On the KAM and Nekhoroshev theorems for symplectic integrators and implications for error growth.

Per Christian Moan*

December 20, 2002

Abstract

In this study an alternative to backward error analysis (BEA)for geometric numerical integration schemes is introduced. It allows us to prove KAM and Nekhoroshev stability theorems for symplectic discretizations of close to integrable Hamiltonian systems, without the restrictions of traditional BEA. The results are qualitative in nature and explain in further detail why symplectic integration of Hamiltonian systems is so successful.

Key words: Numerical integration, Symplectic, KAM, Nekhoroshev theory, error growth.

1 Introduction

Symplectic integration (SI) of Hamiltonian differential equations has received a lot of attention in the last few decades. For long time-span simulations, such as those of celestial mechanics, molecular dynamics and accelerator dynamics SI-schemes have shown outstanding performance, not seen in traditional discretization schemes.

By retaining the geometric properties of symplectic flows, SI often leads to improved error growth and retention of global qualitative properties of the exact flow. The technique of BEA has been used to analyze and explain this behavior.

Suppose we are given an analytic[1] Hamiltonian function $h : \mathbb{R}^{2n} \mapsto \mathbb{R}$. If the flow of h is discretized by a symplectic integration scheme, $\Psi_{\Delta t,h}$, then BEA tells us [3, 40, 17, 39, 16, 18] that there exists a formal autonomous Hamiltonian function $\overline{h}$ so that its time-Δt flow, $\phi_{\Delta t,\overline{h}}$, is equal to $\Psi_{\Delta t,h}$ up to all algebraic orders in Δt. The series defining $\overline{h}$ is generally divergent, but very accurate approximations, $\overline{h}^*_{\Delta t}$, are possible in the limit $\Delta t \to 0$. For the general case, when no assumptions on the dynamics of h are made, one can prove the *worst case bound*

$$\Psi_{\Delta t,h}(x) = \phi_{\Delta t,\overline{h}^*}(x) + \mathcal{O}(\exp(-c_0 r/\Delta t\|\partial h\|_r)), \quad \text{for } |\Delta t| < \Delta t^*, \tag{1}$$

where the positive constant $c_0 \leq 2\pi$ depends on the method $\Psi_{\Delta t,h}$ but not on Δt. The norm $\|\cdot\|_r$ is defined as the supremum of the individual components of the vector $\partial_I h_0$ on complex tubular neighbourhood of radius $r > 0$ of the trajectory.

*School Of Mathematical and Statistical Sciences, La Trobe University, VIC 3086, Australia. email:p.moan@latrobe.edu.au.

[1] By "analytic" we will mean *real* analytic, i.e. an analytic function that is real for real arguments.

When the existence of an optimally truncated $\overline{h}^*_{\Delta t}$ has been established, stability results from the theory of differential equations can be applied to $\overline{h}^*$, and in this manner the stability properties of $\Psi_{\Delta t,h}$ up to the effects of the exponentially small remainder term are found. This idea was explored in different contexts in [17, 39, 19] giving various stability estimates for symplectic integrators. The most immediate consequence is that $\overline{h}^*$ is a virtually conserved quantity restricting the dynamics produced by the numerical scheme to a $2n-1$ dimensional subspace of $\mathbb{R}^{2n}$.

When applying BEA to real calculations it becomes apparent that the bound Δt^* is highly conservative, and that numerical experiments with SI often performe better than this theory predicts. Indeed if $\Delta t \|\partial h\|_r > 2\pi r$ the exponential estimate can no longer be considered small, but experiments indicate good stability even then. Even for small Δt the estimate (1) must be viewed as conservative.

Overall, one can say that the strength of BEA is that one can always construct the Hamiltonian $\overline{h}^*_{\Delta t}$ under mild assumptions on the flow of h but quite restrictive assumptions on Δt, and the implications are often less than optimal.

1.0.1 Close to integrable systems

A more careful analysis of close-to-integrable Hamiltonian systems ($h = h_0(I) + \epsilon r_1(I,\theta)$) shows that the exponential remainder term in (1) originates from step-size resonances, i.e. when some oscillatory component of the exact solution has period equal to the step size Δt [31]. Therefore to improve the stability estimates of BEA, one can make more assumptions on the step size, Δt or on the unperturbed Hamiltonian h_0 so that such resonances or their effects are avoided [48, 18, 30]. Along these lines a KAM theorem for SI-schemes by Shang [47], see also [18, 43], shows that for Kolmogorov non-degenerate Hamiltonians and step sizes belonging to a certain Cantor set, strongly non-resonant orbits are essentially preserved, and therefore the stability in this case is not subject to the exponential estimate (1). Further evidence for the importance of non-resonance conditions is given in [52] where Wisdom and Holman give an alternative analysis of the effect of step size resonances and show examples where stability is actually lost near the resonant step sizes.

In this paper we introduce a different approach to backward error analysis which overcomes the shortcomings of traditional BEA, and use it to study the KAM stability (infinite time stability) [23] and Nekhoroshev stability (exponentially long stability times) [33] of SI-schemes. A similar technique was used by Douady [8] to prove the first general KAM theorems for symplectic mappings close to the identity. The Nekhoroshev stability results are particularly useful for explaining the superior approximations produced by SI-schemes as they avoid the impossible-to-check non-resonance conditions of KAM theory and allow initial values in open subsets of phase space. An improved stability result for orbits close to KAM-tori derived by Morbidelli and Giorgilli [32] can also be derived for SI schemes giving a super-exponential stability result for simulations close to strongly-non resonant orbits.

