On the nonlinear stability of symplectic integrators

Robert I. McLachlan*   Matthew Perlmutter *

G. R. W. Quispel†

October 21, 2003

Abstract

The modified Hamiltonian is used to study the nonlinear stability of symplectic integrators, especially for nonlinear oscillators. We give conditions under which an initial condition on a compact energy surface will remain bounded for exponentially long times for sufficiently small time steps. While this is easy to achieve for non-critical energy surfaces, in some cases it can also be achieved for critical energy surfaces (those containing critical points of the Hamiltonian). For example, the implicit midpoint rule achieves this for the critical energy surface of the Hénon-Heiles system, while the leapfrog method does not. We construct explicit methods which are nonlinearly stable for all simple mechanical systems for exponentially long times. We also address questions of topological stability, finding conditions under which the original and modified energy surfaces are topologically equivalent.

1 Introduction

The general idea of stability of a numerical method for a differential equation is that, if the true solution does not tend to infinity, then the numerical solution does not tend to infinity. For Hamiltonian systems, one can say something about this simply by studying the energy surfaces. Orbits which lie on compact energy surfaces certainly do not tend to infinity. (These are the only ones considered in this paper.) Symplectic integrators do not preserve energy but do almost preserve a quantity (called the modified Hamiltonian) which is close, in some sense, to the original energy. Is this enough to ensure stability?

Let $(M, \{,\})$ be a symplectic manifold, let $H: M \to \mathbb{R}$ be an analytic Hamiltonian function, let $X_H = \{\cdot, H\}$ be its Hamiltonian vector field, and let φ be a symplectic integrator for X_H [12], i.e., a family of symplectic maps with the property that $\varphi(x, \tau) = x + \tau X_H(x) + \mathcal{O}(\tau^2)$; here τ is the time step. The integrator can be viewed as a flow in two ways: either as the exact time-τ flow of a nonautonomous Hamiltonian system, or as the approximate time-τ flow of an autonomous Hamiltonian system $\widetilde{H} = H + \mathcal{O}(\tau)$. In the second case the error in the approximation is exponentially small ($\mathcal{O}(e^{-c/\tau})$) in the time step. The function $\widetilde{H}: M \to \mathbb{R}$ is called the *modified*

*Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand (M.Perlmutter@massey.ac.nz).

†Department of Mathematics, La Trobe University, Bundoora, Melbourne 3083, Australia (R.Quispel@Latrobe.edu.au).

Hamiltonian. It depends on H, on the method φ, and on the time step τ and is usually found by expanding in Taylor series in τ to any desired order [3, 9]. We are primarily interested in the x-dependence of $\widetilde{H}$, and usually regard H, φ, and τ as fixed parameters.

Questions of convergence of the Taylor series of the modified vector field and computing the vector field to which it is asymptotic have been considered for a model nonlinear equation in [3]. Theory and applications for the Hamiltonian case can be found in [1, 8]. Since we are mainly interested in the behaviour of the method for small time steps—in fact properties governed by the first term in the perturbation—we here simply assume the existence of $\widetilde{H}$.

Definition 1 *For a function $H \colon M \to \mathbb{R}$, we write*

$$H_c := H^{-1}(c) = \{x \in M : H(x) = c\}$$

for the level set of H with value c. That is, $H_{H(x)}$ is the set of points whose energy is the same as that of the initial condition x; loosely, it is the energy surface through x.

It has long been realized [6, 13] that the existence of the modified Hamiltonian explains the good energy behaviour of symplectic integrators and their long-time stability. For, if the initial condition x is such that $H_{H(x)}$ and $\widetilde{H}_{\widetilde{H}(x)}$ are both compact, then the energy and global errors are necessarily bounded over exponentially long time intervals. Conversely, if $H_{H(x)}$ is compact but $\widetilde{H}_{\widetilde{H}(x)}$ is not, then it is entirely possible and indeed likely that the numerical orbit will not remain bounded. This is one way to view the loss of stability when integrating linear systems (see the example below) and we propose that it is a major mechanism for loss of stability when integrating nonlinear systems as well. (See also [14], in which nonlinear stability in the neighbourhood of fixed points is studied.) In this paper we explore the relationship between the topology of the true and the modified energy surfaces, and find conditions under which compactness, and hence nonlinear stability, can be guaranteed. (Our results must be interpreted modulo the exponentially small terms, which we shall ignore.)

Example 1 (The harmonic oscillator) *The standard leapfrog method applied to the harmonic oscillator in the plane ($H = (p^2 + q^2)/2$) gives a linear symplectic map of the plane,*

$$\begin{pmatrix} q \\ p \end{pmatrix} \mapsto \begin{pmatrix} 1 & \tau/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\tau & 1 \end{pmatrix} \begin{pmatrix} 1 & \tau/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} =: A(\tau) \begin{pmatrix} q \\ p \end{pmatrix}.$$

In backward error analysis, we seek a symmetric matrix $B(\tau)$ such that $A(\tau) = \exp(\tau J B(\tau))$, so that the map is the flow corresponding to the Hamiltonian

$$\widetilde{H} = \frac{1}{2}(q,p)B(\tau)(q,p)^T.$$

For $0 < \tau < 2$, B can be determined directly to be

$$B = \begin{pmatrix} \theta \csc \theta & 0 \\ 0 & \frac{1}{2}\theta \cot \frac{1}{2}\theta \end{pmatrix} = \begin{pmatrix} 1 - \frac{1}{12}\tau^2 - \frac{1}{120}\tau^4 - \cdots & 0 \\ 0 & 1 + \frac{1}{6}\tau^2 + \frac{1}{30}\tau^4 + \ldots \end{pmatrix}$$

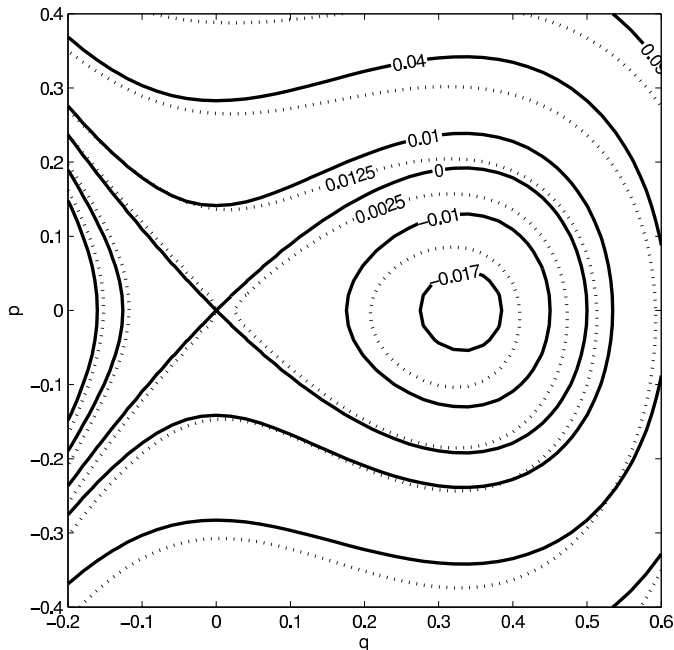*where*

$$\theta = 2\sin^{-1}\frac{\tau}{2}$$

Figure 1: Topological stability. The energy surfaces of the cubic oscillator with Hamiltonian $H = \frac{1}{2}p^2 - \frac{1}{2}q^2 + q^3$ are shown (solid lines), together with the energy surfaces of a small perturbation $\widetilde{H}$ of H (dotted lines).

and the eigenvalues of A are $e^{\pm i\theta}$. Because B is positive, the contours of $\widetilde{H}$ are ellipses, i.e., topologically equivalent to those of H.

