required and the interest lies in obtaining statistical or qualitative properties. In these situations the fastest components typically have the smallest amplitudes and need not be resolved accurately. Then stability is likely to be the factor limiting the stepsize. There is a second reason why stability intervals may be of interest: a large constant in any of the terms of the asymptotic expansion of the global error in powers of the stepsize h is likely to lead to a small stability interval. Therefore by controlling the stability interval it may be possible to obtain methods with small error coefficients at all powers of h.

Molecular dynamics simulations (Allen and Tildesley[1]) are often carried out with the simple second order, symplectic, time-reversible (selfadjoint in the terminology of Hairer, Nørsett and Wanner[6] or Sanz-Serna and Calvo[11]) Verlet method, which possesses a *scaled* stability interval of length 2. This length is known to be (Chawla and Sharma[4]) the largest possible for any symplectic, explicit method (regardless of the order of accuracy). Recall that a numerical method with stability interval of length L has a stability stepsize restriction $h \leq L$ when applied to the model problem $d^2y/dt^2 = -y$. If the method uses m force evaluations per step, then the scaled length L/m measures how many units of time per force evaluation may be advanced with the integrator within a stable simulation. Therefore, the Verlet method may well be the best choice if one is interested in explicit, symplectic integrators of order two.

There is not much point in using symplectic methods of order three: by concatenating an order three method and its adjoint one obtains a method of order four (Sanz-Serna and Calvo[11], Section 8.4.5), so that one may directly consider methods of order four. Here we construct a Runge-Kutta-Nyström (RKN) method that possesses a stability interval that is maximal over all methods (RKN or not) that are explicit, symplectic, effectively fourth order and use three force evaluations per step. The terminology effective order r is taken from Butcher[2]. It implies that the output of the given method is going to be processed so as to enhance its accuracy and that the numerical solution *after processing* possesses $O(h^r)$ global errors. López-Marcos, Sanz-Serna and Skeel[8] have showed how to process order four, symplectic, time-reversible symplectic methods at virtually no cost. Therefore when designing our method, it is meaningful to look at its effective order rather than at its conventional order.

In Section 2 we review the idea of processing. In Section 3 we study stability intervals and find the maximal stability interval for explicit, symplectic methods of effective order four that use three force evaluations per step. We present an interpretation of processing in terms of the eigenvectors and eigenvalues of the amplification matrix. In Section 4 a method is constructed that realizes the maximal stability interval. In Section 5 we study analytically the accuracy of the new method. By comparing the new method with a similar method of conventional order four, we show that, when designing schemes, the use of the notion of effective order leads to more efficient integrators than the notion of conventional order. By imposing the conditions for conventional order many degrees of freedom in the method are wasted. However it turns out that the new method is less efficient than a related method of effective order four introduced by Rowlands that uses the Hessian of the potential. It is shown in López-Marcos, Sanz-Serna and Skeel[8] that, in many problems, the cost of evaluating at a given point the potential *and* the Hessian of the potential is less than that of two evaluations of the potential. The cost per step of the Rowlands method (one evaluation of the gradient and Hessian) is significantly lower than that of the method constructed here (three evaluations of the gradient). Then, even though the new method is more accurate per step than the Rowlands method, it is less accurate per unit of cost. This may well show the interest in investigating further methods that use the Hessian of the potential. Numerical results are presented in the final Section 6.

2. Processing

Throughout the paper we consider Hamiltonian functions of the form

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

where the potential V is a smooth function and M is a constant, invertible, symmetric matrix. The Hamiltonian system corresponding to (1) is given by

$$\frac{d\mathrm{q}}{dt} = M^{-1}\mathrm{p}, \qquad \frac{d\mathrm{p}}{dt} = -V_{\mathrm{q}}(\mathrm{q}). \tag{2}$$

The notation V_{q} means the gradient of V with respect to q. The negative of this gradient is the force.

A one-step method for the integration of (2) is given by a transformation $\psi_{h,H}$ that maps the approximation $(\mathrm{q}_n, \mathrm{p}_n)$ corresponding to a time level $t_n = nh$ into the approximation $(\mathrm{q}_{n+1}, \mathrm{p}_{n+1}) = \psi_{h,H}(\mathrm{q}_n, \mathrm{p}_n)$ at the next time level t_{n+1}. Let us assume that we have been given a method $\psi_{h,H}$, that in what follows is called the basic method. When processing is used, there are two sets of variables being considered. The first set, that we denote by capital letters (Q,P), corresponds to the values computed by the basic method; specifically, we compute the sequence $(\mathrm{Q}_{n+1}, \mathrm{P}_{n+1}) = \psi_{h,H}(\mathrm{Q}_n, \mathrm{P}_n)$, $n = 0, 1, \ldots$, starting from $(\mathrm{Q}_0, \mathrm{P}_0)$. The second set of variables (q,p) is related to the first through a transformation $(\mathrm{Q},\mathrm{P}) = \chi_{h,H}(\mathrm{q},\mathrm{p})$. It is the lower case variables that are seen as the processed numerical approximations to the solutions of (2). Thus there are three steps involved in the processed algorithm:

1. *Preprocessing:* Find, from the initial values $(\mathrm{q}(0), \mathrm{p}(0))$ the starting values for time-stepping $(\mathrm{Q}_0, \mathrm{P}_0) = \chi_{h,H}(\mathrm{q}(0), \mathrm{p}(0))$.

2. *Time-stepping:* Compute $(\mathrm{Q}_{n+1}, \mathrm{P}_{n+1}) = \psi_{h,H}(\mathrm{Q}_n, \mathrm{P}_n)$, $n = 0, 1, \ldots$.

3. *Postprocessing:* If output at time $t = nh$ is desired, then find $(\mathrm{q}_n, \mathrm{p}_n) = \chi^{-1}_{h,H}(\mathrm{Q}_n, \mathrm{P}_n)$, which provides the numerical approximation to $(\mathrm{q}(nh), \mathrm{p}(nh))$.

The cost of preprocessing can be ignored, because preprocessing is performed only once in each integration. Note that

$$(\mathrm{q}_{n+1}, \mathrm{p}_{n+1}) = \chi^{-1}_{h,H}(\psi_{h,H}(\chi_{h,H}(\mathrm{q}_n, \mathrm{p}_n)))$$

and therefore the processed solutions can be interpreted as unprocessed solutions computed with the method

$$\widehat{\psi}_{h,H} = \chi_{h,H}^{-1} \circ \psi_{h,H} \circ \chi_{h,H}. \tag{3}$$

Processing is of interest if $\widehat{\psi}_{h,H}$ is a more accurate method than $\psi_{h,H}$ and the cost of postprocessing is negligible, either because output is not frequently required or because $\chi_{h,H}^{-1}$ is cheaply evaluated. Then, processing provides the accuracy of $\widehat{\psi}_{h,H}$ at the cost of the less accurate method $\psi_{h,H}$. López-Marcos, Sanz-Serna and Skeel[8] have showed how, in many cases, pre- and postprocessing may be carried out at virtually no cost. The idea of processing goes back to Butcher[2]. Further references are given by López-Marcos, Sanz-Serna and Skeel[8].

In what follows we always assume that the basic method $\psi_{h,H}$ is symplectic. Then, *formally,* the basic method provides (Sanz-Serna and Calvo[11], Section 10.1) exact solutions of a perturbed Hamiltonian system whose Hamiltonian function $\tilde{H}_h$ is a perturbation of the true Hamiltonian (1) of the system (2) being integrated. This is a characteristic feature of symplectic integrators; for a nonsymplectic method the computed points also lie on the solutions of a perturbation of the system (2), but the perturbed system is not Hamiltonian. If the preprocessor $\chi_{h,H}$ is a canonical or symplectic mapping, then the processed method (3) is also symplectic and then provides approximations that solve exactly its associated perturbed Hamiltonian system. The modified Hamiltonian of the processed method is (López-Marcos, Sanz-Serna and Skeel[8])

$$\widetilde{\widehat{H}}_h(\mathbf{q},\mathbf{p}) = \tilde{H}_h(\mathbf{Q},\mathbf{P}) = \tilde{H}_h(\chi_{h,H}(\mathbf{q},\mathbf{p})). \tag{4}$$

The aim of processing is then, given $\psi_{h,H}$ (i.e., given $\tilde{H}_h$), to find a symplectic transformation $\chi_{h,H}$ so that the right hand side of (4), that drives the processed solution, is as close as possible to the Hamiltonian H that drives the true solution.

Let us now focus on basic methods that are time-reversible. The corresponding modified Hamiltonian is of the form (López-Marcos, Sanz-Serna and Skeel[8])

$$\begin{aligned}\tilde{H}_h(\mathbf{Q},\mathbf{P}) &= H(\mathbf{Q},\mathbf{P}) \\ &\quad + h^2\frac{A}{2}[\mathbf{P}^T M^{-1} V_{\mathbf{QQ}}(\mathbf{Q}) M^{-1}\mathbf{P}] \\ &\quad + h^2\frac{B}{2}[V_{\mathbf{Q}}(\mathbf{Q})^T M^{-1} V_{\mathbf{Q}}(\mathbf{Q})] + O(h^4),\end{aligned} \tag{5}$$

Here A and B are method-dependent constants and $V_{\mathbf{QQ}}$ is the Hessian matrix of V. The terms in square brackets are so-called elementary Hamiltonians. If $A = B = 0$, then $\tilde{H}_h$ and the true H differ in $O(h^4)$ terms and the basic method has order four; otherwise the order of accuracy is only two.