1.1 Set-up

We consider discretizations of the flow of perturbed integrable Hamiltonian systems. For the analysis we assume the Hamiltonian has been put into the form $h(I,\theta) = h_0(I) + \epsilon r_1(I,\theta)$, and is analytic in the action-angle variables of the unperturbed system $(I,\theta) \in \mathcal{D} \times \mathbb{T}^n$, $\mathcal{D} \subset \mathbb{R}^n$ is an open set and $\mathbb{T} = \mathbb{R}/\mathbb{Z} = \{\theta (\mathrm{mod} 1) : \theta \in \mathbb{R}\}$.

The corresponding differential equations are then

$$\begin{aligned} I' &= -\partial_\theta \epsilon r_1(I,\theta) \\ \theta' &= \partial_I h_0(I) + \partial_I \epsilon r_1(I,\theta) \end{aligned} \tag{2}$$

The perturbation ϵr_1 is assumed small and h_0 will be assumed to satisfy various non-degeneracy conditions so that the exact flow $\phi_{t,h}$ is stable. We say that the flow $\phi_{t,h}$ is T-stable if the following bound holds

$$\sup_{t\in[0,T]} \|I(t) - I(0)\| = R(\|\epsilon r_1\|), \text{ where } R = \mathcal{O}(\|\epsilon r_1\|^b) \tag{3}$$

for some constant $b > 0$[2]. If one can conclude that $T = \infty$ then the flow is stable in the classical sense. However for applications with non-degenerate conservative systems, finite but large T can be useful, in particular exponential bounds, $T = \mathcal{O}(\exp(C\|\epsilon r_1\|^{-a}))$, where $a > 1$.

1.1.1 The numerical method

The exact flow of a Hamiltonian system is well known to be symplectic, i.e. preserves the two-form $\omega = \sum_{i=1}^n dI \wedge d\theta$, $(\phi_{t,h})_*\omega = \omega$. The discretization of the flow $\phi_{t,h}$ is obtained by iterating a map $\Psi_{\Delta t,h}$ which we assume is analytic, i.e. it can be expanded as a convergent Taylor expansion in the step size Δt around $\Delta t = 0$, and is consistent to order p, $\Psi_{\Delta t,h} = \phi_{\Delta t,h} + \mathcal{O}(\Delta t^{p+1})$. We also assume, like for the exact flow, that the map $\Psi_{\Delta t,h}$ is symplectic, $(\Psi_{\Delta t,h})_*\omega = \omega$. We note this assumption can be weakened to assuming that $\Psi_{\Delta t,h}$ is conjugate to a symplectic mapping, i.e. there exists a diffeomorphism, χ, on phase space so that $\chi_{\Delta t,h} \circ \Psi_{\Delta t,h} \circ \chi^{-1}_{\Delta t,h} = \varphi_{\Delta t,h}$ where φ is a canonical symplectic map. Alternatively, one may say that the non-canonical symplectic two form $\omega_* = (\chi_{\Delta t,h})_*\omega$ is left invariant by $\Psi_{\Delta t,h}$. Methods of this type include operator splitting methods such as the Störmer/Verlet method, methods of Runge-Kutta type such as the trapezoidal rule and midpoint rule together with the Gauss methods of arbitrary high order[18] and others.

We also note that the assumption that the Hamiltonian is written in action-angle coordinates is not necessary, for the numerical discretization, we merely rely on the fact that it is possible to write it as above for the stability analysis.

1.2 Non-autonomous backward error analysis

As an alternative to BEA, *non-autonomous backward error analysis* (na-BEA) represents an intermediate step in the stability analysis. By not carrying out the time-averaging[**?**] the exponentially small remainder term are avoided, yet it is sufficient for the application of standard stability results from Hamiltonian perturbation theory. The idea of na-BEA is to construct a non-autonomous perturbation $\epsilon r_2(I,\theta,t/\Delta t)$, Δt-periodic in t so that the Δt-flow map of $h + \epsilon r_2$ is equal to $\Psi_{\Delta t,h}$, i.e.

$$\phi_{\Delta t,h+\epsilon r_2} = \Psi_{\Delta t,h}.$$

[2]The norm of a function is the sup-norm in some complex neighbourhood of $\mathcal{D} \times \mathbb{T}^n$, while that of vectors is the Euclidean norm.

A simple construction of ϵr_2 is found by letting H and ϵR_2 be the Hamiltonian vector fields corresponding to h and ϵr_2 respectively. Then we may define

$$\epsilon R_2 := \frac{\partial \Psi_{t,h}}{\partial t} \circ \Psi_{t,h}^{-1} - H, \tag{4}$$

provided $\Psi_{t,h}$ is an invertible mapping on $\mathcal{D} \times \mathbb{T}^n$ for $t \in [0, \Delta t]$. Enforcing the Δt-periodicity in t gives a Hamiltonian $h + \epsilon r_2(I, \theta, t/\Delta t)$ whose flow exactly interpolates the trajectory of $\Psi_{\Delta t,h}$. A stability theory for SI schemes was derived for this setting in [30], where questions related to the invertibility of $\Psi_{t,h}$ were also addressed. See also the work by Fiedler and Scheurle[12] for other results along these lines.

