SIAM J. Numer. Anal.
Vol. 34, No. 4, pp. 1391–1417, August 1997

ERROR GROWTH IN THE NUMERICAL INTEGRATION OF PERIODIC ORBITS, WITH APPLICATION TO HAMILTONIAN AND REVERSIBLE SYSTEMS[*]

B. CANO[†] AND J. M. SANZ-SERNA[†]

Abstract. We analyze in detail the growth with time (of the coefficients of the asymptotic expansion) of the error in the numerical integration with one-step methods of periodic solutions of systems of ordinary differential equations. Variable stepsizes are allowed. We successively consider "general," Hamiltonian, and reversible problems. For Hamiltonian and reversible systems and under fairly general hypotheses on the orbit being integrated, numerical methods with relevant geometric properties (symplecticness, energy-conservation, reversibility) are proved to have better error growth than "general" methods.

Key words. periodic solutions, Hamiltonian systems, reversible systems, conservation of energy, symplectic integrators, Runge–Kutta methods, error growth, asymptotic expansion of the error

AMS subject classifications. 65L05, 70F05, 70H05

PII. S0036142995281152

1. Introduction.

The purpose of this paper is the analysis of the error growth in the numerical integration by one-step methods of periodic orbits of systems of differential equations. Periodic solutions and solutions that are small perturbations of periodic orbits appear very frequently in the applications. Also, as Butcher [5] points out, "problems with periodic solutions are convenient as test problems for differential equation software because of the ease with which the accuracy of the computed solutions can be assessed." Often, when periodic orbits are used as test problems, the integration is carried out for many periods of the solution. These considerations prove the interest in investigating how integrators perform on periodic orbits and in particular in the long-time integration of periodic orbits.

The second author's interest in periodic orbits started in the paper [7] with Calvo. An optimized, explicit, *symplectic,* fourth-order Runge–Kutta–Nyström formula was developed there and compared with an optimized, explicit, *nonsymplectic,* 3/4 embedded Runge–Kutta–Nyström pair of Dormand, El-Mikkawy, and Prince [9, Table 3]. Kepler's problem was used for numerical tests and the numerical results showed that errors grow quadratically with time for the nonsymplectic, variable stepsize algorithm and only linearly for the symplectic algorithm *with constant stepsizes.* Section 4 of [7] was devoted to the mathematical analysis of these behaviors. Due to the context in which it arose, the analysis in [7] was essentially focused on Kepler's problem, in spite of the fact that the underlying ideas have a much wider applicability. The present paper avoids such a limitation in scope and considers general periodic orbits. Furthermore, the techniques of proof in [7] require a number of unnatural hypotheses; for instance, a period of the orbit was supposed to contain a whole number of timesteps. These artificial hypotheses have been completely removed here. The present paper subsumes the contents of Section 4 in [7].

[*] Received by the editors February 6, 1995; accepted for publication (in revised form) October 5, 1995. This research was supported in part by Dirección General de Investigación Científica y Técnica (DGICYT) grant PB92-254.

http://www.siam.org/journals/sinum/34-4/28115.html

[†] Departamento de Matemática Aplicada y Computación, Facultad de Ciencias, Universidad de Valladolid, Valladolid, Spain (bego@cpd.uva.es, sanzserna@cpd.uva.es).

The article [7] stimulated further work [6], [13] on error growth in the integration of periodic orbits. We explore the relations between [6], [13], and the present research in section 8. The papers [14], [15] present cases of linear versus quadratic growth in partial differential equation problems.

Section 2 contains some preliminary material. We begin (section 2.1) with the introduction of the initial value problem to be solved, which, for simplicity in the presentation, is supposed to be smooth in the whole of the phase space. The inclusion here of the extension to cases where the system is only C^k in a domain would have resulted in a longer paper but not in any new mathematical idea. Section 2.1 also describes the methods being used, which include virtually all one-step methods, and the (variable) stepsize strategies that are allowed. The paper relies throughout on the existence of an asymptotic expansion for the global error. This is presented in section 2 as well. We also provide some background on periodic orbits, including the notions of monodromy matrix and Floquet multipliers.

Section 3 studies, for general periodic problems, the growth of the coefficients e_m of the expansion of the error. It turns out that, if r denotes the order of the integrator, then, for $r \leq m \leq 2r-1$, the value of the e_m when the integration has been going on for N periods is completely determined by (i) the value $e_m^{(1)}$ of e_m after one period and (ii) the monodromy matrix of the orbit being integrated. The latter measures the change in the periodic orbit when its initial value suffers an infinitesimal perturbation. If there is a Floquet multiplier > 1 then the e_m grow exponentially as expected. For a hyperbolic attracting orbit the components of e_m transversal to the orbit remain bounded, but there is a linearly increasing phase error.

Section 4 looks at the important particular case where the system being studied is Hamiltonian. In section 4.1 we show how, for Hamiltonian systems, the monodromy matrix has some specific properties. These properties imply that situations that are "generic," i.e., the rule, for "general" systems cannot arise at all in the Hamiltonian case. Conversely, there are situations that are generic for Hamiltonian systems and exceptional for general systems. As a consequence, the error-growth behavior for the Hamiltonian case is special and deserves a separate study. We have devoted section 4.2 to this point. Often Hamiltonian systems are integrated by numerical methods with conservation properties, such as energy conservation or symplecticness. In section 4.3 we prove that for these methods the values $e_m^{(1)}$ which, along with the monodromy matrix determine the error growth, satisfy some constraints. It then follows that there are many periodic Hamiltonian problems where general integrators possess quadratic error growth, while energy-conserving methods and symplectic methods only lead to linear error growth. We emphasize that in the symplectic case constant stepsizes are necessary for the linear error growth.

Section 5 considers the case of reversible systems. Our treatment is parallel to that given in section 4 to the Hamiltonian case. We show how reversibility constrains the monodromy matrix and the consequences for error growth. For reversible methods the $e_m^{(1)}$ are also constrained, and it is possible to have linear error growth in important cases where general integrators lead to quadratic growth. Unlike the symplectic cases the stepsizes may be variable provided that they are reversible. This has attracted some recent attention on reversible integrators following the work by Hut, Makino, and McMillan [18]. Also in this section we point out an interesting property of nonreversible integrators of even order of accuracy.

Section 6 uses Kepler's problem to illustrate the material in sections 4 and 5. In numerical experiments symplectic formulas with variable stepsizes lead to quadratic error growth, while reversible, variable stepsize algorithms lead to linear error growth. We show that for large eccentricities a reversible, variable stepsize fourth-order method outperforms a standard fourth-order code.

Section 7 is technical. We present there the existence of the asymptotic expansion used throughout the paper.

Section 8 contains some concluding remarks.

Traditionally, the analysis of timestepping methods is based on the ideas of consistency and stability. Consistency means small local errors and stability means that local errors do not have a catastrophic effect on global errors. Stability is usually measured by a *number* (stability constant) that measures the relation between the *size* of the global error and the *size* of the local error. In actual fact, local and global errors are *vectors* in phase space and their *direction* matters: in the integration, a local error of a given size in one direction may be amplified by a large factor while another local error of the same size may be amplified by a smaller factor. As discussed in [14], [15], the geometric properties of the system being integrated (reversibility, Hamiltonian character) determine which directions in phase space are relatively more harmful. Numerical methods with good geometric properties have local errors that avoid those harmful directions. In this paper we witness similar phenomena: not all directions in phase space are treated in the same way by the monodromy matrix and geometric integrators have their errors in the “good” directions. We finally point out that one or another geometric property appears to be essential here: it is not enough for the integrators to have amplification factors of unit modulus along the imaginary axis [21].

2. Preliminaries.

2.1. The numerical method.

We consider an initial value problem

$$\dot{x} = f(x), \tag{1}$$

$$x(t_0) = x_0 \in R^D, \tag{2}$$

where for simplicity we assume that f is smooth (C^∞) in the whole of R^D. All the results in this paper can be easily adapted to the case where f is only of class C^k in a domain $\Omega \subset R^D$. The symbol φ_t refers to the t-flow of the system (1), [22, Section 2.1]. For each fixed t, φ_t is a mapping $\varphi_t : R^D \to R^D$ and, by definition, $\varphi_t(\alpha)$ is the value at time t of the solution of (1) with initial value α at time zero. For simplicity in the presentation, it is assumed that, for each real t, φ_t is defined in the whole of R^D.

We are concerned with a one-step integration formula for (1). This is given ([8, Chapter 4], [22, Chapter 3]) by a mapping $\psi_h : R^D \to R^D$ that advances the solution h units of time. For instance, $\psi_h(x) = x + hf(x)$ corresponds to Euler's rule. For implicit formulas $\psi_h(x)$ is defined implicitly. The mapping ψ_h is assumed to possess the following properties.

(i) A value $h_0 > 0$ exists such that for $|h| \leq h_0$ the domain of ψ_h is the whole of R^D. Practical explicit methods satisfy this assumption. Implicit methods satisfy this assumption if f has bounded derivatives [4, Corollary 341B]; see [22, Section 3.3.3] for a discussion. Note also that negative values of h are allowed in ψ_h; they move the solution backwards in time.

(ii) $\psi_h(x)$ depends smoothly on h and x; for Runge–Kutta and other methods used in practice this smoothness is an automatic consequence of the smoothness of (1).

(iii) The mapping ψ_h is consistent of order, say, $r \geq 1$, r an integer. This means that for each $x \in R^D$ the local error at x defined by $\varphi_h(x) - \psi_h(x)$ is $O(h^{r+1})$ as $h \to 0$.

(iv) The consistency of order r of ψ_h also holds in the C^1 topology, i.e., at each $x \in R^D$ the $D \times D$ Jacobian matrices $\psi_h'(x)$ and $\varphi_h'(x)$ satisfy

$$\varphi_h'(x) - \psi_h'(x) = O(h^{r+1}), \quad h \to 0. \tag{3}$$

For Runge–Kutta and other practical methods (iv) is a consequence of (iii); see, e.g., [11], [12].

Note that from assumptions (ii) and (iii) the local error possesses an asymptotic expansion

$$\varphi_h(x) - \psi_h(x) = h^{r+1}\lambda_{r+1}(x) + h^{r+2}\lambda_{r+2}(x) + \cdots, \tag{4}$$

where the $\lambda' s$ are smooth functions of x.

Of course the numerical solution is found by iteration of the mapping ψ_h,

$$x_{n+1} = \psi_{h_n}(x), \quad t_{n+1} = t_n + h_n, \quad n = 0, 1, 2, \dots, \tag{5}$$

where $h_0, h_1, h_2, \dots$ is a given sequence of positive stepsizes. Then x_n is an approximation to $x(t_n)$, where $x(t) = \varphi_{t-t_0}(x_0)$ denotes the solution of (1), (2).

Throughout the paper, we examine the case where the stepsizes are determined by

$$h_n = \epsilon s(x_n, \epsilon), \quad \epsilon > 0, \quad n = 0, 1, 2, \dots. \tag{6}$$

Here we assume that

(v) $s(x, \epsilon)$ is a smooth real-valued function defined in $R^D \times [-1, 1]$ such that, for suitable positive constants $s_{\min}$ and $s_{\max}$ and all $x \in R^D$ and ϵ, $|\epsilon| \leq 1$,

$$s_{\min} \leq s(x, \epsilon) \leq s_{\max}. \tag{7}$$

Stoffer and Nipp [28] have proved that (6) "almost" holds for all the standard stepsize selection strategies including the use of embedded pairs or Richardson's extrapolation. The reader is strongly advised to read the original paper for a precise understanding of the applicability of the Stoffer and Nipp result to actual codes. The nonstandard stepsize strategy suggested by Hut, Makino, and McMillan [18] and implemented in this paper is actually of the form (6); see section 5.3. Of course, $s \equiv 1$ provides constant stepsizes.

Note that, when (5), (6) are applied to the integration of (1), (2), the user only supplies the value of the parameter ϵ and the integrator returns sequences $x_n = x_n(\epsilon)$, $t_n = t_n(\epsilon)$.