For $\tau \geq 2$, the eigenvalues of $A(\tau)$ are nonpositive and it does not have a real logarithm, i.e., the map is not the flow of any linear Hamiltonian system. In fact, the orbits now lie on (noncompact) hyperbolas, which can be realized as the flow of a hyperbolic linear Hamiltonian system composed with the involution $(q, p) \mapsto (-q, -p)$. This change of topology at $\tau = 2$ corresponds to the loss of stability of the leapfrog method. Note that the Taylor series of $\widetilde{H}$ with respect to τ has radius of convergence 2.

Although many types of nonlinear stability can be considered [10, 15], we concentrate here on two, one topological and one quantitative. These are illustrated in the following example.

Example 2 (A nonlinear oscillator) *Consider the one degree-of-freedom nonlinear oscillator with Hamiltonian $H = \frac{1}{2}p^2 - \frac{1}{2}q^2 + q^3$, whose energy surfaces are shown in Fig. 1. In contrast to Example 1, it has both compact and non-compact energy surfaces. On a compact set (say the square shown in Fig. 1), the energy surfaces of $\widetilde{H}$ for a typical symplectic integrator may look like the dotted lines in Fig. 1. Although the energy surfaces have moved and changed values, the overall picture is the same, and we will say that in this case the integrator is* topologically stable. *Now consider fixing a particular initial condition x on the compact component of the critical energy surface (in this case, on the homoclinic orbit) and fixing the time step τ. The energy*
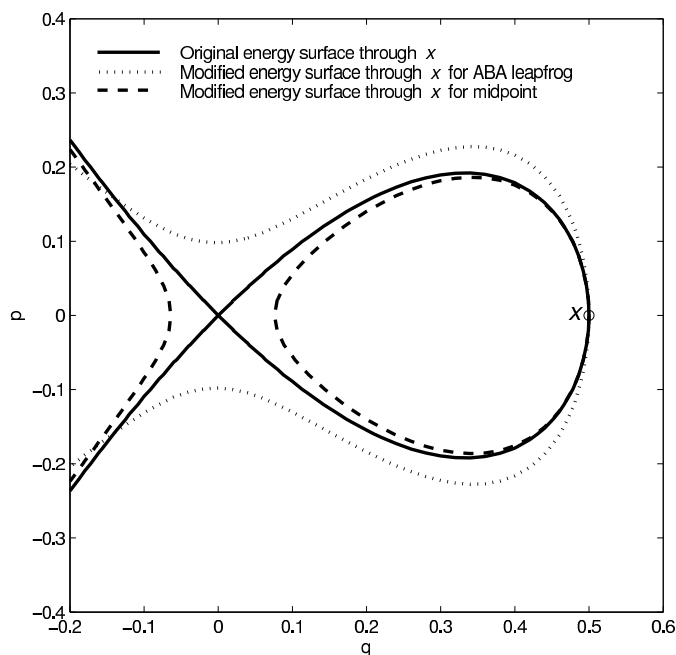
Figure 2: Nonlinear stability. The critical energy set of the cubic oscillator of Fig. 1 is shown together with the perturbations due to two different methods. The curves are drawn for time step $\tau = 1$, although they are qualitatively the same for all sufficiently small time steps.

surface of the modified Hamiltonian through x, $\widetilde{H}_{\widetilde{H}(x)}$, is shown in Fig. 2 (dashed line) for the symplectic midpoint rule: the energy surface is compact and we say the integrator is nonlinearly stable. In contrast, for the symplectic leapfrog method (Fig. 2, dotted line), the surface $\widetilde{H}_{\widetilde{H}(x)}$ is non-compact and the integrator is not nonlinearly stable. Furthermore, the situation of Fig. 2 persists no matter how small τ is.

Our work has been strongly motivated throughout by the Hénon-Heiles system, with phase space $M = \mathbb{R}^4$ and $H(q_1, q_2, p_1, p_2) = T(p) + V(q)$, $T(p) = \frac{1}{2}(p_1^2 + p_2^2)$, $V(q) = \frac{1}{2}(q_1^2 + q_2^2) + q_1^2 q_2 - \frac{1}{3} q_2^3$. The topology of the energy surfaces can be read off the contour plot of the potential energy V (Fig. 3): H_c has a compact component equivalent to a point for $c = 0$ and to S^3 for $0 < c \leq 1/6$, while all energy surfaces are noncompact for $c > 1/6$. We were particularly drawn to the critical energy surface $H_{1/6}$, which contains three fixed points and on which it seems that a tiny perturbation would be enough to push the orbit into the noncompact region, and hence blow up. What happens to these energy surfaces under the perturbation due to the symplectic integrator?

(This example makes it clear that for our purposes, different components of H_c should be considered as different energy surfaces. The noncompact part of H_0, for example, is irrelevant to what happens near the origin. The term 'compact' will therefore always refer in this paper to the relevant compact component.)

It might seem that attempting to integrate with initial conditions on H_c should be avoided because the problem is in a sense ill-posed. In fact, we were eventually
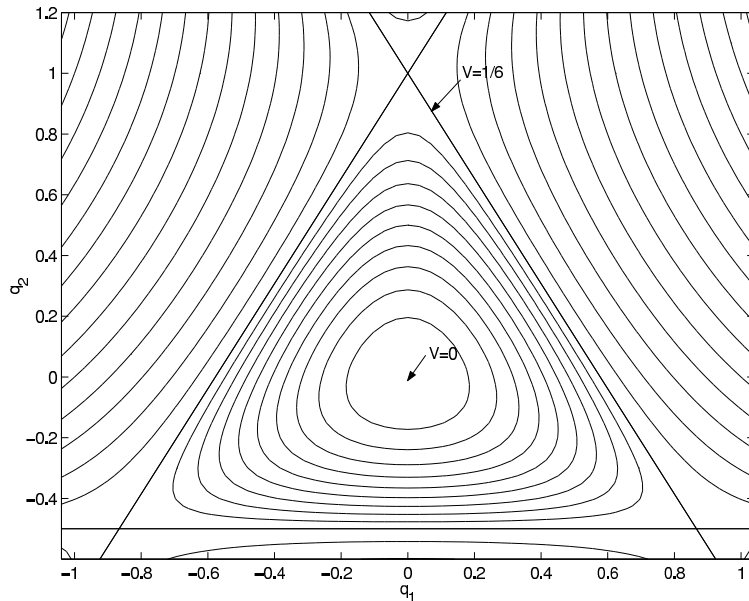
Figure 3: Hill regions (contours of potential energy) for the Hénon–Heiles system. An orbit with energy c must stay in the region $\{x : V(x) \leq c\}$. The 3 hyperbolic critical points of V (on $c = \frac{1}{6}$) correspond to fixed points with 2-dimensional center and 1-dimensional stable and unstable manifolds; the latter two coincide forming separatrices (e.g., $\frac{1}{2}p_2^2 + \frac{1}{2}q_2^2 - \frac{1}{3}q_2^3 = \frac{1}{6}$, $q_1 = p_1 = 0$), so all orbits on the compact component of the critical energy surface are in fact bounded. Orbits starting arbitrarily close to the critical contour become unbounded in a finite time.

led to believe that this is the most interesting case:

1. We show in Section 2 that both topological stability and nonlinear stability are easy to obtain when bounded away from critical energy surfaces. The critical energy surfaces are where this potentially breaks down, and we want to know precisely how this happens.