This simple construction is not sufficient as ϵr_2 is not analytic (or even $C^1([0, \Delta t])$ at $t = 0$ and $t = \Delta t$, which makes the application of fast-converging Fourier series difficult and hence it reduces the usefulness of traditional perturbation techniques. In particular if one is interested in applying results from Hamiltonian perturbation theory, more (e.g. C^k with $k > 2n$) regularity is necessary for KAM-type results [1, 46]. For exponential stability estimates such as the Nekhoroshev estimates [33] it will be necessary that $h + \epsilon r_2$ is analytic in some complex neighborhood of $(I, \theta, t/\Delta t) \in \mathcal{D} \times \mathbb{T}^n \times \mathbb{T}$. The existence of such an analytic perturbation $\epsilon r_2(I, \theta, t/\Delta t)$ is the theme of a paper by Kuksin and Pöschel [25] see also [20, 8]. From [25] the following result follows directly

Theorem 1 *Suppose $\Psi_{t,h}$ is an analytic[3], symplectic numerical method approximating the flow $\phi_{t,h}$, then for sufficiently small ϵ there exists a Δt-periodic perturbation $\epsilon r_2(I, \theta, t/\Delta t)$ so that $\phi_{\Delta t, h+\epsilon r_2} = \Psi_{\Delta t,h}$. Moreover $h + \epsilon r_2$ is analytic in some open neighborhood of $\mathcal{D} \times \mathbb{T}^n \times \mathbb{T}$ and $\epsilon r_2 \to 0$ when $\epsilon r_1 \to 0$.*

Theorem 1 replaces the role of BEA for close-to-integrable Hamiltonian systems, while a result by Pronin and Treschev [49] and Trifonov [50] generalizes the Theorem 1 to all analytic mappings. If standard BEA is desired, then e.g. the time-averaging procedure of Neĭshtadt [34] can be applied to $h + \epsilon r_2$ to remove the time-dependency up to an exponentially small remainder term.

We note that the proof by Kuksin and Pöschel is an existence proof, and it is not clear from their technique how to derive a rigorous bound on the magnitude of ϵr_2. If rigorous bounds are desired the approach by Pronin and Treschev [49] is possibly more useful.

1.2.1 The extended Hamiltonian

Instead of carrying out a time-averaging on $h + \epsilon r_2$ reducing the dynamics to $\mathbb{R}^n \times \mathbb{T}^n$ we now introduce the variable $e \in \mathbb{R}$ which is canonically conjugate to $\tau = t/\Delta t$. In this way we are led to consider the *extended* Hamiltonian

$$\hat{h} := h_0(I) + e/\Delta t + \epsilon r_1(I, \theta) + \epsilon r_2(I, \theta, \tau),$$

on the phase space $(I, \theta, \tau, e) \in \mathcal{D} \times \mathbb{T}^n \times \mathbb{T} \times \mathbb{R}$, with a symplectic two form is given by $\hat{\omega} := \omega + de \wedge d\tau$, left invariant by the flow $\phi_{t,\hat{h}}$, $(\phi_{t,\hat{h}})_* \hat{\omega} = \hat{\omega}$. For the extended Hamiltonian the unperturbed part is $\hat{h}_0 := h_0(I) + e/\Delta t$, while the perturbation is taken as $\epsilon \hat{r}_1 = \epsilon r_1 + \epsilon r_2$.

[3] We say the numerical method Ψ is analytic if when applied to analytic Hamiltonians it produces an analytic mapping in phase space.

The application of stability results from Hamiltonian perturbation theory is now a matter of verifying that the assumptions that hold for the original unperturbed Hamiltonian $h_0(I)$ still hold for the extended unperturbed Hamiltonian $\hat{h}_0(I,e)$. Any discrepancies will then indicate how the phase portrait of the approximation $\Psi_{\Delta t,h}$ differs qualitatively from that of the exact flow, $\phi_{\Delta t,h}$. In other words, the results indicate when numerical discretization introduces instabilities not present in the original system. As the results we are studying are perturbative we will assume throughout that the perturbations $\epsilon\hat{r}_1 \approx \epsilon r_1$ are "sufficiently small".

2 KAM theorems for symplectic integrators

KAM theorems are the best known stability theorems of close-to-integrable Hamiltonian systems of ODEs. They are perturbative in nature, and give the existence and persistence of invariant toral submanifolds in phase space. In other words they give the existence of a symplectic coordinate transform $(I,\theta) = \Phi(\tilde{I},\tilde{\theta})$ so that the perturbed Hamiltonian $h_0(I) + \epsilon r_1(I,\theta)$ in the new coordinates takes the form $\tilde{h}_0(\tilde{I})$. Shang [47] proved a KAM theorem for numerical discretizations of Kolmogorov non-degenerate systems, but we will in this section study KAM theorems for numerical schemes in more generality.

There are essentially three assumptions made before KAM theorems can be applied;

1) A certain smoothness is required for both the unperturbed Hamiltonian $\hat{h}_0$ and for the perturbation, $\epsilon\hat{r}_1$.

2) The unperturbed Hamiltonian $\hat{h}_0$ must satisfy a non-degeneracy condition.

3) The trajectory generated by $\hat{h}_0$ is not close to resonant, i.e. $\forall m \in \mathbb{Z}^n \setminus 0$ we have $| < \partial_I \hat{h}_0(I(0), e(0)), \hat{m} > | \geq \Omega(|\hat{m}|_1) > 0$, for some function Ω that is usually taken as $\Omega(s) = \gamma s^{-\zeta}$ for some constants $\zeta, \gamma > 0$.