2.2. The variational equation.

It is well known that if the initial condition x_0 in (2) is perturbed and becomes $x_0 + \delta_0$, δ_0 small, then the perturbed solution of the initial-value problem is approximately given by $x(t) + \delta(t)$, where $\delta(t_0) = \delta_0$ and δ satisfies the variational equation

$$\dot{\delta} = J(t)\delta, \quad J(t) = f'(x(t)) \tag{8}$$

(f' is the Jacobian matrix of f). The system (8) is linear with variable coefficients, and therefore its solution with initial value δ_0 at time t_0 is given by $\delta(t) = M(t, t_0)\delta_0$, where $M(t, t_0)$ is the associated transition matrix. Recall that the matrix-valued function of two real arguments $M(\cdot, \cdot)$ is characterized by

$$\frac{\partial M(t,s)}{\partial t} = J(t)M(t,s), \quad M(s,s) = I$$

and satisfies

$$M(t_3, t_1) = M(t_3, t_2)M(t_2, t_1) \tag{9}$$

for all t_1, t_2, t_3.

In terms more precise than those used above, the difference $\varphi_{t-t_0}(x_0 + \delta_0) - \varphi_{t-t_0}(x_0)$ between the perturbed and unperturbed solutions is given by $M(t, t_0)\delta_0 + o(\|\delta_0\|)$ as $\delta_0 \to 0$. This shows that $M(t, t_0)$ is the value at x_0 of the Jacobian matrix of the flow φ_{t-t_0}; i.e.,

$$\varphi'_{t-t_0}(x_0) = M(t, t_0). \tag{10}$$

Similarly, we have that, for all t and s,

$$\varphi'_{t-s}(x(s)) = M(t, s). \tag{11}$$

2.3. Asymptotic expansion of the global error. Under the assumptions (i)–(v) in section 2.1, the global errors $x_n - x(t_n)$ of (5), (6) possess an asymptotic expansion in powers of ϵ,

$$\begin{aligned} x_n - x(t_n) = \epsilon^r e_r(t_n) + \epsilon^{r+1} e_{r+1}(t_n) + \cdots + \epsilon^{2r-1} e_{2r-1}(t_n) \\ + \epsilon^{2r-1} R_{2r-1}(t_n, \epsilon), \qquad \epsilon \to 0. \end{aligned} \tag{12}$$

Here the *error functions* e_m are smooth functions of t (independent of ϵ), and R_{2r-1} is a smooth function such that, for $\epsilon \to 0$, $R_{2r-1}(t, \epsilon)$ tends to zero uniformly in bounded time intervals. The existence of (12) is proved in section 7.

The error functions e_m satisfy nonhomogeneous versions of the variational equation (8)

$$\dot{e}_m = J(t)e_m + \sigma_m(t), \quad e_m(t_0) = 0, \tag{13}$$

where the σ_m are suitable source terms (see section 7). Via the variation of constants formula the ${e_m}'s$ can be expressed in terms of the transition matrices M and the corresponding source

$$e_m(t) = \int_{t_0}^{t} M(t, s)\sigma_m(s)ds. \tag{14}$$

Remark. In (12) we have expanded only to order ϵ^{2r-1}, but it is possible to expand to any power of ϵ. We stop at ϵ^{2r-1} because this is what we need later.

2.4. Floquet multipliers. Let us now assume that the solution $x(t)$ of the initial value problem (1), (2) is T periodic, $x(t + T) \equiv x(t)$, $T > 0$. Then the matrix $J(t)$ in the variational equation is also T periodic, and it follows trivially that

$$M(t + T, s + T) \equiv M(t, s). \tag{15}$$

To avoid trivialities, we always assume that $x(t)$ is a genuine periodic solution and not an equilibrium; i.e., $x(t)$ is not constant. This is equivalent to the assumption that, for all t, $f(x(t)) \neq 0$.

From (15) and (9) we can write, for any real t and integer N,

$$\begin{aligned} M(t+NT,t) &= M(t+NT,t+(N-1)T) \\ &\quad \times M(t+(N-1)T,t+(N-2)T) \\ &\quad \times \ \dots \\ &\quad M(t+T,t) \\ &= M(t+T,t)^N. \end{aligned} \tag{16}$$

This is the *group property* of the transition matrices over multiples of the period T: advancing N periods in one go coincides with advancing N times over one period.

The matrix

$$M_t = M(t+T,t)$$

that effects the transition over one period starting at time t is called a *monodromy matrix* of the periodic solution x. Different t lead to different matrices M_t, but from (9) and (15)

$$M_{t_2}M(t_2,t_1) = M(t_2+T,t_1) = M(t_2+T,t_1+T)M(t_1+T,t_1) = M(t_2,t_1)M_{t_1},$$

so that M_{t_2} and M_{t_1} are related by a similarity transformation and share the same eigenvalues and Jordan structure. The eigenvalues of the monodromy matrices are called the *Floquet multipliers of the periodic solution* [16, Section 1.5]. Unity is always a Floquet multiplier because, by differentiation in (11),

$$M(t+T,t)f(x(t)) = f(x(t)), \tag{17}$$

so that $f(x(t))$ is an eigenvector of M_t with eigenvalue 1.

3. Error growth in the integration of periodic orbits: The general case.

Our aim in this section is to study the behavior as $t \to \infty$ of the error functions $e_m(t)$, $r \leq m \leq 2r-1$, of the numerical integrator (5), (6) when the solution $x(t)$ of (1), (2) is T periodic. A crucial observation (see section 7) is that *the sources* $\sigma_m(t)$, $r \leq m \leq 2r-1$ *in* (13) *are also* T *periodic.*

We begin with an auxiliary result.

LEMMA 3.1. *If* $t \in [t_0+(N-1)T, t_0+NT]$, N *an integer, then, for* $r \leq m \leq 2r-1$,

$$\begin{aligned} e_m(t) &= M(t-(N-1)T,t_0)\, e_m(t_0+(N-1)T) \\ &\quad + \int_{t_0}^{t-(N-1)T} M(t-(N-1)T,s)\sigma_m(s)ds. \end{aligned} \tag{18}$$

Proof. By (14) and (9)

$$\begin{aligned} e_m(t) &= \int_{t_0}^{t} M(t,s)\sigma_m(s)ds \\ &= \int_{t_0}^{t_0+(N-1)T} M(t,t_0+(N-1)T)M(t_0+(N-1)T,s)\sigma_m(s)ds \\ &\quad + \int_{t_0+(N-1)T}^{t} M(t,s)\sigma_m(s)ds. \end{aligned}$$

Now (18) follows after using (9) in the first integrand and changing variables $s = (N-1)T + s'$ in the second integral. □

We observe that in (18) the matrix $M(t-(N-1)T, t_0)$ and the integral are bounded uniformly in t, because $t-(N-1)T$ and t_0 differ at most by T. Therefore the growth of $e_m(t)$ as $t \uparrow \infty$ is governed by the growth of $e_m(t_0+(N-1)T)$ as $N \uparrow \infty$, i.e., *it is enough to look at the values of the error functions only when the computation has been going on for a whole number of periods.* Thus, in what follows, we look at the vectors defined by

$$e_m^{(N)} = e_m(t_0 + NT), \quad N = 1, 2, \ldots .$$

A second consequence of the lemma is the following result, which is essentially in [7, Section 4.2].

THEOREM 3.1. *In the situation above, for $r \le m \le 2r-1$,*

$$e_m^{(N)} = M_{t_0} e_m^{(N-1)} + e_m^{(1)}, \quad N = 2, 3, \ldots \tag{19}$$

and

$$e_m^{(N)} = \Big(\sum_{i=0}^{N-1} M_{t_0}^i \Big) e_m^{(1)}, \quad N = 1, 2, 3, \ldots . \tag{20}$$

Proof. Formula (20) follows from (19) by induction in N. To obtain (19) set $t = t_0 + NT$ in (18). □

Formula (20) is the main formula in this paper. It says that the error term $e_m^{(N)}$ after N periods is determined by (i) the error term $e_m^{(1)}$ after one period and (ii) the monodromy matrix M_{t_0}. Note that the latter depends only on the problem being solved and not on the particular numerical procedure (5), (6).

The growth with N of (20) can be investigated by a procedure very similar to that used to analyze the familiar power method for the computation of eigenvalues. The vector $e_m^{(1)}$ is decomposed according to the eigenvectors and generalized eigenvectors of M_{t_0}. This decomposition reduces the study of (20) to the study of expressions $\sum M^i e$, where e is a known vector and M is a Jordan block of M_{t_0}. In turn, the growth of $(\sum M)^i$ is governed by the following lemma.

LEMMA 3.2. *Assume that M is a $\mu \times \mu$ Jordan block with eigenvector λ. Then, as $N \uparrow \infty$,*

(i) *if $|\lambda| \ge 1$, $\lambda \ne 1$, then*

$$\| \sum_{i=0}^{N-1} M^i \| = O(N^{\mu-1} |\lambda|^N);$$

(ii) *if $|\lambda| < 1$, then*

$$\| \sum_{i=0}^{N-1} M^i \| = O(1);$$

(iii) *if $\lambda = 1$, $\mu > 1$, then*

$$\| \sum_{i=0}^{N-1} M^i \| = O(N^\mu);$$

(iv) *if* $\lambda = 1$, $\mu = 1$, *then*

$$\sum_{i=0}^{N-1} M^i = N.$$

Proof. In the first two cases,

$$\sum_{i=0}^{N-1} M^i = (M-I)^{-1}(M^N - I),$$

and the lemma is a consequence of the well-known behavior of M^N as $N \uparrow \infty$. In the third case, the entries of M^{N-1} are polynomials of degree $\mu - 1$ in N, and summation brings in an extra power of N. The final case is trivial. □

We are now ready to discuss the growth of the $e_m^{(N)}$. The following result is a direct consequence of Lemma 3.2 in tandem with (20).

THEOREM 3.2. *Assume that the numerical method* (5), (6) *satisfying the assumptions* (i), (v) *in section* 2.1 *is applied to the integration of the initial value problem* (1), (2) *with* T*-periodic solution. Then the following mutually exclusive possibilities arise:*

(G1) *The solution* $x(\cdot)$ *has a Floquet multiplier of modulus* > 1, *or in other words the monodromy matrix* M_{t_0} *has spectral radius* $R > 1$. *Then, for* $r \le m \le 2r-1$, $N \to \infty$, $e_m^{(N)} = O(N^{\mu-1}R^N)$, *where* μ *is the size of the largest Jordan block of* M_{t_0} *corresponding to the eigenvalues of modulus* R.

(G2) *All Floquet multipliers have modulus* ≤ 1. *Denote by* μ *the size of the largest Jordan block of* M_{t_0} *corresponding to eigenvalues* $\ne 1$ *of modulus* 1, *and denote by* μ_1 *the size of the largest Jordan block of* M_{t_0} *with eigenvalue* 1 ($\mu_1 \ge 1$ *by* (17)). *Then, for* $r \le m \le 2r-1$, $||e_m^{(N)}|| = O(N^\nu)$, *with* $\nu = \max(\mu - 1, \mu_1)$, *so that the growth is polynomial.*

A frequently occurring particular case of (G2) deserves special attention and is considered next.

(G2′) The periodic solution is hyperbolic and attracting; i.e. 1 is a simple Floquet multiplier, and the remaining $D-1$ multipliers have modulus < 1. This corresponds to case (G2) above with $\mu = 0$, $\mu_1 = 1$, and leads to linear error growth $||e_m^{(N)}|| = O(N)$. If we decompose $e_m^{(N)}$ according to eigenvectors and generalized eigenvectors of M_{t_0}, then the components of $e_m^{(N)}$ corresponding to multipliers $\ne 1$ remain bounded by (ii) in Lemma 3.2. In the direction of the eigenvector $f(x_0)$ associated with the multiplier 1 (see (17)), multiplication by $\sum_{i=0}^{N-1} M_{t_0}^i$ amounts to multiplication by N; see (iv) in Lemma 3.2. Hence the component along $f(x_0)$ is in $e_m^{(N)}$ exactly N times larger than in $e_m^{(1)}$. The conclusion is that the error term $e_m^{(N)}$ consists of (i) a bounded component transversal to the orbit and (ii) a component in the direction tangent to the orbit that is linear in N. This second component is a phase error.