2. In numerical analysis one is often faced with a problem which is ill-posed as stated, but becomes well-posed when the problem class is restricted appropriately. In geometric integration this corresponds to imposing the appropriate constraint (e.g. symplectic, volume-preserving) on the integrator. Are critical energy surfaces of this nature?

3. One way to achieve the kind of stability we study here, even on a critical energy surface, is simply to impose energy conservation on the integrator. (This is commonly done in numerical PDEs, for example.) But this is a high price to pay, because the integrator cannot then be symplectic [12]. We want to know if the *kind* of energy conservation enjoyed by symplectic integrators is sufficient to ensure nonlinear stability as good as that enjoyed by energy-conserving integrators.

4. Suppose there is an $x \in H_c$ for which $\widetilde{H}_{\widetilde{H}(x)}$ is noncompact for all sufficiently

small time steps τ; the integrator is not nonlinearly stable. Then for initial conditions y close to x but lying on compact, non-critical energy surfaces, the integrator will suffer a stability restriction: there will exist $\tau^*(y)$ such that for $\tau > \tau^*(y)$, $\widetilde{H}_{\widetilde{H}(y)}$ is non-compact. In addition, $\tau^*(y) \to 0$ as $y \to x$. Studying the critical case also controls what happens in the well-posed but close-to-critical cases.

5. We will show that some methods are in fact nonlinearly stable while some are not. Considering the critical case thus helps one discriminate amongst otherwise equally good methods.

Section 2 studies topological stability. This approach was motivated by the idea that instability may be caused by a change in the topology of the level sets of $\widetilde{H}$. However, we show in Corollary 1 that for generic Hamiltonians, any symplectic integrator is topologically stable for sufficiently small time steps. This result does not apply to the Hénon-Heiles Hamiltonian because it is not generic, having more than one critical point on H_c. In this case not all symplectic integrators are topologically stable. Proposition 2 gives sufficient conditions, satisfied by some common methods, which ensure topological stability in this case. Thus topological stability is a practical criterion which allows one to discriminate amongst methods.

In a theoretical sense, topological stability is the end of the story. If a method is topologically stable, then given the initial condition x there exists y close to x such that the energy surfaces $\widetilde{H}_{\widetilde{H}(y)}$ and $H_{H(x)}$ are topologically equivalent. It would be a good idea to use y as an initial condition (similar to what is done in 'warming up' integrators in molecular dynamics). However, in practice one cannot find y. But one still wants to know what the integrator will do when given the initial condition x.

This is considered in Section 3, Nonlinear Stability. If an orbit of the real system is bounded, does the numerical orbit also remain bounded for exponentially long times for a given time step τ? Proposition 3 gives a criterion which can be used to determine this for all sufficiently small τ, given H and (the Taylor series in τ of) $\widetilde{H}$. As with topological stability, methods differ. We show that leapfrog can be unstable (there are systems with bounded orbits which blow up numerically for all $\tau > 0$); the midpoint rule is more stable; and we also construct an explicit symplectic integrator which is nonlinearly stable for all mechanical systems (in a sense made precise below).

2   Topological stability

Although we are primarily interested in compactness of energy surfaces, as this is related to stability, with Morse theory it is possible to decide much more, namely topological equivalence of the whole set of energy surfaces. If an energy surface changes topology under the perturbation due to the integrator, then the numerical solution cannot possibly be correct.

Definition 2 $H \in C^k(M, \mathbb{R})$ *(a C^k, real-valued function on M), $k \geq 2$, is a* Morse function *if its critical points $\{x : \mathrm{d}H(x) = 0\}$ are isolated (each is contained in a neighbourhood which contains no other critical points), nondegenerate (the Hessian Hess $H(x)$ is nonsingular), and if $H(x) \neq H(y)$ for all distinct pairs (x, y) of critical points (that is, each level set of H contains at most one critical point).*

Definition 3 *Two functions H and $\widetilde{H} \in C^k(M, \mathbb{R})$ are* topologically equivalent *if there are homeomorphisms Φ, Ψ such that $\Psi \circ \widetilde{H} = H \circ \Phi$.*

Looking at the level sets of two topologically equivalent functions, the level sets may have a different shape (determined by the homeomorphism Φ), and the level sets may have different values (determined by the homeomorphism Ψ), but still to each level set H_c there corresponds a level set $\widetilde{H}_{\tilde{c}(c)}$ with the same topology; furthermore the two foliations of M into level sets of H and $\widetilde{H}$ are topologically equivalent.

When the phase space M is not compact then the perturbation to H due to the integrator need not be bounded. But we are only interested in controlling the compact energy surfaces, and on these, the perturbation can be made arbitrarily small by choosing τ small enough.

Definition 4 *A symplectic integrator is* topologically stable *for time step τ on the compact set D, if $\widetilde{H}$ is topologically equivalent to H on D.*

It is useful to keep in mind the main result of Morse theory, that two level sets H_a and H_b of a Morse function H are topologically equivalent if H has no critical values in $[a, b]$. This easily extends to show that nearby Morse functions are equivalent (a local statement, which relies merely on transforming the functions to their normal form in each neighbourhood) which is the same as what we are calling topological stability of an integrator. For a general treatment of the topological stability of functions, and a proof of the following theorem, see du Plessis and Wall [5].

Theorem 1 *Let H be a C^k Morse function, and let $\widetilde{H}$ be sufficiently close to H in the C^k topology (defined by the ball $B(\varepsilon) = \{H \in C^k(M, \mathbb{R}) : |D^\alpha H(x)| < \varepsilon, \forall x \in M, \forall |\alpha| \leq k\}$). Then $\widetilde{H}$ is topologically equivalent to H.*

Corollary 1 *If the Hamiltonian H is a Morse function on a compact set, then any symplectic integrator is topologically stable for all sufficiently small time steps τ.*

This also extends to Hamiltonians with several critical points on the same energy surface, as in the Hénon–Heiles system.

Proposition 1 *If H is C^k on a compact set, with all critical points nondegenerate, and if for each pair of critical points of H with equal values the corresponding critical points of $\widetilde{H}$ have equal values for all sufficiently small time steps, then for all sufficiently small time steps the symplectic integrator is topologically stable.*

Proof   Theorem 1 is largely local; the required homeomorphisms can be chosen in neighbourhoods of critical and noncritical points and patched together. The crucial fact we need is that the homeomorphism Ψ (which adjusts the values of $\widetilde{H}$) can be chosen in a neighbourhood of a critical value to take the simple form $c \mapsto c + H(x) - \widetilde{H}(y)$ where x and y are nearby critical points of H and $\widetilde{H}$. Clearly, if two critical points of H with the same value require different adjustments then we will not be able to build a single homeomorphism Ψ. The hypothesis of the proposition ensures precisely that this does not happen and that the same map $c \mapsto c + H(x) - \widetilde{H}(y)$ aligns the values of all critical points with value $H(x)$. Away from critical points, the result is as for Theorem 1.   ●

Proposition 1 applies, for example, if the multiple critical points arise through a discrete symmetry (as in Hénon–Heiles) which is shared by the integrator. In fact, even this is not necessary, as the following Proposition shows.