Since the preprocessor $(\mathbf{Q},\mathbf{P}) = \chi_{h,H}(\mathbf{q},\mathbf{p}) = id + O(h^2)$ has to be a symplectic transformation, it will be the exact solution flow of a Hamiltonian system. In other words, a Hamiltonian function H_χ has to exist such that $(\mathbf{Q},\mathbf{P})$ is the value at time h of the solution with initial condition $(\mathbf{q},\mathbf{p})$ of the Hamiltonian system associated with H_χ. The expansion of H_χ will be of the form

$$H_\chi = h\lambda[\mathbf{p}^T M^{-1} V_{\mathbf{q}}(\mathbf{q})] + O(h^3),$$

with λ an undetermined parameter; no $O(1)$ nor $O(h^2)$ contribution is included because (5) possesses no $O(h)$ nor $O(h^3)$ term. The $O(h)$ term includes the only elementary Hamiltonian of order two. We conclude that the preprocessor may be sought in the form (López-Marcos, Sanz-Serna and Skeel[8])

$$\begin{aligned}\mathbf{Q} &= \mathbf{q} + h^2\lambda[M^{-1}V_{\mathbf{q}}(\mathbf{q})] + O(h^4), \\ \mathbf{P} &= \mathbf{p} - h^2\lambda[V_{\mathbf{qq}}(\mathbf{q})M^{-1}\mathbf{p}] + O(h^4),\end{aligned} \tag{6}$$

and by substitution in (5) we find, in view of (4),

$$\begin{aligned}\widetilde{\widehat{H}}_h(\mathbf{q},\mathbf{p}) &= H(\mathbf{q},\mathbf{p}) \\ &\quad + h^2\left(\frac{A}{2} - \lambda\right)[\mathbf{p}^T M^{-1} V_{\mathbf{qq}}(\mathbf{q}) M^{-1}\mathbf{p}] \\ &\quad + h^2\left(\frac{B}{2} + \lambda\right)[V_{\mathbf{q}}(\mathbf{q})^T M^{-1} V_{\mathbf{q}}(\mathbf{q})] + O(h^4).\end{aligned} \tag{7}$$

It is clear that, for the processed method to be of order four, i.e., for the basic method to be of effective order four, it is necessary and sufficient that the system $A/2-\lambda = 0$, $B/2 - \lambda = 0$ may be solved for λ. Obviously this happens if and only if

$$A = -B, \tag{8}$$

a condition that should be compared with the condition $A = B = 0$ for the basic method to be of order four.

3. Stability Intervals

When the basic method with modified Hamiltonian (5) is applied to the integration of the harmonic oscillator $H(q,p) = (1/2)(p^2 + q^2)$, the computed points satisfy

$$\begin{bmatrix} Q_{n+1} \\ P_{n+1} \end{bmatrix} = M_\psi \begin{bmatrix} Q_n \\ P_n \end{bmatrix},$$

where M_ψ is the amplification matrix of the method, which should approximate the matrix

$$\begin{bmatrix} \cos h & \sin h \\ -\sin h & \cos h \end{bmatrix} \tag{9}$$

that advances the true solution. For stability, the eigenvalues of M_ψ must have modulus ≤ 1. Since the method is symplectic, $\det(M_\psi) = 1$, and then it is well

known that stability is equivalent to M_ψ having trace of modulus ≤ 2. When this condition is satisfied, both eigenvalues have unit modulus.

Our aim is to discuss the stability of methods that are explicit, symplectic, time-reversible, use three force evaluations per step, and are effectively of order four. What does the trace of M_ψ look like for those methods? For explicit methods the entries (and hence the trace) of M_ψ are polynomials in h. For an explicit method using m force evaluations per step the trace is a polynomial $P(z)$ of degree $\leq m$ in the variable $z = h^2$: the formula for updating $\mathbf{q}$ will have nested evaluations of the force and each evaluation brings along a factor h^2. It remains to ascertain how the coefficients of $P(z)$ are constrained by the requirement of effective order four. The expansion of M_ψ in powers of h may be easily found if we recall that a step of the basic method is equivalent to advancing h units of time with the true solution of the modified Hamiltonian (5), which for the harmonic oscillator reads

$$\tilde{H}_h(Q,P) = \frac{1}{2}\left((1+Ah^2)P^2 + (1+Bh^2)Q^2\right) + O(h^4).$$