The situation (G2′) is *structurally stable:* if a differential system has a periodic orbit in (G2′) then all neighboring differential systems have a periodic orbit in (G2′). Note also that for (G2′) it is known [3], [10], [28] that the numerical integrator possesses an attractive, closed, invariant curve close to the orbit being integrated. It is clear that as the numerical solution describes again and again this invariant curve the numerical error transverse to the orbit remains bounded. There is a linearly growing phase error because the numerical and true period do not coincide.

4. Error growth in the integration of periodic orbits: The Hamiltonian case.

4.1. Hamiltonian systems. We now look at the case where the system (1) being integrated is Hamiltonian. This means [1], [22] that the dimension D is even, $D = 2d$, and the vector field $f(x)$ is of the form

$$f(x) = \Xi^{-1}\nabla H(x), \tag{21}$$

where Ξ is the $2d \times 2d$ skew-symmetric matrix

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

and $\nabla H(x)$ is the gradient of a scalar function H (the Hamiltonian). It is standard to write

$$x = [p_1, \dots, p_d; q_1, \dots, q_d]^T,$$

and then (1), (21) reads

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad \dot{q}_i = \frac{\partial H}{\partial p_i}, \quad i = 1, \dots, d. \tag{22}$$

Among the properties of Hamiltonian systems we need the following two:

(i) For each t, the flow ϕ_t is a symplectic transformation, i.e., a transformation whose Jacobian matrix is, at each $x \in R^{2d}$, a symplectic matrix. Recall that a $2d \times 2d$ matrix M is said to be symplectic if $M^T \Xi M = \Xi$.

(ii) The Hamiltonian is a constant of motion; i.e., if $[p_1(t), \dots, q_d(t)]^T$ is a solution of (22), then $H(p_1(t), \dots, p_d(t), q_1(t), \dots, q_d(t))$ is constant. This is often the mathematical expression of the principle of *conservation of energy.*

These properties have implications in connection with the monodromy matrices M_t of the periodic solution $x(\cdot)$. From symplecticness we have the following.

LEMMA 4.1. *The Floquet multipliers* $\neq \pm 1$ *of a periodic orbit of a Hamiltonian problem appear in quadruples* λ, $\overline{\lambda}$, λ^{-1}, $\overline{\lambda}^{-1}$ $(|\lambda| \neq 1, \Im\lambda \neq 0)$ *or real pairs* $\lambda = \overline{\lambda}$, $\lambda^{-1} = \overline{\lambda}^{-1}$, *or unimodular pairs* $\lambda = \overline{\lambda}^{-1}$, $\overline{\lambda} = \lambda^{-1}$. *The multiplicities and Jordan structure of all four points of a quadruple (or both points of a pair) are the same. The (algebraic) multiplicity of the multiplier* 1 *is even* ≥ 2.

Proof. From (11) and the symplecticness of φ_t, M_t is a symplectic matrix. The eigenvalues of symplectic matrices appear in quadruples and pairs as those described in the lemma in [1, Section 42]. After removing from the $2d$ eigenvalues those $\neq \pm 1$ with their multiplicities, we are left with an even number. Hence the combined algebraic multiplicities of ± 1 are even. The algebraic multiplicity of -1 cannot be odd because this would lead to a negative determinant in M_t and monodromy matrices have positive determinant. Therefore the multiplicity of 1 is even. □

From conservation of energy we have the following.

LEMMA 4.2. *For each* $\delta_0 \in R^{2d}$,

$$\nabla H(x_0)^T \delta_0 = \nabla H(x_0)^T M_{t_0} \delta_0. \tag{23}$$

Proof. By conservation of energy

$$H(x_0 + \delta_0) = H(\varphi_T(x_0 + \delta_0)).$$

Taylor expansion with respect to δ_0 leads to

$$H(x_0) + \nabla H(x_0)^T \delta_0 + \cdots = H(\varphi_T(x_0) + \varphi_T^{'}(x_0)\delta_0 + \ldots),$$

so that periodicity and (10) imply

$$\begin{aligned} H(x_0) + \nabla H(x_0)^T \delta_0 + \cdots &= H(x_0 + M_{t_0}\delta_0 + \cdots) \\ &= H(x_0) + \nabla H^T M_{t_0} \delta_0 + \cdots \end{aligned}$$

and (23) follows. □

Note that (23) implies that M_{t_0} maps the subspace

$$S^\perp = \{v \in R^{2d} : \nabla H(x_0)^T v = 0\}$$

into itself. This is the linearized counterpart of the fact that φ_t maps the energy surface $\{x : H(x) = H(x_0)\}$ into itself. The dimension of $S^\perp$ is $2d-1$: $\nabla H(x_0) \neq 0$ because otherwise x_0 would be an equilibrium and $x(t)$ would not provide a genuine periodic orbit. Note also that the vector $f(x_0)$ tangent at x_0 to the periodic orbit is in $S^\perp$.

Let us denote by $M_{t_0}^\perp$ the linear operator in $S^\perp$ obtained by restriction of the linear operator in R^{2d} associated with the matrix M_{t_0}.

Lemma 4.3. *The Floquet multipliers of a periodic orbit of a Hamiltonian system are* 1 *and the* $2d-1$ *eigenvalues of the operator* $M_{t_0}^\perp$. *Unity is an eigenvalue of* $M_{t_0}^\perp$ *with odd (algebraic) multiplicity.*

Proof. The second assertion follows from the first because the (algebraic) multiplicity of 1 as an eigenvalue of M_{t_0} is even and ≥ 2. To prove the first assertion, we choose an orthonormal basis $e_1, \ldots, e_{2d}$ of R^{2d} formed by $e_1 = ||\nabla H(x_0)||^{-1} \nabla H(x_0)$ followed by an orthonormal basis of $S^\perp$. From (23) with $\delta_0 = \nabla H(x_0)$, we see that $e_1^T(M_t e_1) = 1$, so that the entry $(1,1)$ of the matrix $\tilde{M}$ that expresses the monodromy operator in the basis $\{e_1, \ldots, e_{2d}\}$ is 1. The entries $(1,2), \ldots, (1,2d)$ of $\tilde{M}$ are all zero because M_{t_0} maps $S^\perp$ into itself. Hence the eigenvalues of the monodromy operator are 1 and the $(2d-1)$ eigenvalues of the $(2d-1) \times (2d-1)$ bottom right submatrix of $\tilde{M}$. This submatrix represents $M_{t_0}^\perp$ in the basis $\{e_2, \ldots, e_d\}$. □

As a consequence of the lemma

$$R^{2d} = S^\perp \oplus \operatorname{Span}\{v\}, \tag{24}$$

where v is either an eigenvector or a generalized eigenvector associated with the eigenvalue 1, i.e., either $(M_{t_0} - I)v = 0$ or $(M_{t_0} - I)v \neq 0$, $(M_{t_0} - I)^k v = 0$, $k > 1$. Note that in general $v \neq \nabla H(x_0)$, because the latter need not be an eigenvector or generalized eigenvector; it is a left eigenvector by (23).

4.2. Hamiltonian periodic orbits and general integrators. After Lemma 4.1 it is easy to apply Theorem 3.2 to the Hamiltonian case. There are two possibilities corresponding to (G1) and (G2).

(H1) There is a Floquet multiplier of modulus $\neq 1$. Then, for $r \leq m \leq 2r-1$, $e_m^{(N)}$ grows exponentially with N.

(H2) All Floquet multipliers have modulus 1. Then, for $r \leq m \leq 2r-1$, $e_m^{(N)}$ grows polynomially.

Now recall that 1 is, in the Hamiltonian case, a multiple eigenvalue of M_{t_0}. Typically, it will be a double eigenvalue with a nontrivial 2×2 Jordan block; the situations where 1 is a quadruple or higher eigenvalue of M_{t_0} or 1 is a double nondefective eigenvalue disappear under arbitrarily small perturbations. Also, typically, the eigenvalues $\neq 1$ of M_{t_0} will be simple and by Theorem 3.2 the growth in case (H2) will be typically *quadratic*. Faster polynomial growth rates in (H2) are of course possible. The (atypical) situation leading to linear growth is the following.

(H2′) All Floquet multipliers have modulus 1 and trivial Jordan blocks, i.e., M_{t_0} diagonalizes. Then, for $r \leq m \leq 2r-1$, $||e_m^{(N)}|| = O(N)$. More precisely, the components of $e_m^{(N)}$ in the direction of the eigenvectors of M_{t_0} with eigenvalue 1 are in $e_m^{(N)}$ N times larger than in $e_m^{(1)}$ while the remaining components of $e_m^{(N)}$ remain bounded as $N \uparrow \infty$.

The solutions of the harmonic oscillator are in case (H2′): the monodromy matrix M_{t_0} is the identity.

4.3. Hamiltonian periodic orbits and special integrators. Often Hamiltonian problems are integrated by one-step methods with conservation properties; see [22] and the literature therein. For *energy-conserving methods*, $H(\psi_h(x)) = H(x)$ for all h and x, leading to $H(x_0) = H(x_1) = \cdots$. For *sympletic methods* ψ_h is a sympletic transformation. The error propagation mechanism for these methods is more favorable than that of general methods. This is due to the fact that, for schemes with conservation properties, the values $e_m^{(1)}$, $r \leq m \leq 2r-1$, that determine via (20) the error growth satisfy some constraints.

LEMMA 4.4. *Assume that the numerical method* (5), (6) *satisfying the assumptions* (i)–(v) *in section* 2.1 *is applied to the integration of the Hamiltonian initial value problem* (1), (2) *and* (21) *with* T*-periodic solution. Assume that the method is either symplectic with constant stepsizes or energy conserving. Then*

$$\nabla H(x_0)^T e_m^{(1)} = 0, \quad r \leq m \leq 2r-1. \tag{25}$$

Proof. We look first at the energy-conserving case. Consider the infinitely many values of ϵ for which t_0+T is one of the step points, say, t_n (clearly the value of n grows as $\epsilon \downarrow 0$). By conservation of energy, and noting that $x(t_0+T) = x_0$,

$$\begin{aligned} 0 &= H(x_n) - H(x_0) \\ &= H(x(t_0+T) + \epsilon^r e_r^{(1)} + \cdots + \epsilon^{2r-1} e_{2r-1}^{(1)} + O(\epsilon^{2r})) - H(x_0) \\ &= \nabla H(x_0)^T [\epsilon^r e_r^{(1)} + \cdots \epsilon^{2r-1} e_{2r-1}^{(1)}] + O(\epsilon^{2r}). \end{aligned}$$

This gives (25) in this case, because the $e_j^{(1)}$ are independent of ϵ.