Proposition 2 *For any symplectic Runge-Kutta or symplectic partitioned Runge-Kutta method, the critical points of H are also critical points of $\widetilde{H}$, with the same critical value. For any splitting method of any order given by the composition of the flows of X_{H_i} where $H = \sum H_i$, then any critical points of H which are also critical points of H_i for all i are also critical points of $\widetilde{H}$, with the same critical value.*

Proof　We regard Runge-Kutta methods as a special case of partitioned Runge-Kutta methods, for which the modified Hamiltonian $\widetilde{H}$ is a sum of elementary Hamiltonians, each of which is a derivative of H contracted with the q or p components of elementary differentials of X_H [7]. Apart from the zeroth order elementary differential, which is H itself, each term in $\widetilde{H}$ contains at least two factors H_q and/or H_p. Therefore if $H_q(x_0) = H_p(x_0) = 0$ then $H(x_0) = \widetilde{H}(x_0)$ and (by the product rule) $\widetilde{H}_q(x_0) = \widetilde{H}_p(x_0) = 0$.

For splitting methods, $\widetilde{H} - H$ is an element of L, the Lie algebra of functions generated by the H_i under the Poisson bracket. If $\nabla H_i(x_0) = 0$ for all i, we prove by induction that elements of L^n (the functions formed from n Poisson brackets) all have a critical point at x_0, whose value for $n > 0$ is zero.

For $n = 0$ the induction hypothesis is true by assumption. Suppose it is true for a certain value n. Let $\{H_i, K\}$, $K \in L^n$, be a basis element of L^{n+1}, giving $\{H_i, K\}(x_0) = 0$ immediately. On the other hand, its Poisson bracket with any other function F at the point x_0 is

$$d\{H_i, K\}.X_F(x_0) = \{F, \{H_i, K\}\}(x_0) = -\{H_i, \{K, F\}\}(x_0) - \{K, \{F, H_i\}\}(x_0).$$

The first term on the right is zero by assumption, while the second term is zero by the induction hypothesis. As F ranges over all functions, $X_F(x_0)$ ranges over all tangent vectors at x_0 (since M is symplectic), so $d\{H_i, K\}(x_0) = 0$, i.e., $\{H_i, K\}$ has a critical point at x_0. Therefore the hypothesis is true for L^{n+1}, which completes the proof. ●

Note that the result is not true for all symplectic integrators. For example, it is not true for splitting methods when critical points of H are not critical points of all the H_i. Such methods usually move the fixed points. In general there may also be spurious critical points of $\widetilde{H}$. This suggests studying the relationship between fixed points of the method and critical points of $\widetilde{H}$.

On $M = \mathbb{R}^n$ with H unbounded, $\widetilde{H}$ may have more critical points than H; for example, 'spurious' critical points could be born at infinity at $\tau = 0$ and move inwards as τ increases. For large enough τ they may collide with the non-spurious critical points, resulting in a bifurcation and change in topology of the modified energy surfaces.

However, we believe that for some methods this does not happen and is ruled out as a source of instability. For, if x^* is a critical point of $\widetilde{H}$, then

$$\varphi(x^*) - x^* = \varphi(x^*) - \exp(\tau X_\tau)(x^*) = \mathcal{O}(e^{-c/\tau}).$$

Under mild nondegeneracy conditions, either x^* is a fixed point of the method or there is a fixed point very (exponentially) close to x^*. In that case a method with no spurious fixed points has no spurious critical points of $\widetilde{H}$.

Now consider the leapfrog and midpoint methods. They do not have spurious fixed points [15], so $\widetilde{H}$ does not (typically) have spurious critical points. In addition, the nonspurious critical points do not move or change their value. Thus, $\widetilde{H}$ is then topologically equivalent to H on all (compact subsets of) phase space for sufficiently

small $\tau > 0$. This gives these methods a very strong topological stability and suggests that one possible instability mechanism—a change in the topology of $\widetilde{H}$ due to nonlinear effects—is not in fact significant. The only way the topology can change at larger time steps is by individual critical points changing their type. But near critical points, stability can be determined directly from linear stability theory.

For example, applying leapfrog to the Hénon-Heiles system, the fixed point at the origin is stable for $\tau < 2$, and the type of the fixed point at $p = (0,0)$, $q = (0,1)$ does not change for $\tau < 2/\sqrt{3} \approx 1.155$. For the midpoint rule, these fixed points have the same type for all $\tau > 0$. In practice, one finds that both the midpoint rule and leapfrog lose stability near $(q,p) = (0,1,0,0)$ at $\tau \approx 0.9$. This is likely to be due to nonautonomous effects—the 'exponentially small terms' are now large and any analysis involving $\widetilde{H}$ is irrelevant.

The results of this section apply to Hamiltonians with nondegenerate critical points. It would be interesting to know if a symplectic integrator can be topologically stable even when the Hamiltonian has degenerate critical points.

3   Nonlinear stability

In this section, we assume that the critical points of H are nondegenerate.

In spite of topological stability, if the initial condition x is such that $H(x)$ is close to a critical value, then the original energy surface through x and the modified energy surface through x may be inequivalent. Proposition 1 only guarantees that there is a *nearby* initial condition lying on an equivalent modified energy surface.

One common scenario for numerical integration is to integrate a family of initial conditions on a particular energy surface. This would be done when drawing Poincaré sections, for example. Therefore, we ask when all initial conditions on an energy surface are stable (i.e., bounded) during the numerical integration. This will be true (again, modulo exponentially small terms) when all the sets $\widetilde{H}_{\widetilde{H}(x)}$ are compact for all $x \in H_c$. This implies that when H and the integrator are analytic, the relevant orbits are bounded over exponentially long time intervals.

Definition 5 *A symplectic integrator is* nonlinearly stable at x *if $\widetilde{H}_{\widetilde{H}(x)}$ is compact. It is* nonlinearly stable *if it is nonlinearly stable at all x such that $H_{H(x)}$ is compact.*

Theorem 1 guarantees nonlinear stability for sufficiently small time steps in the neighbourhood of a compact energy surface H_c if c is not a critical value.

We therefore turn to the delicate case in which H_c divides compact and noncompact level sets. This turns out to depend essentially on the sign of the perturbation to the Hamiltonian. The following Proposition reduces the stability criterion to a simple calculation. First note that noncritical level sets carry a canonical volume form (say α, defined by $\alpha \wedge dH = \omega^n$, where ω is the symplectic form and $\dim M = 2n$) and hence are orientable. Their compact components have an inside and an outside. Because of the application to nonlinear oscillators, we shall study H in a neighbourhood of a minimum. For the critical set H_c itself, we define its 'inside' to be $\{x : H(x) < c\}$ and its 'outside' to be $\{x : H(x) > c\}$.

To be specific, we consider the case when the integrator is self-adjoint so that $\widetilde{H}$ contains only even powers of τ.