In this way we find

$$M_\psi = \begin{bmatrix} 1 - \frac{h^2}{2} + \left(\frac{1}{24} - \frac{A+B}{2}\right)h^4 + O(h^6) & h + \left(-\frac{1}{6} + A\right)h^3 + O(h^5) \\ -h + \left(\frac{1}{6} - B\right)h^3 + O(h^5) & 1 - \frac{h^2}{2} + \left(\frac{1}{24} - \frac{A+B}{2}\right)h^4 + O(h^6) \end{bmatrix}.$$

From here, we see that the condition (8) for effective order four is equivalent to

$$\mathrm{trace}(M_\psi) = 2 - h^2 + \frac{h^4}{12} + O(h^6). \tag{10}$$

Summing up, if the basic method is explicit, uses three evaluations per step and is of effective order four, the trace of M_ψ is of the form

$$P(z) = 2 - z + \frac{z^2}{12} + \alpha z^3, \quad z = h^2, \tag{11}$$

where α is a free parameter. Let us determine α so as to have the largest stability region. For $\alpha = 0$, $P(z)$ is a parabola with a minimum value $P(6) = -1$; stability is lost at $z = 12$ when $P = 2$. If $\alpha > 0$, then, for $z > 0$, the graph of P is strictly above the $\alpha = 0$ parabola. Hence $\alpha > 0$ is less stable than $\alpha = 0$. For $\alpha < 0$ and close to 0, the graph of $P(z)$ intersects the line $P = 2$ at $z_0 = 0$, $z_1 \approx 12$, and $z_2 \gg 1$ and the line $P = -2$ at $z_3 > z_2$. Stability is then restricted by the intersection at z_1. As α decreases away from 0, z_1 increases (thereby increasing stability) and z_2 and z_3 decrease. When α reaches the value $-1/576$, the points z_1 and z_2 coalesce and the equation $P = -2$ has a unique real root $z_3 \approx 32.3$, so that stability is restricted by the intersection at z_3. A further decrease in α implies a decrease in z_3 and hence in the length of the stability interval. Therefore the longest stability interval occurs at $\alpha = -1/576$ and the optimal trace is

$$P(z) = 2 - z + \frac{z^2}{12} - \frac{z^3}{576}, \quad z = h^2. \tag{12}$$

The stability interval has length $L \approx 5.69$ and scaled length $L/3 \approx 1.89$. This is within 5% of Verlet's optimal 2.

So far we have just discussed the stability of the basic method $\psi_{h,H}$. What is the stability interval of the processed method $\hat{\psi}_{h,H}$? It is clear that it should be the same as that of the basic method, because it is really the basic method that is being used to propagate the numerical solution. From a more mathematical point of view, we note that from (3), the amplification matrices of the basic and processed methods are related through

$$M_{\hat{\psi}} = M_\chi^{-1} M_\psi M_\chi, \tag{13}$$

where M_χ is the matrix that, for the harmonic oscillator, transforms the variables (q,p) into the variables (Q,P). Thus $M_{\hat{\psi}}$ and M_ψ are related by a similarity transformation and have the same eigenvalues and the same stability properties.

We close this section with two comments on the relation (13). If we think of $\psi_{h,H}$ as given and try to find an optimal processor $\chi_{h,H}$, we see from (13) that, for the harmonic oscillator, the most we can achieve by processing is to change the eigenvectors of M_ψ into the exact eigenvectors, i.e., into the eigenvectors of (9), without changing the eigenvalues. The eigenvectors of the amplification matrix govern the shape of the numerical trajectories on the (q,p) plane; the eigenvalues governs the phase of the numerical solution on its trajectory. Similar considerations apply to any linear problem.

Since M_ψ and $M_{\hat{\psi}}$ have the same trace, if $\hat{\psi}_{h,H}$ is of order four, then the trace of M_ψ differs from the trace $2\cos h$ of the exact (9) in $O(h^5)$ terms (a fourth-order scheme introduces $O(h^5)$ errors in one step). The trace is even in h, so that actually $\mathrm{trace}(M_\psi) = 2\cos h + O(h^6)$. This provides an alternative derivation of the formula (10).

4. Constructing the Method

We now show that the optimal trace polynomial (12) can be realized by an explicit, time-reversible, symplectic RKN method using three function evaluations per step. For three evaluations one may choose between two formats. In the first, the method has three stages. In the second—which we choose—the method has four stages but possesses the FSAL (first same as last) property, whereby the last force evaluation of the current step provides the first force evaluation to be used at the next step. For time-reversible methods choosing between both formats is just a matter of convenience, as we will discuss later.