For symplectic methods with constant stepsizes ϵ it is possible to find a modified Hamiltonian $\tilde{H}_\epsilon(x)$, with $\tilde{H}_\epsilon(x) = H(x) + O(\epsilon^r)$, $\nabla\tilde{H}_\epsilon(x) = \nabla H(x) + O(\epsilon^r)$, and such that the computed points x_n differ by $O(\epsilon^{2r})$ from the values $\tilde{x}_\epsilon(t_n)$ of the solution $x_\epsilon(t)$ of the Hamiltonian system with Hamiltonian $\tilde{H}_\epsilon(x)$, ($\tilde{x}_\epsilon(t_0) = x_0$). If t_0+T is a step point t_n,

$$\begin{aligned} 0 &= \tilde{H}_\epsilon(\tilde{x}(t_0+T)) - \tilde{H}_\epsilon(x_0) = \tilde{H}_\epsilon(x_n) - \tilde{H}_\epsilon(x_0) + O(\epsilon^{2r}) \\ &= \tilde{H}_\epsilon(x_0 + \epsilon^r e_r^{(1)} + \cdots + \epsilon^{2r-1} e_{2r-1}^{(1)}) - \tilde{H}_\epsilon(x_0) + O(\epsilon^{2r}) \\ &= \nabla\tilde{H}_\epsilon(x_0)^T [\epsilon^r e_r^{(1)} + \cdots + \epsilon^{2r-1} e_{2r-1}^{(1)}] + O(\epsilon^{2r}). \\ &= \nabla H(x_0)^T [\epsilon^r e_r^{(1)} + \cdots + \epsilon^{2r-1} e_{2r-1}^{(1)}] + O(\epsilon^{2r}). \end{aligned}$$

This leads to (25). □

Let us focus on the following situation.

(H2$''$) The periodic orbit $x(\cdot)$ of the Hamiltonian system (22) has Floquet multipliers of unit modulus and is such that (see Lemma 4.3) the operator $M_{t_0}^{\perp}$ diagonalizes.

Note that while M_{t_0} has necessarily a multiple eigenvalue and typically does not diagonalize, this is not the case for $M_{t_0}^{\perp}$. Therefore, while (H2$'$) is rather exceptional, (H2$''$) is commonly occurring. Indeed the following cases are within (H2$''$): (1) All periodic orbits of systems with one degree of freedom. In this case $M_{t_0}^{\perp}$ is one dimensional and certainly diagonalizes. (2) All periodic orbits where 1 is a double multiplier and the remaining multipliers are simple with unit modulus. Here (Lemma 4.3) $M_{t_0}^{\perp}$ has distinct eigenvalues. This situation is structurally stable, i.e., persistent under Hamiltonian perturbations. (3) Some special cases, including Kepler's problem, where all solutions in the energy surface $\{x : H(x) = H(x_0)\}$ are periodic with the same period T. Then $M_{t_0}^{\perp}$ is the identity.

Clearly, (H2$'$) is the (exceptional) particular case of (H2$''$), where the vector v in (24) is an eigenvector rather than a generalized eigenvector. For a case in (H2$''$) but not in (H2$'$), there is a 2×2 Jordan block for the eigenvalue 1, v is the only generalized eigenvector of M_{t_0}, $(M_{t_0} - I)^2 v = 0$, and the error growth is, for general methods, quadratic. However, some *special* methods still yield linear error growth. This will happen, according to (20) and Lemma 3.4, when (25) holds and hence $e_m^{(1)} \in S^{\perp}$. The quadratic error growth is not excited because the $e_m^{(1)}$'s have no component in the direction of the generalized eigenvector v in (24). Therefore, from the lemma we conclude the following result.

THEOREM 4.1. *Assume that the numerical method* (5), (6) *satisfying the assumptions* (i)–(v) *in section* 2.1 *is applied to the integration of the Hamiltonian initial value problem* (1), (2) *and* (21) *with T-periodic solution in the case* (H2$''$) *above. Assume that the method is either energy conserving or symplectic with constant stepsizes. Then, for $r \le m \le 2r-1$, $||e_m^{(N)}|| = O(N)$ as $N \uparrow \infty$. More precisely, if $e_m^{(N)}$ is decomposed according to the eigenvectors and generalized eigenvectors of M_{t_0}, the components of $e_m^{(N)}$ along the eigenvectors of M_{t_0} with eigenvalue* 1 *are N times larger than those of $e_m^{(1)}$, while the remaining components of $e_m^{(N)}$ remain bounded.*

We emphasize that in the symplectic case constant stepsizes are required [7], [22, Section 9.2], and the references therein.

5. Error growth in the integration of periodic orbits: The reversible case.

5.1. Reversible systems.

In this section we consider the case where the system (1) being integrated is reversible [2], [23], [25], [26], [27], [19]. Let ρ be an *involution* in R^D, that is, a linear mapping in R^D such that $\rho^2 = I$. To avoid trivial cases, we assume that $\rho \neq \pm I$. Then R^D is decomposed as a direct sum $R^D = X_+ \oplus X_-$, where $\rho v = v$ if $v \in X_+$, $\rho v = -v$ if $v \in X_-$, and the subspaces X_+, X_- have dimension ≥ 1. Vectors in X_+ (respectively, in X_-) are called symmetric (respectively, antisymmetric). The system (1) is said to be ρ-reversible if

$$f(\rho x) = -\rho f(x). \tag{26}$$

An important example is given by the Newton equations of mechanics

$$\dot{p}_i = F_i(q_1, \ldots, q_d), \quad \dot{q}_i = p_i, \quad i = 1, \ldots, d. \tag{27}$$

These are reversible with respect to

$$\rho[p_1, \ldots, p_d; q_1, \ldots, q_d]^T = [-p_1, \ldots, -p_d; q_1, \ldots, q_d]^T. \tag{28}$$

They are also Hamiltonian if the vector of forces $[F_1, F_2, \ldots, F_d]^T$ is the gradient of a scalar potential; then the Hamiltonian is given by $H = (1/2)p^Tp + V(q)$. It is clear that the example (27), (28) does not exhaust all reversible systems that are of interest in the applications; other examples may be seen in [23]. Even for the Newton equations (27), reversibility may hold for choice of ρ different from (28); for the Kepler problem studied later we essentially use the involution given by reflection with respect to the major semiaxis of the Keplerian ellipse instead of (28).

For each t, the flow φ_t of a ρ-reversible system is a ρ-reversible mapping; i.e.,

$$(\varphi_t)^{-1} = \rho\varphi_t\rho. \tag{29}$$

Since for flows $(\varphi_t)^{-1} = \varphi_{-t}$, an alternative formulation is $\varphi_{-t} = \rho\varphi_t\rho$. In the example (27), (28), this means that evolving backward in time is the same as changing the sign of the initial velocities, evolving forward in time, and reversing the sign of the final velocities.

Yet another equivalent formulation of the reversibility of the flow is

$$\rho\varphi_t\rho\varphi_t = Id. \tag{30}$$

This is convenient because it does not involve inverses or evolutions with $-t$.

A nontrivial periodic orbit $x(\cdot)$ of a ρ-reversible system (1), (26) is called *symmetric* [2] if the corresponding trajectory in phase space R^D intersects the invariant subspace X_+ of ρ. It is very easy to check that the trajectory of a symmetric T-periodic solution intersects X_+ at exactly two points; the solution takes $T/2$ units of time to move from one of the intersection points to the other. Furthermore, ρ maps the symmetric trajectory into itself.

Just as in the Hamiltonian case (Lemmas 4.1 to 4.3), monodromy matrices of reversible orbits have some special properties. These are studied next.

LEMMA 5.1. *Assume that the solution $x(\cdot)$ of the reversible initial value problem* (1), (2) *and* (26) *is periodic and symmetric and that $x_0 \in X_+$. Then the monodromy matrix M_{t_0} is a ρ-reversible matrix; i.e.,*

$$M_{t_0}{}^{-1} = \rho M_{t_0}\rho.$$

Proof. From (30),

$$\rho\varphi_T\rho\varphi_T = Id,$$

and computation of the Jacobian at x leads to

$$\rho \cdot \varphi_T^{'}(\rho\varphi_T(x)) \cdot \rho \cdot \varphi_T^{'}(x) = I.$$

When $x = x_0$, this reads

$$\rho M_{t_0}\rho M_{t_0} = I,$$

because $\rho x_0 = x_0$ and $\varphi_T(x_0) = x_0$. □

LEMMA 5.2. *The Floquet multipliers $\neq \pm 1$ of a symmetric periodic orbit of a reversible system appear in quadruples λ, $\overline{\lambda}$, λ^{-1}, $\overline{\lambda}^{-1}$ ($|\lambda| \neq 1$, $\Im\lambda \neq 0$) or real pairs $\lambda = \overline{\lambda}$, $\lambda^{-1} = \overline{\lambda}^{-1}$, or unimodular pairs $\lambda = \overline{\lambda}^{-1}$, $\overline{\lambda} = \lambda^{-1}$. The multiplicities and Jordan structure of all four points of a quadruple (or both points of a pair) are the* *same. If the dimension D is even, then the (algebraic) multiplicity of the multiplier* 1 *is even* ≥ 2.

Proof. Since the different monodromy matrices of a periodic orbit are related by similarity transforms (section 2.4), we may assume that the initial value x_0 is one of the two points where the trajectory intersects X_+. Then, by Lemma 5.1, M_{t_0} is a reversible matrix, and we may apply known results [23] on the spectral properties of reversible matrices. The issue of the multiplicity of 1 is settled as in Lemma 4.1. □

In the discussion of the error growth we shall need the following results.

LEMMA 5.3. *In the situation of Lemma* 5.1*,*

$$\mathrm{Ker}(M_{t_0} - I) = [\mathrm{Ker}(M_{t_0} - I) \cap X_+] \oplus [\mathrm{Ker}(M_{t_0} - I) \cap X_-].$$

In other words, the eigenvectors of M_{t_0} with eigenvalue 1 *may be spanned by a basis consisting of symmetric eigenvectors ($\rho v = v$) and antisymmetric eigenvectors ($\rho v = -v$).*

Proof. We first show that there is at least an eigenvector that is in X_+ or X_-. Choose $v \in \mathrm{Ker}(M_{t_0} - I)$, $v \neq 0$. If $v \in X_+$ we are done; in the other case $w = v - \rho v \neq 0$ and

$$M_{t_0} w = M_{t_0} v - M_{t_0} \rho v = v - \rho {M_{t_0}}^{-1} v = v - \rho v = w,$$

so that w is an eigenvector in X_-.

In a similar way we may prove that if $v_1{}^+, \ldots, v_j{}^+, v_1{}^-, \ldots, v_k{}^-$ are linearly independent eigenvectors, $v_i{}^\pm \in X_\pm$, and $v \notin X_+$ is an eigenvector independent of $v_1{}^+, \ldots, v_j{}^+, v_1{}^-, \ldots, v_k{}^-$, then $w = v - \rho v \in X_-$, $w \neq 0$ provides an eigenvector independent of $v_1^+, \ldots, v_j^+, v_1^-, \ldots, v_k^-$. Iteration of this argument constructs the required basis. □

LEMMA 5.4. *In the situation of Lemma* 5.1*,*

$$\mathrm{Ker}(M_{t_0} - I)^2 = [\mathrm{Ker}(M_{t_0} - I)^2 \cap X_+] \oplus [\mathrm{Ker}(M_{t_0} - I)^2 \cap X_-];$$

i.e., the kernel $\mathrm{Ker}(M_{t_0} - I)^2$ *can be spanned by symmetric and antisymmetric vectors. More precisely, there is a basis of* $\mathrm{Ker}(M_{t_0} - I)^2$ *of the form* $v_1{}^+, \ldots, v_j{}^+$*;* $v_1{}^-, \ldots, v_k{}^-$*;* $w_1{}^+, \ldots, w_l{}^+$*;* $w_1{}^-, \ldots, w_m{}^-$*, where*

(i) *the $v_i{}^+$ are symmetric eigenvectors $M_{t_0} v_i{}^+ = v_i{}^+ = \rho v_i{}^+$;*

(ii) *the $v_i{}^-$ are antisymmetric eigenvectors $M_{t_0} v_i{}^- = v_i{}^- = -\rho v_i{}^-$;*

(iii) *the $w_i{}^+$ are symmetric generalized eigenvectors $\rho w_i{}^+ = w_i{}^+$, $(M_{t_0} - I) w_i{}^+ \neq 0$;*

(iv) *the $w_i{}^-$ are antisymmetric generalized eigenvectors $\rho w_i{}^- = -w_i{}^-$,$(M_{t_0} - I) w_i{}^- \neq 0$.*

Furthermore, $(M_{t_0} - I) w_i{}^+ \in X_-$, $i = 1, \ldots, l$, $(M_{t_0} - I) w_i^- \in X_+$, $i = 1, \ldots, l$.