Proposition 3 *Let H_d be compact for $d \leq c$ and noncompact for $d > c$, and let $\widetilde{H} = H + \tau^2 H^{(2)} + \mathcal{O}(\tau^4)$.*

(i) If the critical points of H on H_c are also critical points of $\widetilde{H}$ with the same value, then all initial conditions on H_c lie on compact level sets of $\widetilde{H}$ for sufficiently small time steps τ if $d\widetilde{H}/d\tau \leq 0$ for sufficiently small τ for all $x \in H_c$. This is true if $H^{(2)}(x) < 0$ for all x in H_c which are not critical points, and false if there is an $x \in H_c$ with $H^{(2)}(x) > 0$.

(ii) If, on the other hand, the critical points of $\widetilde{H}$ can move and change their values as τ varies, then all initial conditions on H_c lie on compact level sets of $\widetilde{H}$ for sufficiently small time steps τ if (a) all critical points on H_c move outside H_c for $\tau > 0$ and (b) the maximum of $H^{(2)}$ on H_c is attained only at the critical points.

Proof   (i) By Proposition 1, the compact level sets of $\widetilde{H}$ are equivalent to those of H, and the only question is whether the unperturbed level set H_c intersects only compact level sets of $\widetilde{H}$. This is true iff $H_c \subset \{x : \widetilde{H}(x) \leq c\}$, that is, iff the critical level set $\widetilde{H}_c$ lies nowhere inside H_c. Since H and $\widetilde{H}$ are increasing in the outward direction, this is true iff $\widetilde{H}|_{H_c} \leq c$, which is true for sufficiently small time steps iff $\frac{\partial \widetilde{H}}{\partial \tau}|_{H_c} \leq 0$, which is true if $H^{(2)}|_{H_c} \leq 0$ and equals zero only at the critical points.

(ii) Now suppose the critical points move and change their values. Then $\widetilde{H}$ has several critical sets close to H_c. If any of the critical points lie inside H_c, then a noncompact level set of $\widetilde{H}$ will intersect H_c, indicating instability; hence (a) is necessary. Let the innermost of these critical sets contain $\tilde{x}(\tau)$, the critical point of $\widetilde{H}$ with the least critical value. Repeating the above calculation, we need

$$\widetilde{H}|_{H_c} \leq \widetilde{H}(\tilde{x})$$
$$\Leftrightarrow \frac{\partial \widetilde{H}}{\partial \tau}\bigg|_{H_c} \leq \frac{d\widetilde{H}(\tilde{x}, \tau)}{d\tau}$$
$$= \left\langle \nabla \widetilde{H}(\tilde{x}, \tau), \frac{d\tilde{x}}{d\tau} \right\rangle + \frac{\partial \widetilde{H}}{\partial \tau}(\tilde{x}, \tau)$$
$$= \frac{\partial \widetilde{H}}{\partial \tau}(x, \tau)$$

(the implications being true for all sufficiently small time steps). This gives the criterion that we have stability if $H^{(2)}$ attains its maximum value on H_c at the critical points, condition (b). ●

If $H^{(2)}|_{H_c}$ had any extra maxima with the same value, higher order terms would need to be examined as well.

Note that if the maximum values of $\widetilde{H}$ are distinct, then condition (b) cannot hold, so for this test to apply it is necessary for all critical values of $\widetilde{H}$ (corresponding to critical points on H_c) to be equal.

3.1   Stability of leapfrog and midpoint methods

It is impossible to ensure stability for completely general Hamiltonians. Surprisingly, however, it is possible to ensure stability for simple mechanical systems, even with an explicit method. Determining stability in any particular case is reduced by Proposition 3 to a calculus exercise.

We now specialize to the case of simple mechanical systems, i.e. $H = T(q,p) + V(q)$ with $T = \frac{1}{2}M(q)(p,p)$ and $M(q)$ positive definite. Then for any second order symplectic Runge-Kutta or splitting method,

$$H^{(2)} = \alpha\{T, \{T, V\}\} + \beta\{V, \{V, T\}\} = \alpha\widetilde{M}(q)(p,p) + \beta M(V', V')$$

for certain coefficients α and β, where the modified metric $\widetilde{M}$ is given by $\widetilde{M} = M^T V'' M + \frac{1}{2}M'MV'$. The coefficients α and β will play an important role in the sequel.

Denoting the method $\exp(\frac{\tau}{2}X_T)\exp(\tau X_V)\exp(\frac{\tau}{2}X_T)$ by 'ABA leapfrog', where $\exp(\tau Z)$ is the time-τ flow of the vector field Z, we get the following values for α and β:

ABA leapfrog: $\alpha = -\frac{1}{24}$, $\beta = \frac{1}{12}$

BAB leapfrog: $\alpha = \frac{1}{12}$, $\beta = -\frac{1}{24}$

Midpoint rule: $\alpha = -\frac{1}{24}$, $\beta = -\frac{1}{24}$.

Stability depends only on the angle of $(\alpha, \beta) \in \mathbb{R}^2$. Note that $M(V', V') \geq 0$, so the second term in $H^{(2)}$ is positive (causing instability) for $\beta > 0$ and negative (tending to help stability) for $\beta < 0$. We see immediately that the midpoint rule is "more nonlinearly stable" than the ABA leapfrog method. However, the most stable second order method in this sense would have $\alpha = 0$ and $\beta < 0$: we construct a composition method with this property below.

We need to evaluate $H^{(2)}$ on the energy surface H_c. We will consider Hamiltonians with $M = I$ (i.e. $T = \frac{1}{2}p^2$), various potentials $V(q)$, and various methods. We want to maximize $H^{(2)} = \alpha V''(p,p) + \beta|V'|^2$ subject to $\frac{1}{2}p^2 + V(q) = c$, where c is a critical value of H. For each q in the Hill region

$$R_c = \{q : V(q) \leq c\},$$

p lies on a sphere, and the first term in $H^{(2)}$ is maximized when p is an eigenvector corresponding to the maximum (for $\alpha > 0$) or minimum (for $\alpha < 0$) eigenvalue $\lambda(V''(q))$ of the symmetric matrix $V''(q)$. That is,

$$\max H^{(2)}|_{H_c} = \max_{q \in R_c} \Delta(q),$$

$$\Delta(q) := 2\alpha(c - V(q))\lambda(V''(q)) + \beta|V'(q)|^2.$$

Example 1. One degree of freedom.

For systems with one degree of freedom, $V''(q)$ is a 1×1 matrix, so $\lambda(V''(q)) = V''(q)$. Without loss of generality we can take $c = 0$ and $V(q) \sim -\frac{1}{2}q^2$ near a hyperbolic critical point. Then $\Delta \sim (\beta - \alpha)q^2$, so $\beta - \alpha \leq 0$ is necessary for stability. Thus the ABA leapfrog method is always unstable, while the midpoint rule (with $\alpha = \beta$) just passes this test (although it is unstable for the degenerate potentials $V(q) \sim -q^a$, $a > 2$, which require $\beta a \leq 2\alpha(a - 1)$).

Any single well potential will have an interior minimum of V, at which $\Delta = -2\alpha V(q)V''(q)$, so $\text{sgn}(\Delta) = \text{sgn}(\alpha)$, so $\alpha \leq 0$ is necessary for stability.