A qualitative formulation of a KAM theorem is then

Theorem 2 (KAM, Generic) *Suppose* **1)-3)** *above are satisfied for a particular initial value $I(0) \in \mathcal{D}$ and that the perturbation, $\epsilon\hat{r}_1$ is sufficiently small, then there exists a symplectic coordinate transform $(I,\theta) = \Phi(\tilde{I},\tilde{\theta})$ so that in the new coordinates the Hamiltonian $h_0(I) + \epsilon\hat{r}_1(I,\theta)$ only depends on the new action variables, $\tilde{I}$, when represented in the coordinate system $(\tilde{I},\tilde{\theta})$.*

Arnol'ds version of the KAM theorem shows that the measure of the initial conditions for which the tori are destroyed goes to zero as $\epsilon \to 0$ [1].

By Theorem 1 ϵr_2 can be made analytic, and thus assumption **1)** is that of analyticity.[4] The question we are interested in is how assumptions **2)** and **3)** might be affected by numerical discretization. We will call degeneracies introduced by numerical discretization, i.e. when going from h_0 to $\hat{h}_0$, *numerical degeneracies*, while violations of the resonance condition will be referred to as *numerical resonances.*

[4]It is unknown if the construction of ϵr_2 is possible for the C^k category.

2.1 Degeneracy conditions

2.1.1 Kolmogorov vs isoenergetic degeneracy

To start we consider the two classical non-degeneracy conditions from Hamiltonian KAM theory. Let the unperturbed Hamiltonian h_0 satisfy the *Kolmogorov* non-degeneracy condition [1],

$$\det(\partial_I^2 h_0(I)) \neq 0, \qquad \text{for } I \in \mathcal{D}.$$

It is not difficult to see that $(\det(\partial_{I,e}^2 \hat{h}_0) = 0)$ Kolmogorov non-degeneracy is lost through numerical discretization. However, if h_0 is Kolmogorov non-degenerate then one can show that $\hat{h}_0$ is isoenergetically non-degenerate:

$$\det \begin{pmatrix} \partial_{I,e}^2 \hat{h}_0 & \partial_{I,e} \hat{h}_0 \\ \partial_{I,e} \hat{h}_0^T & 0 \end{pmatrix} = \det \begin{pmatrix} \partial_I^2 h_0 & 0 & \partial_I h_0 \\ 0 & 0 & 1/\Delta t \\ \partial_I h_0^T & 1/\Delta t & 0 \end{pmatrix} = -\frac{1}{\Delta t^2} \det(\partial_I^2 h_0(I)) \neq 0,$$

and therefore numerical degeneracy is avoided for Kolmogorov non-degenerate h_0 which implies the results of Shang[47]. So the question that remains in this classical case is what happens to isoenergetic non-degenerate Hamiltonians such as $h_0 = I_1^2/2 + I_2$.

Note 1 *To detect the effect of numerical degeneracy in real simulations is hard as the perturbation $\epsilon \hat{r}_1$ might in fact restore the non-degeneracy. Furthermore from standard BEA or time-averaging [34] it follows that the drift in the action variables due to numerical degeneracy is at most $\mathcal{O}(t \exp(-Cr/\Delta t \|h\|_r))$ or even $\mathcal{O}(t \exp(-Cr/\Delta t \|\epsilon \hat{r}_1\|_r))$ [51] when the step size satisfies certain non-resonance conditions.*

2.1.2 Rüssmann non-degeneracy

The most general non-degeneracy condition for analytic systems which includes the Kolmogorov and isoenergetic non-degeneracy condition is the Rüssmann non-degeneracy condition [41].

Theorem 3 (Rüssmann non-degeneracy) *For analytic h_0, the following condition is necessary and sufficient for KAM non-degeneracy: the image $\partial_I h_0 : \mathcal{D} \mapsto \mathbb{R}^n$ does not lie in any hyperplane passing through the origin.*

It was the important observation by Sevryuk that Rüssmann non-degeneracy is both necessary and sufficient [46]. I.e. for a Rüssmann degenerate Hamiltonian there exists an arbitrary small perturbation ϵr_1 so that all the invariant tori of the unperturbed system are destroyed. For our analysis we will apply the following characterization on Rüssmann non-degeneracy using standard multi-index notation.

Theorem 4 *[41] Let $\mathcal{D} \subset \mathbb{R}^n$ be open and connected. Then for an analytic frequency vector $\omega : \mathcal{D} \mapsto \mathbb{R}^n$ to be non-degenerate it is sufficient that the Taylor expansion*

$$\partial_I h_0(I) = \omega(I) = \sum_{l \in \mathbb{N}^n} \frac{(I - I_0)^l}{l!} \omega^{(l)}(I_0)$$

of $\omega(I)$ at some point *$I_0 \in \mathcal{D}$ contains n linearly independent coefficients $\omega^{(l)}(I_0)$ $l = l_1, l_2, \ldots, l_n \in \mathbb{N}^n$, and it is necessary that among the Taylor coefficients $\omega^{(l)}(I_0)$ there are n linearly independent coefficients at* each point *$I_0 \in \mathcal{D}$.*

As a corollary of this theorem we have

Corollary 1 *Suppose* h_0 *is Rüssmann non-degenerate, then* $\hat{h}_0$ *is Rüssmann non-degenerate.*

The proof is trivial, and follows from the e-independence of $h_0(I)$. This striking result shows that within the most general definition of non-degeneracy, numerical degeneracy does not occur.