With the requirements of symplecticness and time reversibility, an explicit, four-stage, FSAL RKN method is given by (Okunbor and Skeel[9], Sanz-Serna and Calvo[11], Section 8.5)

$$\mathbf{P}_n^1 \;=\; \mathbf{P}_n - h\left(\frac{1}{2} - b\right) V_{\mathbf{Q}}(\mathbf{Q}_n^*),$$

$$
\begin{aligned}
\mathbf{Q}_n^1 &= \mathbf{Q}_n + h\left(\frac{1}{2}-\gamma\right) M^{-1}\mathbf{P}_n^1,\\
\mathbf{P}_n^2 &= \mathbf{P}_n^1 - hbV_{\mathbf{Q}}(\mathbf{Q}_n^1),\\
\mathbf{Q}_n^2 &= \mathbf{Q}_n^1 + 2h\gamma M^{-1}\mathbf{P}_n^2,\\
\mathbf{P}_n^3 &= \mathbf{P}_n^2 - hbV_{\mathbf{Q}}(\mathbf{Q}_n^2),\\
\mathbf{Q}_{n+1} &= \mathbf{Q}_n^2 + h\left(\frac{1}{2}-\gamma\right) M^{-1}\mathbf{P}_n^3,\\
\mathbf{P}_{n+1} &= \mathbf{P}_n^3 - h\left(\frac{1}{2}-b\right) V_{\mathbf{Q}}(\mathbf{Q}_{n+1}),
\end{aligned}
\tag{14}
$$

where γ and b are free parameters. In the compact notation of Sanz-Serna and Calvo[11], Sections 8.4, 8.5, this method is described as

$$
[\tfrac{1}{2}-b, b, b, \tfrac{1}{2}+b], \qquad (\tfrac{1}{2}-\gamma, 2\gamma, \tfrac{1}{2}-\gamma, 0); \tag{15}
$$

the coefficients of the first group are used for the p variables and those of the second group for the q variables. The square brackets surrounding the p coefficients indicate that the first updating in (14) affects the p variables.

Note that if we define $\mathbf{Q}_n^* = \mathbf{Q}_n^1 + h\gamma M^{-1}\mathbf{P}_n^2$, $\mathbf{P}_n^* = \mathbf{P}_n^2$, then $(\mathbf{Q}_n^*, \mathbf{P}_n^*)$ is an approximation to the solution at $t = (n+1/2)h$, i.e., halfway through the step. Furthermore it is trivial to check that the transformation $(\mathbf{Q}_n^*, \mathbf{P}_n^*) \to (\mathbf{Q}_{n+1}^*, \mathbf{P}_{n+1}^*)$ that maps one halfway approximations into the next is in fact a RKN step with the method

$$
(b, 1-2b, b, 0), \qquad [\gamma, \tfrac{1}{2}-\gamma, \tfrac{1}{2}-\gamma, \gamma], \tag{16}
$$

that only has three stages (see formula (8.19) in Sanz-Serna and Calvo[11]). Therefore the four-stage, FSAL method (14) is related through the change of variables $(\mathbf{Q}_n, \mathbf{P}_n) \to (\mathbf{Q}_n^*, \mathbf{P}_n^*)$ to the three stage method (16). By considerations similar to those we presented towards the end of the preceding section, both methods then have the same stability properties. This proves that when looking for RKN methods that realize the optimal trace polynomial, the two formats mentioned above (three stages or four stages with FSAL) are equivalent.

After applying the method (14) to the harmonic oscillator, we find that the trace of the amplification matrix is given by

$$
2 - z + 2b(1-2\gamma)\left(\frac{1}{4}-\frac{b}{2}+\frac{\gamma}{2}\right) z^2 - \frac{b^2}{2}(1-2b)\gamma(1-2\gamma)^2 z^3,
$$

and, after comparing with (12), we have to consider the system

$$
\begin{aligned}
2b(1-2\gamma)\left(\frac{1}{4}-\frac{b}{2}+\frac{\gamma}{2}\right) &= \frac{1}{12}, &(17)\\
-\frac{b^2}{2}(1-2b)\gamma(1-2\gamma)^2 &= -\frac{1}{576}. &(18)
\end{aligned}
$$

This has the solution

$$
\gamma_{\text{opt}} = \frac{2 + 2^{1/3} + 2^{-1/3}}{6} \approx 0.6756, \quad b_{\text{opt}} = \frac{1 - 2^{1/3} - 2^{-1/3}}{6} \approx -0.1756. \tag{19}
$$

We recall from Section 4 that (17) is necessary and sufficient for effective order four. Hence the method (14) with parameter values (19) provides an explicit, symplectic, time-reversible RKN method with effective order four and optimal stability interval.