For cubic potentials, we can take $V(q) = -\frac{1}{2}q^2 + q^3$ and study Δ on the Hill region $[0, \frac{1}{2}]$. We have

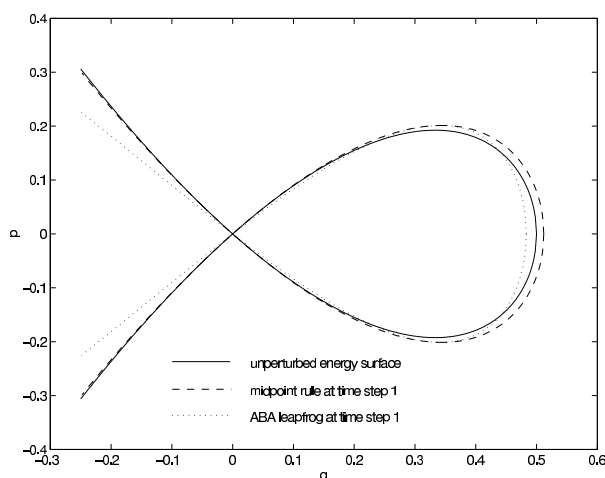$$\Delta = -q^2((12\alpha - 9\beta)q^2 + (6\beta - 8\alpha)q + \alpha - \beta)$$

Figure 4: The critical energy surface H_0 for a cubic oscillator $H = \frac{1}{2}p^2 - \frac{1}{2}q^2 + q^3$ is shown, together with its modification under two integrators at time step $\tau = 1$. It moves out under the midpoint rule, which is therefore stable, but in and out under the ABA leapfrog method: some initial conditions on H_0 lie outside the modified surface and hence can escape to infinity.

which one can show is negative on the Hill region iff $\beta - \alpha \leq 0$. Thus, the midpoint rule is stable for all cubic potentials. Figure 4 shows the original and perturbed critical energy surfaces in this case.

The midpoint rule is also stable for all quartics with no cubic term, but it is *not* stable for all quartics. It fails for the right-hand well of the double-well potential $V(q) = -\frac{1}{2}q^2 - q^3 + q^4$. To see this, note that for the midpoint rule, $\Delta' = \frac{1}{12}VV'''$, so there is a chance of instability if $V'''(q) = 0$ for some q in the Hill region. In this example this occurs at $q = \frac{1}{4}$, and here $\Delta(\frac{1}{4}) = \frac{5}{12288} > 0$ (just!).

The midpoint rule (and any method with $\alpha < 0$) is unstable if there are extra hyperbolic critical points of V *inside* the Hill region. For then $\Delta = -2\alpha V\lambda(V'') > 0$. It is interesting, though, that the midpoint rule is as stable as it is.

The condition $\beta \leq 2\alpha$ is also useful for stability. Note that

$$\frac{\Delta}{V^2} = -2\alpha(\ln|V|)'' + (\beta - 2\alpha)(\ln|V|)')^2$$

so that if $\beta \leq 2\alpha$ and $\alpha \leq 0$, the method is stable for all potentials for which $\ln|V|$ is convex, such as the pendulum $V(q) = 1 - \cos q$ and the quartic potential $V = -\frac{1}{2}q^2 - q^3 + q^4$ for which the midpoint rule failed.

Example 2. The two degree-of-freedom Hénon–Heiles system.

Similar calculations can be done in more dimensions. For the Hénon-Heiles potential $V(q_1, q_2) = \frac{1}{2}(q_1^2 + q_2^2) + q_1^2 q_2 - \frac{1}{3}q_2^3$ and the eigenvalues of V'' are $1 \pm 2\sqrt{q_1^2 + q_2^2}$. For the midpoint rule ($\alpha = \beta = -\frac{1}{24}$) for the critical value $c = \frac{1}{6}$ this gives

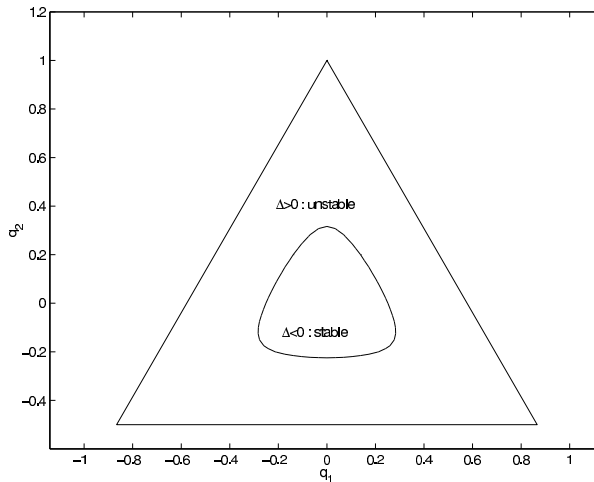$$24\Delta = -(1 - 2\sqrt{q_1^2 + q_2^2})(\frac{1}{3} - 2V) - |V'|^2$$

Figure 5: Stability region for ABA leapfrog applied to the Hénon-Heiles system. Initial conditions with q in the region shown and energy $H = \frac{1}{6}$ will be bounded for exponentially long time intervals.

which has exactly 7 critical points: 4 at the critical points of V (maxima with value 0 at $(q_1, q_2) = (0, 1)$ and its symmetric images, and a minimum with value $-\frac{1}{3}$ at $(0, 0)$) and three with value ≈ -0.2270 at $(q_1, q_2) = (0, (1 - \sqrt{57})/28)$ and its symmetric images. On the boundary $q_2 = -\frac{1}{2}$ we have $24\Delta = -(q_1^2 - \frac{3}{4})^2 \leq 0$. Therefore (using the threefold symmetry), $\Delta \leq 0$ on the unperturbed Hill region and the midpoint rule is stable. The ABA leapfrog method is stable for small $|q|$ (the precise region is shown in Figure 5) and unstable for large $|q|$, while the BAB leapfrog method is stable only for large enough $|q|$ (Figure 6). This prediction is confirmed numerically in Figure 7: initial conditions on the critical energy set $H_{1/6}$ with $p = 0$ blow up even for small time steps, while initial conditions with $q = 0$ (inside the safe region of Figure 5) are bounded even for quite large time steps.

If a different splitting is used, leapfrog can be stable. Consider the popular linear plus nonlinear splitting, $A = (|q|^2 + |p|^2)/2$, $B = q_1^2 q_2 - \frac{1}{3} q_2^3$. Then composition methods will in general move the fixed points, and we have to check that they move outwards. By symmetry, they will all retain the same critical value; one can calculate from $H^{(2)}$ that the critical point initially at $p = 0, q = (0, 1)$ moves to $p = 0, q = (0, 1 + (3\beta + 4\alpha)\tau^2 + \mathcal{O}(\tau^4))$, i.e., it moves outwards for $3\beta + 4\alpha > 0$. Both ABA and BAB leapfrog satisfy this. Secondly, one has to check that $H^{(2)}$ is maximized at the critical points. This also turns out to be the case for both ABA and BAB leapfrog. Therefore, both these methods are stable. In a sense, allowing the critical points to move out has actually made it easier for the entire level set to move out, ensuring stability.