It is appropriate to mention that different types of degeneracies have different consequences for which tori are destroyed by perturbations. If the unperturbed Hamiltonian is Kolmogorov or isoenergetic non-degenerate it is known that the measure of the destroyed tori is $\mathcal{O}(\epsilon^{1/2})$ [46] while for Rüssmann non-degenerate h_0 it is $\mathcal{O}(\epsilon^{1/(2N)})$ where N is the highest derivative needed to achieve linear independence in Theorem 4. From the definition of the Rüssman non-degeneracy condition it follows that N is equal for $\hat{h}_0$ and h_0 and the measure of the preserved tori is asymptotically the same order of magnitude. This does not exclude the possibility that some tori invariant in h are destroyed in $\hat{h}$, although we now know that the mechanism for this must be something other than degeneracy.

2.2 The resonance condition

In most presentations of KAM theorems a strong non-resonance condition

$$| < \partial_I h_0(I), m > | \geq \gamma |m|_1^{-\zeta}, \quad \forall m \in \mathbb{Z}_+^n \tag{5}$$

with $\zeta, \gamma > 0$ typically I-dependent is assumed[5]. The resonance condition restricts the action variables I to $\mathcal{D}_\infty \subset \mathcal{D}$, which is a Cantor set, i.e. a closed, totally disconnected and perfect subset of $\mathcal{D}$. For $\mathcal{D}_\infty$ to be of positive measure we can assume $\zeta > n$. When analyzing $\hat{h}_0 + \epsilon r_1$ numerically it is also of interest to study how one must choose the step size Δt not to lose stability for $I \in \mathcal{D}_\infty$, that is, we need a condition on Δt for a strong non-resonance bound like (5) to hold for $| < \partial_{I,e} \hat{h}_0(I, e), \hat{m} > |$ assuming (5) holds. To do that we follow Shang [47] and first define the set of resonant steps

$$\mathcal{R}(\partial_I h_0) := \left\{ \Delta t \in \mathbb{R} : \exists m \in \mathbb{Z}_+^n, k \in \mathbb{Z}_+ \Rightarrow \Delta t = \frac{k}{< \partial_I h_0, m >} \right\}, \tag{6}$$

with $\mathcal{R} = \mathbb{R}$ if for some $m \in \mathbb{Z}_+^n$ $< \partial_I h_0, m >= 0$. The set of non-resonant Δt is then $\mathbb{R} \setminus \mathcal{R}$. Among the non-resonant Δt we have strongly non-resonant steps which we define by

$$\mathcal{NR}(\partial_I h_0) := \left\{ \Delta t \in \mathbb{R} : |\Delta t + \frac{k}{< \partial_I h_0, m >}| \geq \frac{\hat{\gamma} |m|^\zeta}{\gamma (|m|_1 + |k|)^{\hat{\zeta}}} \forall m \in \mathbb{Z}_+^n, k \in \mathbb{Z}_+ \right\}, \tag{7}$$

where ζ and γ are defined through (5), while $\hat{\zeta}$ and $\hat{\gamma}$ positive constants to be determined. We assume for simplicity that $\Delta t < 1$. To show that $\mathcal{NR}$ has positive Lebesgue measure, μ in $\mathbb{R}$, we find

$$\mu(\mathbb{R} \setminus \mathcal{NR}) \quad \leq \quad \frac{2\hat{\gamma}}{\gamma} \sum_{m \in \mathbb{Z}_+^n, k \in \mathbb{Z}_+} \frac{\hat{\gamma} |m|^\zeta}{\gamma (|m|_1 + |k|)^{\hat{\zeta}}}$$

[5] $\mathbb{Z}_+^n = \{m \in \mathbb{Z}^n : |m|_1 \neq 0\}$

$$
\begin{aligned}
&= \frac{2\hat{\gamma}}{\gamma} \sum_{m\in\mathbb{Z}_+^n, k\in\mathbb{Z}_+} \frac{1}{(1+|k|/|m|_1)^{\varsigma}(|m|_1+|k|)^{\hat{\varsigma}-\varsigma}} \\
&\le \frac{2\hat{\gamma}}{\gamma} \sum_{m\in\mathbb{Z}_+^n, k\in\mathbb{Z}_+} \frac{1}{(|m|_1+|k|)^{\hat{\varsigma}-\varsigma}} = \frac{2\hat{\gamma}}{\gamma} \sum_{\hat{m}\in\mathbb{Z}_+^{n+1}} \frac{1}{|\hat{m}|_1^{\hat{\varsigma}-\varsigma}}.
\end{aligned}
$$

The number of vectors $\hat{m} \in \mathbb{Z}_+^{n+1}$ such that $|\hat{m}|_1 = l$ is bounded by $2^{n+1}l^n$ and it follows that

$$
\sum_{\hat{m}\in\mathbb{Z}_+^{n+1}} \frac{1}{|\hat{m}|_1^{\hat{\varsigma}-\varsigma}} \le 2^{n+1} \sum_{l=1}^{\infty} l^{n+\varsigma-\hat{\varsigma}} < \infty,
$$

if $\hat{\varsigma} > n+1+\varsigma$. Since $\mu(\mathbb{R}\setminus\mathcal{NR})$ is bounded $\mu(\mathcal{NR})$ is positive and it can be shown [47, 18] that density of $\mathcal{NR}$ increases when $\Delta t \to 0$. With this definition of $\mathcal{NR}$ we find