A final observation. From (19), $b_{\text{opt}} = 1/2 - \gamma_{\text{opt}}$ and therefore the compact notation (15) of the optimal method can be also be written as

$$
[\gamma_{\text{opt}}, \tfrac{1}{2}-\gamma_{\text{opt}}, \tfrac{1}{2}-\gamma_{\text{opt}}, \gamma_{\text{opt}}], \qquad (b_{\text{opt}}, 1-2b_{\text{opt}}, b_{\text{opt}}, 0).
$$

Comparison with (16) shows that the optimal method for the three-stage format can be obtained from the optimal method in the FSAL, four-stage format by swapping the roles of the $\mathbf{Q}$ and $\mathbf{P}$ variables.

5. The Accuracy of the New Method

It is of interest to compare analytically the method (14), (19) with other explicit, symplectic, time-reversible, (effectively) fourth-order methods.

To simplify the discussion we assume that the methods are applied to a linear Hamiltonian problem with potential $V(\mathbf{q}) = (1/2)\mathbf{q}^T S\mathbf{q}$, where S is a constant stiffness matrix. Then the modified Hamiltonian $\tilde{H}_h$ of the unprocessed method is of the form (López-Marcos, Sanz-Serna and Skeel[8])

$$
\begin{aligned}
\tilde{H}_h(\mathbf{Q},\mathbf{P}) &= H(\mathbf{Q},\mathbf{P})\\
&\quad + h^2\frac{A}{2}[\mathbf{P}^T M^{-1} S M^{-1}\mathbf{P}]\\
&\quad + h^2\frac{B}{2}[\mathbf{Q}^T S M^{-1} S\mathbf{Q}]\\
&\quad + h^4\frac{C}{2}[\mathbf{P}^T M^{-1} S M^{-1} S M^{-1}\mathbf{P}]\\
&\quad + h^4\frac{D}{2}[\mathbf{Q}^T S M^{-1} S M^{-1} S\mathbf{Q}] + O(h^6).
\end{aligned}
$$

Here A, B, C and D are method-dependent constants. This differs from the earlier expansion (5) in that we have substituted for the Hessian of V its current constant value S and we have displayed the $O(h^4)$ terms in the expansion. Recall from (8) that, since we are dealing with methods of effective order four, we may set $A = -B$.

The transformation $\chi_{h,H}$ is sought in the form (López-Marcos, Sanz-Serna and Skeel[8])

$$
\mathbf{Q} = \mathbf{q} + h^2\lambda[M^{-1}S\mathbf{q}] + h^4\left(\frac{\lambda^2}{2}+\mu\right)[M^{-1}SM^{-1}S\mathbf{q}] + O(h^6),
$$

Table 1. Error constant and stability interval of effectively fourth-order methods

Method	Error constant E	Stability Interval L
FRCR unprocessed	0.054	1.57
FRCR processed	0.047	1.57
Optimal stability (processed)	0.00037	5.69
Rowlands (processed)	0.00098	3.46

$$\mathbf{P} = \mathbf{p} - h^2\lambda[SM^{-1}\mathbf{p}] + h^4\left(\frac{\lambda^2}{2} - \mu\right)[SM^{-1}SM^{-1}\mathbf{p}] + O(h^6).$$

Again this differs from the earlier (6) in that we have substituted for V and displayed the $O(h^4)$ terms. By substituting as required by (4), we find

$$\begin{aligned}\widetilde{\widetilde{H}}_h(\mathbf{q},\mathbf{p}) &= H(\mathbf{q},\mathbf{p}) \\ &+ h^2\left(\frac{A}{2} - \lambda\right)[\mathbf{p}^T M^{-1}SM^{-1}\mathbf{p}] \\ &- h^2\left(\frac{A}{2} - \lambda\right)[\mathbf{q}^T SM^{-1}S\mathbf{q}] \\ &+ h^4\left(\frac{C}{2} + \lambda^2 - A\lambda - \mu\right)[\mathbf{p}^T M^{-1}SM^{-1}SM^{-1}\mathbf{p}] \\ &+ h^4\left(\frac{D}{2} + \lambda^2 - A\lambda + \mu\right)[\mathbf{q}^T SM^{-1}SM^{-1}S\mathbf{q}] + O(h^6).\end{aligned}$$

Clearly we have to set $\lambda = A/2$ to achieve a processed method $\widehat{\psi}_{h,H}$ of order four. Furthermore, as in López-Marcos, Sanz-Serna and Skeel[8], we set $\mu = (C - D)/4$ so as to minimize

$$E = \left(\left(\frac{C}{2} + \lambda^2 - A\lambda - \mu\right)^2 + \left(\frac{D}{2} + \lambda^2 - A\lambda + \mu\right)^2\right)^{1/2}, \qquad (20)$$

a measure of the size of the $O(h^4)$ error coefficients in the Hamiltonian $\widetilde{\widetilde{H}}_h$ of the processed method.