3.2   Visualizing the modified energy surfaces

For simple mechanical systems the modified Hill region (the projection of an energy surface to the configuration (q) space) gives some idea of the perturbation to the
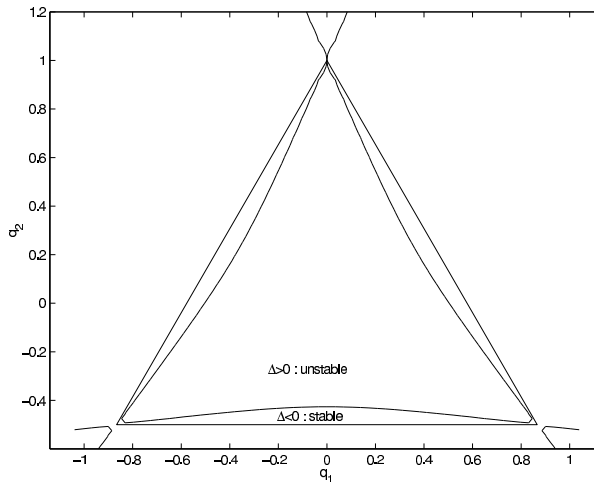
Figure 6: Stability region for BAB leapfrog applied to the Hénon-Heiles system. This method has $\beta < 0$, helping it near critical points of V on the critical level set, but $\alpha > 0$, which hurts it elsewhere. Overall it is less stable than ABA leapfrog.

energy surfaces. Considering only the first term $H^{(2)}$ we have

$$
\begin{aligned}
c &= H + \tau^2 H^{(2)} \\
&= T(p) + V(q) + \tau^2\big(\alpha\widetilde{M}(p,p) + \beta M(V',V')\big). \\
&= \frac{1}{2}(M + 2\alpha\tau^2\widetilde{M})(p,p) + V + \beta\tau^2 M(V',V')
\end{aligned}
$$

For sufficiently small τ, the modified metric $M + 2\alpha\tau^2\widetilde{M}$ is still positive definite (provided $\widetilde{M}$ is bounded) so the modified Hill region is given by

$$\{q : V + \beta\tau^2 M(V',V') \le c\}$$

which is easy to draw. This immediately indicates whether the modified energy surface moves in or out as τ increases from zero, at least at $p = 0$. For $p \ne 0$ the modified energy surface consists of an ellipsoid attached to each point q in the Hill region, instead of a sphere.

Figure 8 shows the path in configuration space for the ABA leapfrog method with initial condition $q = (0,0)$, $p = (1/\sqrt{6}, 1/\sqrt{6})$, at the center of the stable region shown in Figure 5. The time step is $\tau = 1$. The modified critical Hill region lies inside the unmodified Hill region, so that initial conditions lying between them are unstable. However, for this initial condition, $\widetilde{H} = \frac{1}{6}(1 + 2\alpha\tau^2) < \frac{1}{6}$, so the computation is stable, and remains within the modified Hill region shown.

3.3   Methods nonlinearly stable for all potentials

A method with $\alpha = 0$ and $\beta < 0$ will be stable for all potentials $V(q)$ (with the proviso that the higher-order terms in $\widetilde{H}$ become relevant at critical points of V in the interior of the Hill region). A method with $\alpha \le 0$ and $\beta \le 2\alpha$ will be stable for
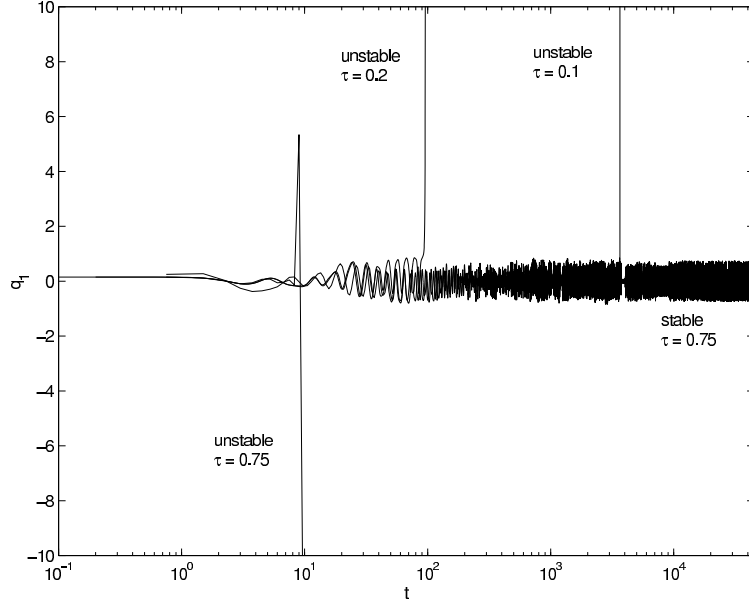
Figure 7: ABA leapfrog with 2 different initial conditions and different time steps. The 'stable' initial condition is $q = 0$, $p = (\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}})$, at the center of Fig. 5. The 'unstable' initial condition is $q = (0.1, -0.5)$, $p = 0$, on the base of the triangle in Fig. 5.

many single-well potentials. It is possible to design composition methods with these properties.

A first attempt, increasing the total number of A and B stages from 3 to 5, ends in a failure reminiscent of an order barrier. Letting $X_H = A + B$, we consider the method

$$e^{a\tau A}e^{\frac{1}{2}\tau B}e^{(1-2a)\tau A}e^{\frac{1}{2}\tau B}e^{a\tau A} \tag{1}$$

and compute using the BCH formula that

$$\alpha = \frac{1}{12}(1 - 6a + 6a^2), \quad \beta = \frac{1}{24}(6a - 1).$$

So $\beta < 0$ requires $a < \frac{1}{6}$, but $\alpha = 0$ only at $a = \frac{1}{2} \pm \frac{\sqrt{3}}{6}$. Eliminating a, we see that only methods with $\alpha = 8\beta^2 - \frac{4}{3}\beta + \frac{1}{72}$ are possible. Swapping the positions of A and B in (1) gives methods with $\beta = 8\alpha^2 - \frac{4}{3}\alpha + \frac{1}{72}$. These two curves are shown in Figure 9—they do not come anywhere near the desired stability region.

Including a corrector [2] of the form $\exp(\tau^2\gamma[A,B])$ shifts the parameters from (α, β) to $(\alpha + \gamma, \beta - \gamma)$. Choosing $\gamma = -\alpha$ we get $\beta' = \beta + \alpha$. So a method can be corrected to be nonlinearly stable for all potentials if $\beta + \alpha < 0$. The midpoint rule satisfies this (see Figure 9), but none of the 5-stage composition methods do.

Therefore, more stages are required, so we try 7 (4 As and 3 Bs):

$$e^{a\tau A}e^{b\tau B}e^{(1/2-a)\tau A}e^{(1-2b)\tau B}e^{(1/2-a)\tau A}e^{b\tau B}e^{a\tau A}. \tag{2}$$