$$
\begin{aligned}
|<\partial_{I,e}\hat{h}_0, \hat{m}>| &= \frac{1}{\Delta t}|<\partial_I h_0, m>|\,|\Delta t + \frac{k}{<\partial_I h_0, m>}| \\
&\ge \frac{\hat{\gamma}}{\Delta t(|m|_1+|k|)^{\hat{\varsigma}}} \ge \frac{\hat{\gamma}}{|\hat{m}|_1^{\hat{\varsigma}}},
\end{aligned} \tag{8}
$$

thus $\partial_{I,e}\hat{h}_0$ also satisfies a strong non-resonance condition provided $\Delta t \in \mathcal{NR}$. The estimate $\hat{\varsigma} > 2n+1$ is not sharp, and this originates from the choice of $\mathcal{NR}$. If we consider frequencies and time steps together as frequencies in $\mathbb{R}^{n+1}$ it can be shown that the set

$$
\left\{(\partial_I h_0(I), \Delta t) : |<\partial_{I,e}\hat{h}_0, \hat{m}>| \ge \tilde{\gamma}|\hat{m}|_1^{-\tilde{\varsigma}}, \quad \forall m \in \mathbb{Z}_+^{n+1}\right\} \tag{9}
$$

has a large measure in $\mathbb{R}^{n+1}$ provided $\tilde{\varsigma} > n+1$.

Furthermore the lower bound (5) is one of convenience, and a careful study of weaker assumptions made by Rüssmann [41] has revealed that it is sufficient in the analytic category to assume the lower bound

$$
|<\partial_I h_0, m>| \ge \Omega_I(|m|_1) > 0 \tag{10}
$$

with $\Omega_I(s)$ satisfying

$$
\int_1^{\infty} \Omega_I(s)\log(s)/s^2 ds < \infty.
$$

This weaker assumption enlarges the Cantor set $\mathcal{D}_\infty$, and a similar analysis to that above would be interesting to pursue. With this comment we now formulate

Corollary 2 (KAM, Numerical) *Consider the analytic Hamiltonian $h = h_0(I) + \epsilon r_1(I,\theta)$ where h_0 satisfies the Rüssmann non-degeneracy condition and ϵr_1 is sufficiently small. Suppose $\partial_I h_0(I)$ satisfies (5) and define the Cantor sets $\mathcal{NR}(\partial_I h_0) \subset \mathbb{R}$ for appropriate $\hat{\gamma}$ and $\hat{\varsigma}$. Let the flow of h be discretized by a consistent symplectic, numerical method $\Psi_{\Delta t,h}$ with a sufficiently small $\Delta t \in \mathcal{NR}(\partial_I h_0)$. Then a large measure of the invariant tori in $\mathcal{D}_\infty$ of the original Hamiltonian persist in the discretized dynamics.*

The reason why not all the tori in $\mathcal{D}_\infty$ are preserved is because the perturbation changes from ϵr_1 to $\epsilon \hat{r}_1$.

Note 2 *Other adaptations of KAM theorems as we have done here for the classical case are possible for lower dimensional tori or around elliptic equilibrium points [25]. Such results for lower dimensional tori may then be used to explain the good preservation of first integrals by SI schemes applied to systems that do not have a complete set of n integrals in involution.*

3 Nekhoroshev theorems for symplectic integrators

A weakness of the KAM theory is that the subset of the phase space filled by invariant tori is not open, and determining if an initial value is in $\mathcal{D}_\infty$ requires an infinite number of inequalities to be verified. Further difficulties related to this are introduced by the finite precision effects of computers, guaranteeing that all steps are resonant. Furthermore, it has also been numerically verified that roundoff effects are sufficient for a trajectory to cross a torus [9], whereby showing that invariant tori are not strict barriers to the flows approximated by $\Psi_{\Delta t,h}$ on computers. Roundoff effects aside, when we consider resonant initial points (i.e. $I \in \mathcal{D} \setminus \mathcal{D}_\infty$) it is known that through a process named Arnold diffusion that $\|I(t) - I(0)\|$ might grow as $\mathcal{O}(t)$ when $n > 2$, see e.g. examples in [1, 27]. For specific systems, bounds on the Arnol'd diffusion have been worked out [26]. By making certain assumptions on h_0 one can prove the boundedness of $\|I(t) - I(0)\|$ for exponentially long times for initial values on open subsets of $\mathcal{D}$. Nekhoroshev theorem gives such a bound for the generic class of steep Hamiltonian systems.(For definition, see below.)

Theorem 5 (Nekhoroshev, Generic) *Let $h_0(I)$ be a* steep *analytic function on some domain $I \in \mathbb{R}^n$, and $\epsilon r_1(I, \theta)$ a sufficiently small perturbation, then there exists constants $1 > a, b > 0$ depending only on h_0 so that the action variables $I \in \mathbb{R}^n$ can be bounded as*

$$\|I(0) - I(t)\| = R(\|\epsilon r_1\|), \text{ where } R(\|\epsilon r_1\|) = \mathcal{O}(\|\epsilon r_1\|^b),$$

for times bounded as

$$t < T = \mathcal{O}(\exp(C\|\epsilon r_1\|^{-a})),$$

for some positive constant C that depends on the neighbourhood of $\mathcal{D} \times \mathbb{T}^n$ on which we define the function norm and h_0.