Proof. The proof is similar to that of the previous lemma and will not be given. □

Another useful result is the following.

LEMMA 5.5. *In the situation of Lemma* 5.1*, assume that I is a conserved quantity of the differential system being integrated and that ρ is orthogonal. If I is ρ invariant, i.e., $I(x) \equiv I(\rho x)$, then $\nabla I(x_0) \in X_+$.*

Proof. From $I(x) \equiv I(\rho x)$, it follows that $\nabla I(x) \equiv \rho^T \nabla I(\rho x)$, so that $\nabla I(x_0) \equiv \rho^T \nabla I(x_0)$. For an orthogonal involution, $\rho = \rho^{-1} = \rho^T$. □

For the Newton case (28), ρ is obviously orthogonal. Since the energy $H = (1/2)p^T p + V(q)$ is obviously ρ invariant, then we have that $\nabla H(x_0) \in X_+$.

5.2. Symmetric periodic orbits and general integrators. After Lemma 5.2 it is straightforward to apply Theorem 3.2 to symmetric periodic orbits. There are two possibilities corresponding to (G1) and (G2):

(R1) There is a Floquet multiplier of modulus $\neq 1$. Then, for $r \leq m \leq 2r-1$, $e_m^{(N)}$ grows exponentially with N.

(R2) All Floquet multipliers have modulus 1. Then, for $r \leq m \leq 2r-1$, $e_m^{(N)}$ grows polynomially.

As in section 4.2, *if the dimension D is even* (as in (27)), then the growth in (R2) is typically quadratic. The situation leading to linear error growth is the following:

(R2′) All Floquet multipliers have modulus 1 and trivial Jordan blocks. Then, for $r \leq m \leq 2r-1$, $||e_m^{(N)}|| = O(N)$.

We emphasize that for D even (R2′) is exceptional.

In section 5.5 below we return again to general integrators for symmetric periodic orbits.

5.3. Reversible integrators. Reversible systems (1)–(26) may be integrated by means of *reversible integrators*, [25], [26], [27]. The integrator in section 2.1 applied to a reversible system is called reversible if (i) it uses a reversible formula, i.e., for each h,

$$(\psi_h)^{-1} = \rho\psi_h\rho \tag{31}$$

(cf. (29)), and (ii) it uses a reversible stepsize selection function

$$s(x,\epsilon) \equiv s(\rho y, \epsilon), \tag{32}$$

where

$$y = \psi_{\epsilon s(x,\epsilon)}(x). \tag{33}$$

For reversible integrators,

$$\rho\psi_{\epsilon s(\rho y,\epsilon)}(\rho y) = \rho\psi_{\epsilon s(x,\epsilon)}(\rho y) = \left[\psi_{\epsilon s(x,\epsilon)}\right]^{-1}(y) = x,$$

so that the mapping $\chi(x) = \psi_{\epsilon s(x,\epsilon)}(x)$ that advances the numerical solution satisfies

$$\rho\chi(\rho(\chi(x))) \equiv x \tag{34}$$

or, in other words, is reversible (cf. (30)).

Stoffer [26], [27] has studied conditions under which the formula ψ_h satisfies (31). For Runge–Kutta and all other standard formulas, it is straightforward to check that

$$\rho\psi_h\rho = \psi_{-h} = (\psi_h^*)^{-1}, \tag{35}$$

where the adjoint ψ_h^* of ψ_h is, by definition, the scheme such that $\psi_h^* = (\psi_{-h})^{-1}$ [17, Section II.8]. In view of (35), it is clear that, for such standard formulas, (31) holds if and only if ψ_h is self-adjoint, i.e., is its own adjoint. Obviously self-adjoint schemes satisfy $\psi_{-h} = (\psi_h)^{-1}$; i.e., they are time reversible. Examples of Runge–Kutta methods that are self-adjoint and therefore ρ reversible include all collocation methods with symmetric abscissas [17]. Also note that if ψ_h is a method of order r, then $\bar{\psi}_h = \psi_{h/2}\psi_{h/2}^*$ defines a new method of order $\geq r$ that is self-adjoint; see, e.g., [7, Section 3.6.2].

Several ways of ensuring the reversibility (32) of the stepsize function are discussed by Stoffer [27]. Here we focus on the technique suggested by Hut, Makino, and McMillan [18]. Let $\tau(x)$ be a positive function defined in the phase space R^D. We may think that $\tau(x)$ provides a local characteristic time in such a way that, from the point of view of local accuracy, $\epsilon\tau(x)$ is a suitable steplength at x. However, the choice $s(x,\epsilon) = \tau(x)$ does not satisfy (32), and to ensure reversibility the steplength $h = \epsilon s(x,\epsilon)$ to be used at x is determined by iteratively solving the equation

$$h = \frac{\epsilon}{2}[\tau(x) + \tau(\psi_h(x))]. \tag{36}$$

This involves attempting several steps $\psi_{h^{(1)}}(x), \psi_{h^{(2)}}(x), \ldots$ from x, with

$$h^{(n)} = \frac{\epsilon}{2}[\tau(x) + \tau(\psi_{h^{(n-1)}}(x))]. \tag{37}$$

This iteration is stopped once a value $h = h^n$ is found such that $h^{(n)} \approx h^{(n-1)}$. Clearly, if (37) has to be applied ν times at a given x_n, then the cost of finding x_{n+1} is ν times the cost of evaluating ψ_h once. Therefore, this is a rather expensive technique. Alternative techniques described in [27] also require iteration.

It is straightforward to check that, by the implicit function theorem, (36) defines, for ϵ small, a steplength function $s(x,\epsilon)$ such that $s(x,\epsilon) = \tau(x) + O(\epsilon)$. Thus $s(x,\epsilon)$ is a small perturbation of the characteristic time $\tau(x)$. Furthermore, this construction ensures reversibility as the following result shows.

LEMMA 5.6. *If the formula ψ_h is reversible, i.e.,* (31) *holds, and*

$$\tau(x) \equiv \tau(\rho(x)), \tag{38}$$

then the stepsize function $s(x,\epsilon)$ defined implicitly by (36) *satisfies the reversibility condition* (32).

Proof. By (36) and (33),

$$\begin{aligned}\epsilon s(x,\epsilon) &= \frac{\epsilon}{2}[\tau(x) + \tau(\psi_{\epsilon s(x,\epsilon)}(x))]\\ &= \frac{\epsilon}{2}[\tau(\psi^{-1}_{\epsilon s(x,\epsilon)}(y)) + \tau(y)]\end{aligned}$$

and (31) yields

$$\epsilon s(x,\epsilon) = \frac{\epsilon}{2}[\tau(\rho\psi_{\epsilon s(x,\epsilon)}(\rho y)) + \tau(y)].$$

Now, by (38)

$$\epsilon s(x,\epsilon) = \frac{\epsilon}{2}[\tau(\rho y) + \tau(\psi_{\epsilon s(x,\epsilon)}(\rho y))],$$

an equality that, when compared with (36), shows that $\epsilon s(x,\epsilon)$ is the steplength to be used at ρy. □

5.4. Symmetric orbits and reversible integrators.

When the integrator is reversible, it is possible to have linear growth in situations far more general than (R2′). Let us first consider the following lemma that constrains the $e_m^{(1)}$ and thus plays here the role played by Lemma 4.4 in the Hamiltonian case.

LEMMA 5.7. *Assume that the solution of the reversible initial value problem* (1), (2) *and* (26) *is a symmetric periodic orbit, that $x_0 \in X_+$, and that the integrator*

(5), (6), *satisfying the assumptions* (i)–(v) *in section* 2.1 *is reversible. Then, for* $r \leq m \leq 2r-1$,

$$e_m^{(1)} = -M_{t_0}\rho e_m^{(1)}. \tag{39}$$

Proof. Restrict the attention to the values of ϵ for which t_0+T is a step point t_n and $n=n(\epsilon)$, and consider the mapping ψ_ϵ defined by

$$\psi_\epsilon = \psi_{h_n} \circ \cdots \circ \psi_{h_2} \circ \psi_{h_1},$$

where h_n is the sequence of the stepsizes used to move from x_0 to x_n. From hypothesis (iv) in section 2.1 and (10), $\psi_\epsilon{}' = M_{t_0} + O(\epsilon^r)$. From the reversibility of the method, $\psi_\epsilon(\rho x_n) = \rho x_0 = x_0$. Therefore,

$$\begin{aligned} x_0 &= \psi_\epsilon(x_0 + \epsilon^r \rho e_r^{(1)} + \cdots + \epsilon^{2r-1}\rho e_{2r-1}^{(1)} + O(\epsilon^{2r})) \\ &= \psi_\epsilon(x_0) + \epsilon^r M_{t_0}\rho e_r^{(1)} + \cdots + \epsilon^{2r-1} M_{t_0}\rho e_{2r-1}^{(1)} + O(\epsilon^{2r}), \end{aligned}$$

or, in view of the relation $\varphi_{t_n}(x_0) = x_0$,

$$-(x_n - \varphi_{t_n}(x_0)) = \epsilon^r M_{t_0}\rho e_r^{(1)} + \cdots + \epsilon^{2r-1} M_{t_0}\rho e_{2r-1}^{(1)} + O(\epsilon^{2r}),$$

and the result follows. □

We now look at the following situation.

(R2″) All Floquet multipliers of the symmetric orbit have modulus 1. The multipliers $\neq 1$ have trivial Jordan blocks. The sizes of the Jordan blocks for the multiplier 1 are ≤ 2 and furthermore, with the notation of Lemma 5.4, there are no antisymmetric generalized eigenvectors w_i^-.

(R2″) includes the following cases. (1) All symmetric orbits in two dimensions $D=2$. There is a symmetric direction and an antisymmetric direction, but the antisymmetric direction corresponds to the eigenvector $f(x_0)$ by (26). (2) All symmetric orbits where 1 is a double multiplier and the remaining multipliers are simple with unit modulus. An antisymmetric generalized eigenvector for the eigenvalue 1 would imply, via Lemma 5.4, the existence of a *symmetric* eigenvector. These two vectors and the antisymmetric eigenvector $f(x_0)$ would require at least a triple Floquet multiplier at 1. This case is structurally stable; i.e., it remains under reversible perturbations. (3) Some special cases, where ρ is orthogonal, there is an invariant conserved quantity I and all solutions on the surface $\{x : I(x) = I(x_0)\}$ are periodic with the same period T. Then all vectors orthogonal to $\nabla I(x_0)$ are eigenvectors with eigenvalue 1 and $\nabla I(x_0)$ is an eigenvector or generalized eigenvector that is symmetric by Lemma 5.5. Note that in all this discussion we have assumed that $x_0 \in X_+$. This involves no loss of generality. By the symmetry of the orbit there is a point on the trajectory that is in X_+; if this point is the initial value then we have $x_0 \in X_+$. Changing the initial value along the trajectory changes the monodromy matrix by a similarity transformation and preserves the Jordan structure.

THEOREM 5.1. *Assume that the solution of the reversible initial value problem* (1), (2) *and* (26) *is a symmetric periodic orbit of type* (R2″), *that* $x_0 \in X_+$, *and that the integrator* (5), (6) *satisfying the assumptions* (i)–(v) *in section* 2.1 *is reversible. Then, for* $r \leq m \leq 2r-1$, $||e_m^{(N)}|| = O(N)$ *as* $N \uparrow \infty$. *More precisely if* $e_m^{(N)}$ *is decomposed according to the eigenvectors and generalized eigenvectors of* M_{t_0}, *then*

(i) $e_m^{(N)}$ *has no component in the direction of the symmetric eigenvectors* v_i^+ *and symmetric generalized eigenvectors* w_i^+ *of the multiplier* 1;

(ii) *the components of* $e_m^{(N)}$ *in the direction of the antisymmetric eigenvectors of the multiplier* 1 *are* N *times larger than the corresponding components in* $e_m^{(1)}$*;*

(iii) *the remaining components of* $e_m^{(N)}$ *remain bounded as* $N \uparrow \infty$.