We find

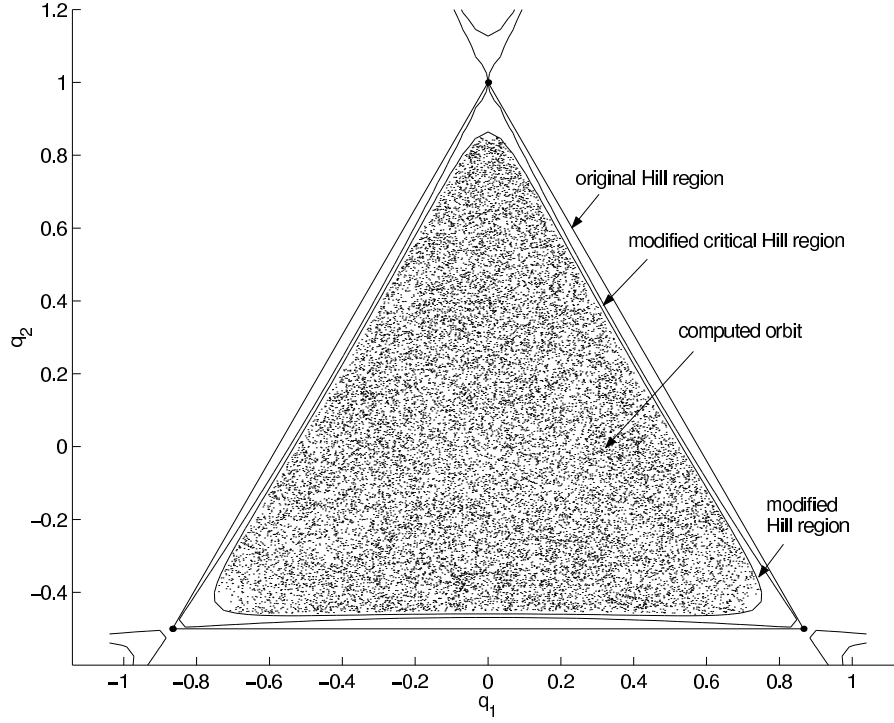$$\alpha = -\frac{1}{24} + b(\frac{1}{4} - a + a^2), \quad \beta = \frac{1}{12} + \frac{1}{2}b(1-2a)(b-1)$$

Figure 8: A stable orbit for the leapfrog method.

so that $\alpha = 0$ gives $b = \frac{1}{6}(2a-1)^{-2}$ and

$$\beta = \frac{48a^3 - 48a^2 - 12a - 1}{72(2a-1)^3}.$$

The root $a = a^* := \frac{1}{2}(2 - 2^{1/3})^{-1} \approx 0.675$ is the familiar fourth-order method, and the range $a \in (\frac{1}{2}, a^*)$ gives stable methods. One convenient solution is

$$e^{\frac{2}{3}\tau A} e^{\frac{3}{2}\tau B} e^{-\frac{1}{6}\tau A} e^{-2\tau B} e^{-\frac{1}{6}\tau A} e^{\frac{3}{2}\tau B} e^{\frac{2}{3}\tau A},$$

for which $\beta = -\frac{1}{24}$, a respectable value compared with leapfrog. No doubt more efficient stable methods with more stages also exist.

3.4 Higher order methods

The methods of this paper can be used to check the stability of any method, or to determine the coefficients that the terms in $\widetilde{H}$ would need to be stable for a given potential. At fourth order, the terms in $H^{(4)}$ for $H = \frac{1}{2}p^2 + V(q)$ are

$$V''''(p,p,p,p), \; V'''(V',p,p), \; V''(V',V'), \text{ and } |V''(p)|^2.$$

Amongst these, the sign of only the last can be controlled for any reasonably general class of potentials. This suggests that the stability situation may be analogous to
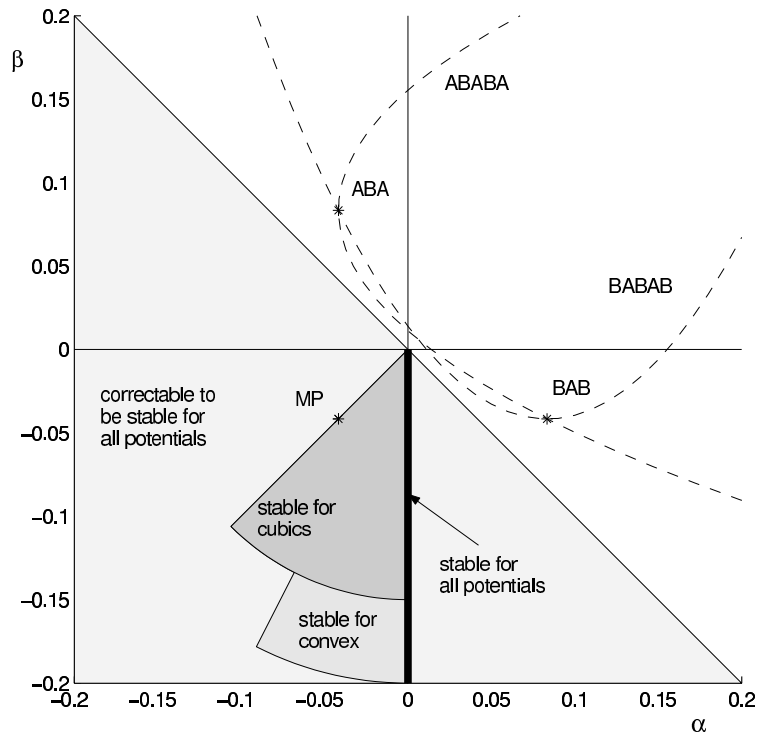
Figure 9: Stability parameters of various methods. The principal error coefficients α and β appearing in the modified Hamiltonian $\widetilde{H}$ are shown for 2 3-stage composition methods (ABA and BAB leapfrog), 2 families of 5-stage composition methods, and the midpoint rule. Also shown are four sectors in which a method would be stable for four classes of problems.

the one at second order but slightly harder to ensure. This pattern persists at higher order: at sixth order, amongst the 6 terms in $H^{(6)}$ we find the two positive terms $|V''(V')|^2$ and $|V'''(p,p)|^2$. We have not yet done further calculations for these cases, however.

Acknowledgements

References

[1] G. Benettin and A. Giorgilli, On the Hamiltonian interpolation of near to the identity symplectic mappings with application to symplectic integration algorithms, *J. Stat. Phys.* **74** (1994), 1117–1143.

[2] S. Blanes, F. Casas, and J. Ros, Symplectic integration with processing: a general study, *SIAM J. Sci. Comput.* **21** (1999), 711-727.

[3] Jonathan M. Borwein and Robert M. Corless, Emerging tools for experimental mathematics, *Am. Math. Monthly* **106**(10) (1999), 889-909.

[4] Th. Bröcker, *Differential Germs and Catastrophes*, LMS Lect. Notes **17**, Cambridge University Press, Cambridge, 1975.

[5] Andrew du Plessis and Terry Wall, *The Geometry of Topological Stability*, Clarendon, Oxford, 1995.

[6] K. Feng, Formal power series and numerical algorithms for dynamical systems, *Proceedings of International Conference on Scientific Computation (Hangzhou, 1991)*, 28–35, Ser. Appl. Math., 1, World Scientific, River Edge, NJ, 1992.

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

[8] E. Hairer and Ch. Lubich, The life-span of backward error analysis for numerical integrators, *Numer. Math.* **76**(4) (1997), 441–462.

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

[10] A. M. Lyapunov, *The general problem of the stability of motion*, Taylor & Francis, London, 1992.

[11] J. Milnor, *Morse Theory*, Annals of Mathematics Studies 51, Princeton University Press, Princeton, 1963.

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

[13] Robert D. Skeel and David J. Hardy, Practical construction of modified Hamiltonians, *SIAM J. Sci. Comput.* **23 (2001)**, 1172–1188.

[14] R. D. Skeel and K. Srinivas, Nonlinear stability analysis of area-preserving integrators, *SIAM J. Numer. Anal.* **38** (2000), no. 1, 129–148.

[15] A. Stuart and A. Humphries, *The Numerical Analysis of Dynamical Systems*, Cambridge University Press, Cambridge, 1996.