3.0.1 Quasi-convexity vs convexity

The simplest class of steep systems are given by convex and quasi-convex Hamiltonians [2, 36, 7]. A function is said to be m-convex in $\mathcal{D}$ if

$$| < \partial_I^2 h_0 \cdot v, v > | \geq m\|v\|_2^2, \quad m > 0, \quad \forall I \in \mathcal{D}, v \in \mathbb{R}^n,$$

while it is $\hat{m}$-quasi-convex in $\mathcal{D}$ if

$$| < \partial_I^2 h_0 \cdot v, v > | \geq \hat{m}\|v\|_2^2 \ \forall v \in \mathbb{R}^n, v \perp \partial_I h_0, \forall I \in \mathcal{D}.$$

Nekhoroshev theorem holds for both these assumptions with stability exponents of $a = 1/(2n)$ and $b = 1/(2n)$.

Theorem 6 *Suppose h_0 satisfies the m-convexity assumption and that $||\partial_I h_0||_2 \leq M_1$ on $\mathcal{D}$, then $\hat{h}_0$ is quasi-convex on $\mathcal{D}$ with constant $\hat{m} = m/(1+\Delta t^2 M_1^2)$.*

Proof Let $\hat{v} = (v, z)$ with $v \in \mathbb{R}^n$ and $z \in \mathbb{R}$. Then we have

$$\left| < \begin{bmatrix} \partial_I^2 h_0 & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \hat{v}, \hat{v} > \right| = | < \partial_I^2 h_0 \cdot v, v > | \geq m||v||_2^2$$

by the assumption of m-convexity. Now, consider $\partial_{(I,e)}\hat{h}_0 = (\partial_I h_0, 1/\Delta t)$ which gives, assuming $\hat{v} \perp \partial_{(I,e)}\hat{h}_0$, that $< \partial_I h_0, v > + z/\Delta t = 0$. From $||\hat{v}||_2^2 = ||v||_2^2 + |z|^2$ we have $|z| = \Delta t| < \partial_I h_0, v > | \leq \Delta t ||\partial_I h_0||_2 ||v||_2$. From this it follows that

$$\begin{aligned} m||v||_2^2 &= m(||\hat{v}||_2^2 - |z|^2) \geq m(||\hat{v}||_2^2 - \Delta t^2 ||\partial_I h_0||_2^2 ||v||_2^2) \\ \Rightarrow m||v||_2^2 &\geq \frac{m}{1+\Delta t^2 ||\partial_I h_0||_2^2} ||\hat{v}||_2^2. \end{aligned}$$

So on the domain on which $||\partial_I h_0||_2 \leq M_1$ we can conclude that $\hat{h}_0$ is $\hat{m}-$quasi-convex with $\hat{m} = m/(1+\Delta t^2 M_1^2)$. □

This result is the analogue of the Kolmogorov degeneracy and isoenergetic degeneracy in the KAM case. While Theorem 6 implies the Nekhoroshev stability for SI schemes that are applied to a convex h_0, in the case of quasi-convex systems a Nekhoroshev theorem for numerical methods was proved in [30].

Note 3 *This result can be used to explain the behaviour of SI simulations of e.g. weakly coupled rotators with $h_0 = \sum_{j=1}^n I_j^2$ without the tricky resonance condition enforced by KAM theorems. In particular when simulations of systems with Kolmogorov degeneracy in the form of convexity are simulated it may very well be a manifestation of the Nekhoroshev stability the numerical analyst is observing.*

As in the KAM case with Rüssmann degeneracy we now study the generic case for Nekhoroshev theorem to hold.

3.0.2 Steepness

The most general Hamiltonians to which Nekhoroshevs theorem can be applied are the so-called steep systems [33, 35][6] These were first defined by Nekhoroshev [33], who showed that steep functions are generic in the class of C^∞ functions. A Hamiltonian function $h_0(I)$ is called steep on $\mathcal{D}$ with indicies $p_i \geq 1$ and constants C_i and $\delta > 0$ if

$$\max_{0\leq\eta\leq\xi} \min_{||I-I_0||=\eta, I\in\mathcal{D}\cap\Lambda} ||\partial_I(h_0|_\Lambda)(w)|| \geq C_i \xi^{p_i},$$

for all planes $\Lambda \subset \mathbb{R}^n$ with $\dim\Lambda = i$ and for each $\xi \in [0, \delta]$. $\partial_I(h_0|_\Lambda)$ denotes the gradient of h_0 projected onto the plane Λ. Note that it is sufficient to consider only the planes orthogonal to $\partial_I h_0$. Il'yashenko [21] gives the following, simpler, characterization of steep analytic functions.

Theorem 7 *Let $h_0 : \mathcal{D} \mapsto \mathbb{R}$ be an analytic function defined in some neighbourhood of a closed bounded domain $\mathcal{D} \subset \mathbb{R}^n$. Suppose that h_0 does not have critical points and that the restriction of h_0 to each plane Λ ($1 \leq \dim\Lambda < n$) has only $\mathbb{C}$-isolated critical points. Then h_0 is steep in $\mathcal{D}$.*

[6]For orbits close to elliptic equilibrium points a weakening of steepness is possible[14].

From this theorem it follows that

Corollary 3 *If the analytic function* h_0 *is steep in* $\mathcal{D}$*, then* $\hat{h}_0$ *is steep in* $\mathcal{D}$.

This is the analogue of Corollary 1 and shows that SI schemes do not destroy the steepness properties of the original system.