Proof. Let V_m^+, V_m^-, and W_m^+, respectively, be the components of $e_m^{(1)}$ in the direction of the symmetric eigenvectors, antisymmetric eigenvectors, and symmetric generalized eigenvectors of the multiplier 1. Recall that for (R2$''$) there is no antisymmetric generalized eigenvector. By (39)

$$\begin{aligned} V_m^+ + V_m^- + W_m^+ &= -M_{t_0}(V_m^+ - V_m^- + W_m^+) \\ &= -V_m^+ + V_m^- - M_{t_0} W_m^+ \end{aligned}$$

and, from Lemma 5.4, $M_{t_0} W_m^+ = W_m^+ + V^{*-}$, where V^{*-} is an antisymmetric eigenvector. Then

$$V_m^+ + V_m^- + W_m^+ = -V_m^+ + V_m^- - W_m^+ - V^{*-};$$

so that the uniqueness of the decomposition implies $W_m^+ = -W_m^+$ or $W_m^+ = 0$ and accordingly $V^{*-} = 0$. Thus $V_m^+ + V_m^- = -V_m^+ + V_m^-$ and hence $V_m^+ = 0$. This proves (i). Parts (ii) and (iii) are covered by Lemma 3.2. □

Remark. In the theorem the hypothesis $x_0 \in X_+$ is not essential; for a reversible integration of a symmetric orbit satisfying (R2$''$) the growth is $O(N)$ even if $x_0 \notin X_+$. To see this, assume that $x_0 \notin X_+$ and let t^* be the smallest time $t^* > t_0$ for which $x(t^*) \in X_+$. For $r \le m \le 2r-1$, the integral (14) that gives $e_m^{(N)}$ may be decomposed as follows:

$$\begin{aligned} e_m^{(N)} &= \int_{t_0}^{t_0+NT} M(t_0+NT, s)\sigma_m(s)ds \\ &= \int_{t^*+(N-1)T}^{t_0+NT} + \int_{t^*}^{t^*+(N-1)T} + \int_{t_0}^{t^*} = I_1 + I_2 + I_3. \end{aligned}$$

The integral I_2 may be rewritten as

$$\begin{aligned} &\int_{t^*}^{t^*+(N-1)T} M(t_0+NT, s)\sigma_m(s)ds \\ &= M(t_0+NT, t^*+(N-1)T) \int_{t^*}^{t^*+(N-1)T} M(t^*+(N-1)T, s)\sigma_m(s)ds \\ &= M(t_0+NT, t^*+(N-1)T)\hat{I}_2 = M(t_0+T, t^*)\hat{I}_2. \end{aligned}$$

The integral $\hat{I}_2$ is easily interpreted. Imagine a second integration of the symmetric orbit, starting from the initial condition $x(t^*) \in X_+$ at the initial time t^*. Then $\hat{I}_2$ is clearly the value of $e_m^{(N-1)}$ for this auxiliary integration. Since Theorem 5.1 applies, $||\hat{I}_2|| = O(N)$ and hence $||I_2|| = O(N)$.

For I_1 we may write, by periodicity,

$$I_1 = \int_{t^*}^{t_0+T} M(t_0+T, s)\sigma_m(s)ds$$

so that I_1 is independent of N. Finally,

$$I_3 = M_{t_0}^{N-1} \int_{t_0}^{t^*} M(t_0+T, s)\sigma_m(s)ds$$

grows like $O(N)$ in view of the Jordan structure of M_{t_0}.

5.5. Symmetric orbits and general integrators of even order. If the orbit is symmetric and satisfies (R2$''$) and the integrator, without being reversible, satisfies (35) and is of even order r, then it is possible to have linear growth in the *leading* error term $e_r^{(N)}$. This was first noticed in [7] for the particular case of Kepler's problem. For a proof, note that if r is even,

$$\psi_h^* - \psi_h = (\phi_h - \psi_h) - (\phi_h - \psi_h^*) = O(h^{r+2}),$$

because the leading terms of the local error of ψ_h and ψ_h^* coincide [17, Theorem 8.6]. From here we conclude that (31) is satisfied except for an $O(h^{r+2})$ remainder. If, with the notation of (32), (33), we consider stepsize selection functions such that

$$s(x, \epsilon) = s(\rho y, \epsilon) + O(\epsilon), \tag{40}$$

then (34) holds except for an $O(\epsilon^{r+2})$ remainder, and the proofs of Lemma 5.7 and Theorem 5.1 apply to the leading error term. We note that (40) holds if

$$s(x, \epsilon) = s(\rho x, \epsilon),$$

a natural condition that would be satisfied for the stepsize strategies used in standard codes.

THEOREM 5.2. *Assume that the solution of the reversible initial value problem* (1)*,* (2) *and* (26) *is a symmetric periodic orbit of type* (R2$''$)*. Assume that the integrator* (5)*,* (6) *satisfying the assumptions* (i)–(v) *in section* 2.1*,* (35)*, and* (40) *is of even order* r*. Then the conclusions of Lemma* 5.7 *and Theorem* 5.1 *hold for* $m = r$*.*

6. A numerical illustration.

The planar Kepler system is the Hamiltonian system (22) with two degrees of freedom ($d = 2$) and Hamiltonian function

$$H = \frac{1}{2}(p_1^2 + p_2^2) - \frac{1}{R}, \quad R = \sqrt{q_1^2 + q_2^2}.$$

This is integrated with the initial conditions

$$p_1 = 0, \quad p_2 = \sqrt{\frac{1+e}{1-e}}, \quad q_1 = 1 - e, \quad q_2 = 0,$$

where e is a parameter ($0 \le e < 1$). The solution is periodic with period $T = 2\pi$, and its projection onto the configuration (q_1, q_2)-plane is an ellipse with eccentricity e. The major axis of the ellipse lies along the $0q_1$ axis and the motion starts at $t_0 = 0$ from the pericenter of the ellipse. Recall that the pericenter lies on the major semiaxis and is the point on the ellipse closest to the focus, i.e., to the center of the attracting force.

Since we are dealing with a particular instance of the Newton equations (27), the system is ρ reversible for the involution ρ in (28). The invariant subspace X_+ for (28) is given by $p_1 = p_2 = 0$ and does not intersect the periodic trajectory under consideration. (For an elliptic Keplerian motion the components p_1, p_2 of the velocity never vanish simultaneously; planets never stand still.) Therefore, the periodic orbit is not symmetric with respect to (28). However, it is symmetric with respect to the alternative orthogonal involution

$$\bar{\rho}(p_1, p_2, q_1, q_2) = (-p_1, p_2, q_1, -q_2).$$

Clearly $\bar{\rho}$, restricted to the configuration space, represents the symmetry $(q_1, q_2) \rightarrow (q_1, -q_2)$ with respect to the major axis of the configuration ellipse. Note also that the hypothesis $x_0 \in X_+$ in Theorem 5.1 holds.

The monodromy matrix M_{t_0} is given by [7]

$$M_{t_0} = I + W_0 \nabla H(x_0)^T,$$

where $\nabla H(x_0)$ is the energy gradient at the initial point x_0 in R^4 and W_0 is a nonzero vector in R^4 in the direction of $f(x_0)$, i.e., tangent in phase space at x_0 to the trajectory being integrated. For vectors v orthogonal to $\nabla H(x_0)$, $M_{t_0} v = v$, so that with the notation in section 4, $M_{t_0}^{\perp}$ is the identity and all vectors in $S^{\perp}$ are eigenvectors with eigenvalue 1. On the other hand,

$$M_{t_0} \nabla H(x_0) = \nabla H(x_0) + ||\nabla H(x_0)||^2 W_0$$

and $\nabla H(x_0)$ is a generalized eigenvector with eigenvalue 1. Thus the solution is case (H2$''$) (but not case (H2$'$)). Furthermore, by Lemma 5.5, $\nabla H(x_0)$ is in the invariant subspace X_+ of $\bar{\rho}$. Thus the solution is in (R2$''$) (but not in (R2$'$)). From sections 4 and 5 we know that the $e_m^{(N)}$, $r \leq m \leq 2r-1$, grow quadratically for general integrators and linearly for energy-conserving integrators, constant stepsize symplectic integrators, and reversible integrators. As a side remark we point out that angular momentum [22, Section 1.2.4] is also a $\bar{\rho}$-invariant conserved quantity. Its gradient at x_0 is therefore a symmetric eigenvector. By Theorem 5.1, for reversible integrators, $e_m^{(N)}$, $r \leq m \leq 2r-1$, has no component in the direction of the energy or angular momentum gradients, which implies that the errors in energy and angular momentum at time $t_0 + NT$ behave as $O(\epsilon^{2r})$.

Since it is not possible to find an energy-preserving integration formula for this problem within the standard classes of one-step methods (Runge–Kutta, Runge–Kutta–Nyström, Taylor, etc.) we concentrate on symplectic and reversible methods.

The following formulas are considered.

(NRF). Nonreversible formula. We have chosen an optimized fourth-order, explicit, symplectic Runge–Kutta–Nyström formula developed by Calvo and Sanz-Serna [7]. This formula is not self-adjoint and hence is not reversible.

(RF). Reversible formula. A step of length h of this is defined by the concatenation $\psi^*_{h/2}\psi_{h/2}$ of a step of length $h/2$ with the formula of Calvo and a step of length $h/2$ with the adjoint formula. The concatenation is a fourth-order, symplectic, explicit [20] Runge–Kutta–Nyström formula that is self-adjoint and hence is reversible.

These two formulas were combined with the following stepsize functions $s(x, \epsilon)$.

(NRS). Nonreversible stepsize function. This is defined by $s(p_1, p_2, q_1, q_2, \epsilon) = \tau(p_1, p_2, q_1, q_2)$, where

$$\tau(p_1, p_2, q_1, q_2) = \min \left\{ \frac{\sqrt{q_1^2 + q_2^2}}{\sqrt{p_1^2 + p_2^2}}, \frac{\pi}{2\sqrt{2}} (q_1^2 + q_2^2)^{3/4} \right\}.$$

The first function in the min represents the time required to cover, at the current speed, the current distance to the center of the forces. The second function is the time for a free (zero initial velocity) fall into the center from the current configuration. This choice of τ is suggested in [18].

(RS). Reversible stepsize function. This is defined by (36) with τ as in NRS.

We have implemented three integrators, (NRF)-(RS), (RF)-(NRS), (RF)-(RS). All three use symplectic formulas, but only (RF)-(RS) is reversible in the sense of section 5.3. Figure 1 corresponds to the (NRF)-(RS) algorithm. The (logarithmic)

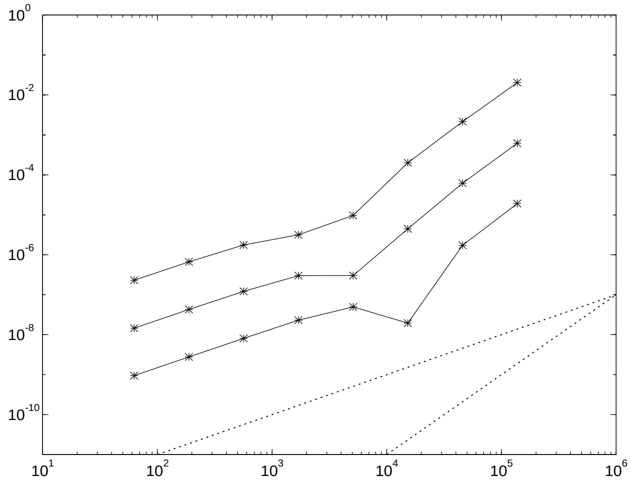

Fig. 1.