We now study the error constant and the stability interval of different methods. We first (Table 1) do so without taking into account the work per step and then (Table 2) report values scaled by work.

In Table 1 we have provided the size of the error constant (20) and the length L of the stability interval. The acronym FRCR refers to the three-stage, fourth-order, symplectic RKN method constructed by Forest and Ruth[5] and Candy and Rozmus[3]. This is the only RKN method that with three stages achieves order four. It has been noted before (Sanz-Serna and Calvo[11], Section 9.1) that FRCR possesses large error constants. We now see in the table that its stability interval is short. The method only benefits slightly by processing.

Table 2. Scaled error constant and scaled stability interval of effectively fourth-order methods

Method	Error constant $mE^{1/4}$	Stability Interval L/m
FRCR unprocessed	1.45	0.52
FRCR processed	1.40	0.52
Optimal stability (processed)	0.42	1.90
Rowlands (processed)	0.35	1.73

A comparison of FRCR with the optimal stability method bears out the advantages of processing and of the associated concept of effective order. When looking for methods that use three force evaluations per step, the requirement of order four leads to FRCR; by relaxing this requirement to effective order four, we have been able to find a method with error constant one hundred times smaller and stability intervals almost four times larger.

'Rowlands' refers to the method of effective order four introduced by Rowlands[10] and given by (the superscripts on the square brackets refer to evaluation at $\mathbf{Q}_n$ or $\mathbf{Q}_{n+1}$)

$$\begin{aligned}\mathbf{P}^{n+1/2} &= \mathbf{P}^n + \frac{h}{2}[-V_{\mathbf{Q}} + \frac{1}{12}h^2 V_{\mathbf{QQ}}M^{-1}V_{\mathbf{Q}}]^n, \\ \mathbf{Q}^{n+1} &= \mathbf{Q}^n + hM^{-1}\mathbf{P}^{n+1/2}, \\ \mathbf{P}^{n+1} &= \mathbf{P}^{n+1/2} + \frac{h}{2}[-V_{\mathbf{Q}} + \frac{1}{12}h^2 V_{\mathbf{QQ}}M^{-1}V_{\mathbf{Q}}]^{n+1}. \end{aligned} \qquad (21)$$

This uses the Hessian $V_{\mathbf{QQ}}$ of the potential and has an FSAL property: the last square bracket in (21) will be reused at the next step. Therefore, per step, one needs one evaluation of $V_{\mathbf{Q}}$ and one evaluation of $V_{\mathbf{QQ}}$. For many problems the cost per step of (21) is the same as that of an RKN method using two force evaluations per step (López-Marcos, Sanz-Serna and Skeel[8]). The trace of (21) is given by (11) with $\alpha = 0$, so that $L = 2\sqrt{3}$, and the scaled stability interval has length $\sqrt{3}$. This is only 13% smaller than Verlet's 2.

It useful to point out that (21) can be seen as a limiting case of methods of the form (14). Assume that we impose the constraint (17) on the parameters γ and b in (14). This leaves a family of methods of effective order four depending on a single parameter, say γ. After expressing b as a function of γ, let γ tend to 1/2 in (14). Then a little analysis shows that the limiting scheme is given by (21). As γ decreases from the optimal value in (19) to Rowlands' 1/2, the error constant E increases and the stability interval L decreases. We see in the table that N and L are almost twice as good for the optimal method than for Rowlands. However we should take into account that while for $\gamma \neq 1/2$, (14) costs *three force evaluations* per step, the limiting scheme only costs *two*.

Table 2 contains essentially the same information that Table 1, but the error constant E and stability length L have been replaced by their normalized counterparts

$mE^{1/4}$ and L/m (m is the number of force evaluations per step) to account for the fact that different methods require different amounts of work per step. Note that $mE^{1/4}$ measures the work required to achieve a target error. From the table we conclude that Rowlands method is more efficient than the optimal stability method. This suggests the need for further investigating symplectic methods that use the Hessian of the potential.