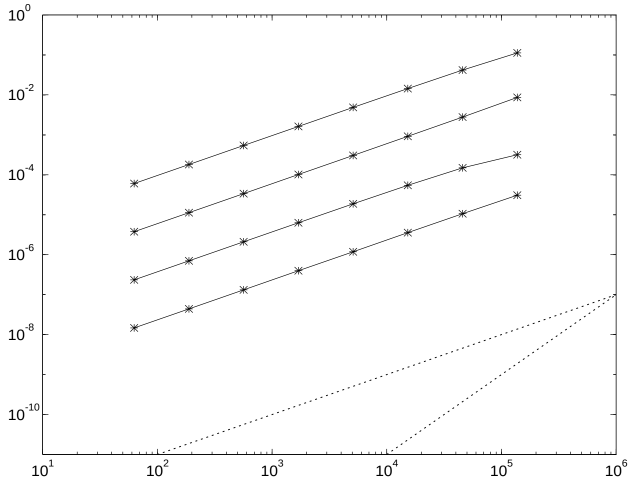

Fig. 2.

horizontal axis is the time t and the (logarithmic) vertical axis gives errors measured in the Euclidean norm of $\mathcal{R}^4$. The eccentricity is $e = 0.5$. The stars give the errors after 10, 30, 90, ..., 21870 periods of the motion for $\epsilon = 1/80$, $1/160$, $1/320$. Clearly the error growth is quadratic for $t > 10,000$ (for comparison we have included dotted lines with slopes 1 and 2 corresponding to linear and quadratic error growth). This figure confirms the results in [7]: symplectic integrators like NRF with variable stepsizes behave as general nonsymplectic integrators; see also [22, Section 9.2] and references therein. For t small, the quadratic growth does not manifest itself for reasons analyzed in section 5.5. For even-order methods $O(\epsilon^r)$ error term only grows linearly; for t small this term dominates the expansion, and the quadratic growth of the remaining terms is not yet visible.

Figure 2 is as Fig. 1, but now the algorithm is (RF)-(RS) and $\epsilon = 1/20$, $1/40$, $1/80$. Clearly the error growth is linear in agreement with the reversibility of the algorithm.

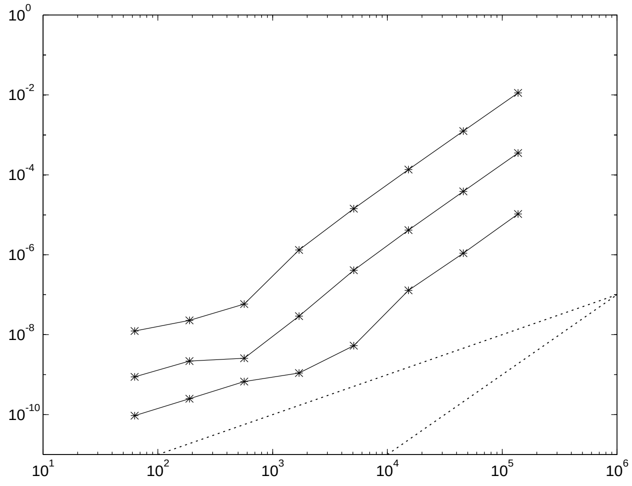

Fig. 3.

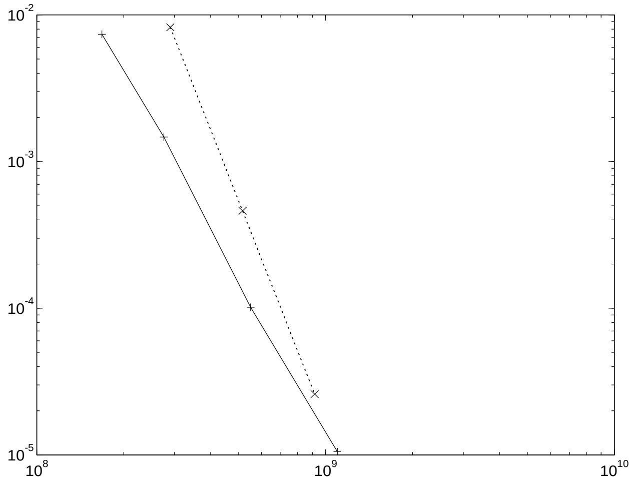

Fig. 4.

In Fig. 3 the algorithm is (RF)-(NRS) and $\epsilon = 1/80$, $1/160$, $1/320$. The growth is again quadratic. A reversible formula is not enough for linear growth: a reversible stepsize selection is also required.

As discussed in detail in [7], for symplectic formulas the advantages of linear growth and variable stepsizes cannot be combined. For fourth-order integrators and orbits of low or moderate eccentricity, it pays to use a symplectic formula with constant stepsize. For high eccentricities, say $e = 0.9$, the motion is very fast near the pericenter and very slow near the apocenter, and the best choice of a fourth-order integrator consists of a nonsymplectic formula with constant stepsizes. We have shown in Fig. 2 that with reversible integrators it is possible to have simultaneously linear growth and variable stepsizes. Unfortunately, the reversible stepsize function (RS) involves iteration, see (37), and makes each step of the algorithm rather expensive. Even though the purpose of this paper is not to discuss the relative efficiency of different kinds of algorithms, we present in Fig. 4 an efficiency plot for $e = 0.9$. The vertical axis is error after 21,870 periods ($t = 21,870\times 2\pi$) and the horizontal axis is work measured by the number of force evaluations. The solid line corresponds to the fourth-order Runge–Kutta–Nyström algorithm (RF)-(RS) run with $\epsilon = 1/40$, $1/80$, $1/160$, $1/320$. The dotted line corresponds to a standard, optimized 3–4, Runge–Kutta–Nyström embedded pair by Dormand, El-Mikkawy, and Prince [9] and [22, Example 5.1] run with tolerances 10^{-11}, 10^{-12}, 10^{-13}. The reversible algorithm is more efficient than the standard code. More experiments in this direction can be seen in [27].

Remark. Strictly speaking, Kepler's system of differential equations does not satisfy the hypotheses in section 2: there is a singularity at $R = 0$. However, Kepler's system would be covered by a modification of the theory in this paper catering to the case in which the domain of f is not the whole of $\mathcal{R}^D$. Alternatively, one may change the potential energy $-1/R$ away from the periodic orbit being integrated so as to render it globally smooth and even with bounded derivatives of all orders. Specifically, for the orbit at hand, $1-e \le R \le 1+e$, and we could change the potential for $R < (1-e)/2$ and $R > 2(1+e)$. For all the runs reported in Figs. 1–4 with the true, singular Kepler problem, the computed points have $(1-e)/2 < R < 2(1+e)$, so that the various methods are yielding the same solution they would have yielded had we used the regularized potential. Therefore, for the analysis, we can pretend we ran a regular problem.

7. Technical results. In the situation of section 2.1 let us prove the existence of the asymptotic expansion (12). We first need the following auxiliary result.

LEMMA 7.1. *Let v and w be the solutions of the initial value problems*

$$\dot v(t) = J(t)v(t), \quad v(t_n) = \alpha,$$
$$\dot w(t) = J(t)w(t) + \mu(t), \quad w(t_n) = \alpha,$$

where J is a smooth $D\times D$ matrix, μ is a smooth source term, t_n is a real value, and α is a given vector in R^D. Then we have the asymptotic expansion

$$w(t_n + h) - v(t_n + h) = hc_1(t_n) + h^2 c_2(t_n) + \cdots, \quad h \to 0,$$

where $c_1(t_n) = \mu(t_n)$ and the c_i, $i \ge 2$, are smooth functions that depend on J and μ but are independent of α. Furthermore, if J and μ are T periodic, then so are the functions c_i.

Proof. If $\delta(t) = w(t) - v(t)$, then $\dot\delta = J\delta + \mu$, $\delta(t_n) = 0$, and

$$w(t_n + h) - v(t_n + h) = h\dot\delta(t_n) + \frac{h^2}{2}\ddot\delta(t_n) + \cdots.$$

Differentiation in the differential equation leads succesively to

$$\begin{aligned}
\dot\delta(t_n) &= J(t_n)\delta(t_n) + \mu(t_n) = \mu(t_n),\\
\ddot\delta(t_n) &= \dot J(t_n)\delta(t_n) + J(t_n)\dot\delta(t_n) + \dot\mu(t_n) = J(t_n)\mu(t_n) + \dot\mu(t_n),\\
\dddot\delta\,(t_n) &= \ddot J(t_n)\delta(t_n) + 2\dot J(t_n)\dot\delta(t_n) + J(t_n)\ddot\delta(t_n) + \ddot\mu(t_n)\\
&= 2\dot J(t_n)\mu(t_n) + J(t_n)^2\mu(t_n) + J(t_n)\dot\mu(t_n) + \ddot\mu(t_n),
\end{aligned}$$

etc. This concludes the proof. □

To investigate the existence of the asymptotic expansion (12), fix a final time $t_{max} > t_0$ and consider the sequences $\{x_n\}$, $\{t_n\}$ delivered by the method with $t_0 \le t_n \le t_{max}$. Set

$$\hat x_n = x_n - \epsilon^r e_r(t_n) - \cdots - \epsilon^{2r-1}e_{2r-1}(t_n),$$

where the functions e_m, $r \le m \le 2r-1$, satisfy the variational problems (13) with (smooth) sources σ_m yet to be determined.

The $\hat{x}_n$ satisfy the one-step recursion

$$\hat{x}_{n+1} = \Psi(\hat{x}_n, t_n, h_n, \epsilon)$$

with

$$\begin{aligned}\Psi(\hat{x}, t, h, \epsilon) = &-\epsilon^r e_r(t+h) - \cdots - \epsilon^{2r-1} e_{2r-1}(t+h)\\ &+\psi_h(\hat{x} + \epsilon^r e_r(t) + \cdots + \epsilon^{2r-1} e_{2r-1}(t)).\end{aligned}$$

Therefore the "global errors" $\hat{x}_n - x(t_n)$, $t_0 \le t_n \le t_{max}$ can be bounded by a product CL, where C is a constant (that increases with t_{max}) and L is a bound for the quantities ${h_n}^{-1}\rho_n$, with

$$\begin{aligned}\rho_n &= x(t_{n+1}) - \Psi(x(t_n), t_n, h_n, \epsilon)\\ &= x(t_{n+1}) + \epsilon^r e_r(t_{n+1}) + \cdots + \epsilon^{2r-1} e_{2r-1}(t_{n+1})\\ &\quad -\psi_{h_n}(x(t_n) + \epsilon^r e_r(t_n) + \cdots + \epsilon^{2r-1} e_{2r-1}(t_n));\end{aligned} \tag{41}$$

the ρ_n play the role of "local errors." Our aim is to show that the sources σ_m can be chosen so that $\rho_n = h_n O(\epsilon^{2r})$.