6. Numerical Experiments

The analytical comparisons of the preceding section were based on a linear problem. We complemented our assessment of the meth ods considered by carrying out numerical experiments for a highly nonlinear problem. We integrated the liquid argon problem described in López-Marcos, Sanz-Serna and Díaz[7] (see also López-Marcos, Sanz-Serna and Skeel[8]). We implemented the optimal stability method and the Rowlands method, both with the cheap processing suggested by López-Marcos, Sanz-Serna and Skeel[8]. The FRCR method was not considered: its poor practical performance has already been discussed in Sanz-Serna and Calvo[11], Chapter 9.

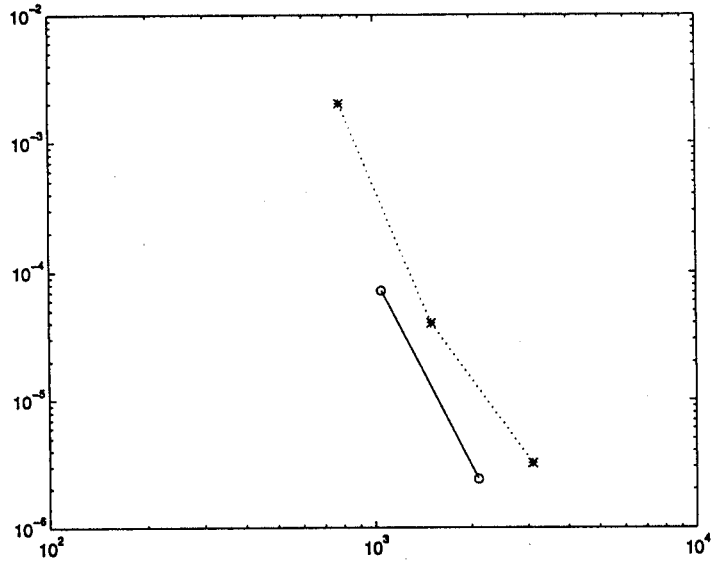

Fig. 1. Relative energy error against number of function evaluations. Rowlands method (solid line) and optimal stability method (broken line)

Figure 1 is an efficiency plot as those presented in López-Marcos, Sanz-Serna and Skeel[8]. The vertical axis gives the relative error in the Hamiltonian or total energy H. The energy error is very relevant in symplectic integration: it measures the difference

between the exact and modified Hamiltonians. The horizontal axis measures work. For the RKN method with optimal stability region, the work equals the number of force evaluations (i.e., is three times as large as the number of steps). For the Rowlands method, work is measured as twice the number of steps. The solid line and circles corresponds to Rowlands method ran with $h = 0.128$ and $h = 0.064$ ($h = 0.256$ was unstable). The broken line with stars corresponds to the optimal stability interval with $h = 0.256, 0.128, 0.0064$ ($h = 0.512$ was unstable). We first note that, in this nonlinear problem, the optimal stability method is able to operate successfully with steplength $h = 0.256$, while the Rowlands method is not. This matches the fact that the linear analysis yields a shorter stability interval for Rowlands' than for the optimal stability method. The plot also shows that, again in agreement with the linear analysis, the Rowlands method is more efficient than the optimal stable scheme.

7. Acknowledgements

MALM and JMS have been partly supported by grant DGICYT PB92-254. This work was done while RDS was Iberdrola Visiting Professor at the University of Valladolid, with additional support from DOE/NSF grant DE-FG02-91ER25099/DMS-9304268, NIH grant P41RR05969 and NSF/ARPA grant ASC-9318159.

8. References

1. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Clarendon Press, Oxford, 1987.
2. J. Butcher, in *Conference on the Numerical Solution of Differential Equations*, Lecture Notes in Math., Vol. 109, Springer, Berlin, 1969.
3. J. Candy, and W. Rozmus, *J. Comput. Phys.* **92**, 230–256, 1991.
4. M. M. Chawla and S. R. Sharma, *BIT* **21**, 455–464, 1981.
5. E. Forest and R. D. Ruth, *Physica* **D43**, 105–117, 1990.
6. E. Hairer, S. P. Nørsett and G. Wanner, *Solving Ordinary Differential Equations I, Nonstiff Problems*, 2nd. Ed., Springer, Berlin, 1993.
7. M. A. López-Marcos, J. M. Sanz-Serna and J. C. Díaz, *J. Comput. Appl. Math.* (1995, to appear).
8. M. A. López-Marcos, J. M. Sanz-Serna and R. D. Skeel, in *Numerical Analysis 1995*, Proceedings of the 1995 Dundee Conference, (1996, to appear).
9. D. Okunbor and R. D. Skeel, *Math. Comput.* **59**, 439–455, 1992.
10. G. Rowlands, *J. Comput. Phys.* **97**, 235–239, 1991.
11. J. M. Sanz-Serna, and M. P. Calvo, *Numerical Hamiltonian Problems*, Chapman & Hall, London, 1994.