We begin by adding and subtracting $\psi_{h_n}(x(t_n))$ in (41). This brings in the local error $x(t_{n+1}) - \psi_{h_n}(x(t_n))$ at t_n that may be expanded as in (4). Thus

$$\begin{aligned}\rho_n = &h_n^{r+1}\lambda_{r+1}(x(t_n)) + \cdots + h_n^{2r}\lambda_{2r}(x(t_n)) + O(h_n^{2r+1})\\ &+\epsilon^r e_r(t_{n+1}) + \cdots + \epsilon^{2r-1} e_{2r-1}(t_{n+1})\\ &+\psi_{h_n}(x(t_n)) - \psi_{h_n}(x(t_n) + \epsilon^r e_r(t_n) + \cdots + \epsilon^{2r-1} e_{2r-1}(t_n)).\end{aligned} \tag{42}$$

Next, in (42),

$$\begin{aligned}&\psi_{h_n}(x(t_n)) - \psi_{h_n}(x(t_n) + \epsilon^r e_r(t_n) + \cdots + \epsilon^{2r-1} e_{2r-1}(t_n))\\ &= -\psi'_{h_n}(x(t_n)) \cdot [\epsilon^r e_r(t_n) + \cdots + \epsilon^{2r-1} e_{2r-1}(t_n))] + h_n O(\epsilon^{2r}).\end{aligned} \tag{43}$$

Here we have noted that the second derivatives of ψ_h with respect to x are $O(h)$ because $\psi_{h=0}(x) = x$ by consistency. Furthermore, in (43) we may replace $\psi'_{h_n}(x(t_n))$ by $\varphi'_{h_n}(x(t_n))$, because these Jacobian matrices differ in terms $O(h_n^{r+1})$, i.e., $h_n O(\epsilon^r)$, that can be absorbed in the remainder. On combining (43) and (42), we obtain

$$\begin{aligned}\rho_n = &h_n^{r+1}\lambda_{r+1}(x(t_n)) + \cdots + h_n^{2r}\lambda_{2r}(x(t_n)) + O({h_n}^{2r+1})\\ &+\epsilon^r[e_r(t_{n+1}) - \varphi^{'}_{h_n}(x(t_n)) e_r(t_n)]\\ &\cdots\\ &+\epsilon^{2r-1}[e_{2r-1}(t_{n+1}) - \varphi^{'}_{h_n}(x(t_n)) e_{2r-1}(t_n)] + h_n O(\epsilon^{2r}).\end{aligned} \tag{44}$$

Here we note, see (11), that $\varphi'_{h_n}(x(t_n)) = M(t_{n+1}, t_n)$, so that, for $r \le m \le 2r-1$, $\varphi'_{h_n}(x(t_n)) e_m(t_n)$ is the value $v_m(t_n + h_n)$ at t_{n+1} of the solution of the homogeneous variational problem $\dot{v}_m = J v_m$, $v_m(t_n) = e_m(t_n)$. Since e_m satisfies a nonhomogeneous variational problem, Lemma 7.1 applies and

$$e_m(t_{n+1}) - \varphi'_{h_n}(x(t_n)) e_m(t_n) = h_n c_{m,1}(t_n) + h_n^2 c_{m,2}(t_n) + \cdots .$$

Substitution in (44) yields

$$\begin{aligned} {h_n}^{-1}\rho_n = {} & h_n^r\lambda_{r+1}(x(t_n)) + \cdots + h_n^{2r}\lambda_{2r}(x(t_n)) \\ & +\epsilon^r c_{r,1}(t_n) + \epsilon^r h_n c_{r,2}(t_n) + \cdots + \epsilon^r h_n^{r-1} c_{r,r}(t_n) \\ & +\epsilon^{r+1} c_{r+1,1}(t_n) + \epsilon^{r+1} h_n c_{r+1,2}(t_n) + \cdots + \epsilon^{r+1} h_n^{r-2} c_{r+1,r-1}(t_n) \\ & +\cdots \\ & +\epsilon^{2r-1} c_{2r-1,1}(t_n) + O(\epsilon^{2r}). \end{aligned} \tag{45}$$

In the right-hand side of (45), $h_n = \epsilon s(x_n, \epsilon)$ may be replaced by $\epsilon s(x(t_n), \epsilon)$ because $s(x_n, \epsilon) - s(x(t_n), \epsilon) = O(\epsilon^r)$. The result is

$$\begin{aligned} {h_n}^{-1}\rho_n = {} & \epsilon^r s(x(t_n),\epsilon)^r \lambda_{r+1}(x(t_n)) + \cdots + \epsilon^{2r} s(x(t_n),\epsilon)^{2r}\lambda_{2r}(x(t_n)) \\ & +\epsilon^r c_{r,1}(t_n) + \epsilon^{r+1} s(x(t_n),\epsilon) c_{r,2}(t_n) + \cdots \\ & +\epsilon^{r+1} c_{r+1,1}(t_n) + \epsilon^{r+2} s(x(t_n),\epsilon) c_{r+1,2}(t_n) + \cdots \\ & +\epsilon^{2r-1} c_{2r-1,1}(t_n) + O(\epsilon^{2r}). \end{aligned}$$

We now collect like powers of ϵ. At order ϵ^r we have a coefficient

$$s(x(t_n),0)^r \lambda_{r+1}(x(t_n)) + c_{r,1}(t_n)$$

or, by Lemma 7.1,

$$s^r(x(t_n),0)\lambda_{r+1}(x(t_n)) + \sigma_r(t_n) = 0, \tag{46}$$

so that the choice $\sigma_r(t) = s^r(x(t),0)\lambda_{r+1}(x(t))$ ensures that ${h_n}^{-1}\rho_n = O(\epsilon^{r+1})$. When σ_r has been determined, $c_{r,2}, c_{r,3}, \ldots$ become known functions of t, as in Lemma 7.1. At the next, ϵ^{r+1} order, we have a coefficient

$$\begin{aligned} & \frac{\partial}{\partial\epsilon} s^r(x(t_n),0)\lambda_{r+1}(x(t_n)) \\ & + s^{r+1}(x(t_n),0)\lambda_{r+2}(x(t_n)) + s(x(t_n),0)c_{r,2}(t_n) + \sigma_{r+1}(t_n) = 0, \end{aligned} \tag{47}$$

which determines σ_{r+1} and therefore the $c_{r+1,i}$. Iterating this procedure we determine all the required sources, and the existence of the asymptotic expansion is established.

Now suppose that the solution $x(\cdot)$ is T periodic. Clearly σ_r as given by (46) is also T periodic. By Lemma 7.1, the $c_{r,i}$ are also T periodic, so that (47) implies that σ_{r+1} is also T periodic, etc. This periodicity only holds for the lower sources σ_m, $r \le m \le 2r-1$.

8. Concluding remarks.

For constant stepsizes, the existence of the asymptotic expansion of the global error (12) is of course well known. A detailed classical treatment is given by Stetter [24]. A simpler proof is given in [17, Section II.8] following an idea of Hairer and Lubich. These known proofs may be extended to various variable stepsize algorithms. However, none of the available proofs meets our needs here. For us it is crucial to know that for periodic problems the sources σ_m, $r \le m \le 2r-1$ are periodic. Existing proofs write σ_m, $m > r$, in terms of the periodic local error contribution λ_{m+1} and of the earlier computed $e_r, \ldots, e_{m-1}$, and these are not periodic!

It is clear that the expansion (12) is not uniform in time, in the sense that $R(t, \epsilon)$ will in general grow with time. The techniques in this paper have enabled us to study the growth of the $e_m(t)$, $r \leq m \leq 2r - 1$, but give no indication of the growth of the reminder and therefore cannot settle the question of how the global error $x_n - x(t_n)$ behaves as t increases with fixed ϵ. This limitation is a direct consequence of the generality of our hypotheses. Throughout, we have only assumed a very small knowledge of the differential system: essentially we have only required information on the monodromy matrix of the orbit being integrated. It is clear that the behavior of the global errors is determined not only by the solution being integrated but also by the geometry of all solutions of the differential system, and therefore we do not really have enough information to estimate the whole global error $x_n - x(t_n)$. When this information is available it is possible to also estimate the remainder. This is done by Calvo and Hairer [6], but the price to pay is that they need to assume that any two solutions of the system deviate linearly from one another as t increases. Even stronger hypotheses are introduced in the work by Estep and Stuart [13], who demand that the problem be Hamiltonian and all solutions be periodic with a period depending only on the energy; an exceptional situation indeed.

The fact that no indication is given here as to the growth of the remainder does not at all imply that our results are irrelevant: we have ascertained the size of the errors as ϵ decreases with t large but fixed!

Acknowledgment. We are grateful to Professor R. D. Skeel for reading the manuscript.

REFERENCES

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

[2] V. I. Arnold and M. B. Sevryuk, *Oscillations and bifurcations in reversible systems,* in Nonlinear Phenomena in Plasma Physics and Hydrodynamics, R. Z. Sageev, ed., Mir, Moscow, 1986, pp. 31–64.

[3] W.-J. Beyn, *On invariant closed curves for one-step methods,* Numer. Math., 51 (1987), pp. 103–122.

[4] J. C. Butcher, *The Numerical Analysis of Ordinary Differential Equations,* Wiley, New York, 1987.

[5] J. C. Butcher, *Some orbital test problems,* Computing, 53 (1994), pp. 75–94.

[6] M. P. Calvo and E. Hairer, *Accurate long-term integration of dynamical systems,* Appl. Numer. Math., 18 (1995), pp. 95–105.

[7] M. P. Calvo and J. M. Sanz-Serna, *The development of variable-step symplectic integrators, with application to the two-body problem,* SIAM J. Sci. Comput., 14 (1993), pp. 936–952.

[8] P. Deuflhard and F. Borneman, *Numerische Mathematik* II*, Integration gewöhnlicher Differentialgleichungen,* Walter de Gruyter, Berlin, 1994.

[9] J. R. Dormand, M. E. A. El-Mikkawy, and P. J. Prince, *Families of Runge-Kutta-Nyström formulae,* IMA J. Numer. Anal., 7 (1987), pp. 235–250.

[10] T. Eirola, *Invariant curves of one-step methods,* BIT, 28 (1988), pp. 113–122.

[11] T. Eirola, *Two concepts for numerical periodic solutions of ODE's,* Appl. Math. Comput., 31 (1989), pp. 121–131.

[12] T. Eirola, *Aspects of backward error analysis for numerical ODEs,* J. Comput. Appl. Math., 45 (1993), pp. 65–73.

[13] D. J. Estep and A. M. Stuart, *The rate of error growth in Hamiltonian-conserving integrators,* Z. Angew. Math. Phys., 46 (1995), pp. 407–418.

[14] J. de Frutos and J. M. Sanz-Serna, *Erring and being conservative,* in Numerical Analysis 1993, D. F. Griffiths and G. A. Watson, eds., Longman Press, London, 1994, pp. 75–88.

[15] J. de Frutos and J. M. Sanz-Serna, *Accuracy and conservation properties in numerical integration: The case of the Korteweg-de Vries equation,* Numer. Math., 75 (1997), pp. 421–445.

[16] J. Guckenheimer and Ph. Holmes, *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields,* Springer, New York, 1983.

[17] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations* I*, Nonstiff Problems,* 2nd ed., Springer, Berlin, 1993.

[18] P. Hut, J. Makino, and S. McMillan, *Building a better leapfrog,* Astrophys. J., 443 (1995), pp. L93–L96.

[19] R. S. MacKay, *Some aspects of the dynamics and numerics of Hamiltonian systems,* in The Dynamics of Numerics and the Numerics of Dynamics, D. D. Broomhead and A. Iserles, eds., Clarendon Press, Oxford, 1992, pp. 137–193.

[20] D. Okunbor and R. D. Skeel, *An explicit Runge-Kutta-Nyström method is canonical if and only if its adjoint is explicit,* SIAM J. Numer. Anal., 29 (1992), pp. 521–527.

[21] A. Portillo and J. M. Sanz-Serna, *Lack of dissipativity is not symplecticness,* BIT, 35 (1995), pp. 269–276.

[22] J. M. Sanz-Serna and M. P. Calvo, *Numerical Hamiltonian Problems,* Chapman and Hall, London, 1994.

[23] M. B. Sevryuk, *Reversible Systems,* Lecture Notes in Math. No. 1211, Springer, Berlin, 1986.

[24] H. J. Stetter, *Analysis of Discretization Methods for Ordinary Differential Equations,* Springer, Berlin, 1973.

[25] D. Stoffer, *Some Geometric and Numerical Methods for Perturbed Integrable Systems,* Ph.D. thesis, Department of Mathematics, Eidgenoessische Technische Hochsshule (ETH), Zürich, 1988.

[26] D. Stoffer, *On Reversible and Canonical Integration Methods,* Res. report no. 88–05, Applied Mathematics, Eidgenoessische Technische Hochsshule (ETH), Zürich, 1988.

[27] D. Stoffer, *Variable steps for reversible integration methods,* Computing, 55 (1995), pp. 1–22.

[28] D. Stoffer and K. Nipp, *Invariant curves for variable step size integrators,* BIT, 31 (1991), pp. 169–180.