Quantitatively the most interesting estimates of the Nekhoroshev theorem are the stability exponents a and b. Niederman [35] showed that, assuming $\|\partial_I^2 h_0^{-1}\| < M$ and $1/M \leq \|\partial_I h_0\| \leq M$ for $M > 1$ one may take

$$a = b = 1/((2n-1)p_1 \cdot p_2 \cdots p_{n-1} + 1)$$

as a generalization of the believed-to-be optimal estimates for convex and quasi-convex Hamiltonians where $p_i = 1$.

Theorem 8 *Suppose the Hamiltonian* h_0 *has stability indices* $\{p_1, \ldots, p_{n-1}\}$ *and that* $\|\partial_I^2 h_0^{-1}\| < M$ *on* $\mathcal{D}$*, then* $\hat{h}_0$ *has indices* $\{p_1, \ldots, p_{n-1}, 1\}$ *on* $\mathcal{D}$.

Proof Let $\hat{\Lambda}$ be a plane in $\mathbb{R}^{n+1}$ with $1 \leq \dim(\hat{\Lambda}) \leq n$, and $\hat{P}$ an orthogonal projection onto $\hat{\Lambda}$ so that $\hat{P}(\partial_{I,e}\hat{h}_0) = 0$. Partitioning $\hat{P} = [\tilde{P}; \tilde{q}]$, gives $\tilde{P}\partial_I h_0(I_0) + \tilde{q}/\Delta t = 0$, and we find

$$\|\hat{P}\partial_{I,e}h_0(I)\| = \|\tilde{P}(\partial_I h_0(I) - \partial_I h_0(I_0))\|.$$

Expanding $\tilde{P}$ and $\tilde{q}$ in powers of Δt we find that $\tilde{P} = \tilde{P}_0 + \mathcal{O}(\Delta t)$ and $\tilde{q} = \Delta t \tilde{q}_1 + \mathcal{O}(\Delta t^2)$, and it follows that $\tilde{P}_0$ must be an orthogonal projection operator. The following bound is trivial

$$\|\tilde{P}(\partial_I h_0(I) - \partial_I h_0(I_0))\| \geq \|\tilde{P}_0(\partial_I h_0(I) - \partial_I h_0(I_0))\| - \|(\tilde{P} - \tilde{P}_0)(\partial_I h_0(I) - \partial_I h_0(I_0))\|.$$

Because $\|(\tilde{P} - \tilde{P}_0)(\partial_I h_0(I) - \partial_I h_0(I_0))\| \to 0$ when $\Delta t \to 0$ it follows that there exists a Δt^* and a constant $c_1 < 1$ so that

$$\begin{aligned}\|\hat{P}\partial_{I,e}h_0(I)\| &= \|\tilde{P}_0(\partial_I h_0(I) - \partial_I h_0(I_0))\| \\ &\geq c_1\|\tilde{P}_0(\partial_I h_0(I) - \partial_I h_0(I_0))\|, \quad \Delta t \leq \Delta t^*.\end{aligned}$$

When $i = \operatorname{rank}(\tilde{P}_0) < n$ by the steepness assumption on h_0 we have that the following bound holds

$$\max_{0 \leq \eta \leq \xi} \min_{\|I - I_0\| = \eta, (I,e) \in \hat{\Lambda}} c_1\|\tilde{P}_0(\partial_I h_0(I) - \partial_I h_0(I_0))\| \geq C_i \xi^{p_i}.$$

However when $\dim(\hat{\Lambda}) = n$ $\tilde{P}_0$ is full rank, the steepness can not be used, but instead

$$\begin{aligned} & \max_{0 \leq \eta \leq \xi} \min_{\|I - I_0\| = \eta, (I,e) \in \hat{\Lambda}} c_1\|\tilde{P}_0(\partial_I h_0(I) - \partial_I h_0(I_0))\| \\ = & \max_{0 \leq \eta \leq \xi} \min_{\|I - I_0\| = \eta, (I,e) \in \hat{\Lambda}} c_1\|(\partial_I h_0(I) - \partial_I h_0(I_0))\| \\ = & \max_{0 \leq \eta \leq \xi} \min_{\|I - I_0\| = \eta, (I,e) \in \hat{\Lambda}} c_1\|\partial_I^2 h_0(I^*)(I - I_0)\| \\ \geq & \max_{0 \leq \eta \leq \xi} \min_{\|I - I_0\| = \eta, (I,e) \in \hat{\Lambda}} c_1\|\partial_I^2 h_0(I^*)^{-1}\|^{-1}\|(I - I_0)\| \\ \geq & \; c_1\|\partial_I^2 h_0(I^*)^{-1}\|^{-1}\xi = c_1 M \xi = C_n \xi, \end{aligned}$$

thus $p_n = 1$.

To complete the proof of steepness we need to show that the above bounds hold, with possibly different constants C_i, when replacing

$$\min_{(I,e)\in\Lambda,\|I-I_0\|=\eta} \quad \text{with} \quad \min_{(I,e)\in\Lambda,\|(I,e)-(I_0,e_0)\|=\eta}.$$

To do this we consider the Taylor expansion of $\hat{h}_0(I,e)$ around $(I_0,e_0) \in \Lambda$ with $(I,e) \in \Lambda$

$$\begin{aligned}\hat{h}_0(I,e) &= \hat{h}_0(I_0,e_0) + \hat{P}(\partial_{I,e}\hat{h}_0)(I_0,e_0)(I-I_0,e-e_0) + \partial_I^2(h_0(I^*|_{\hat{\Lambda}}))(I-I_0,I-I_0)\\ &= \hat{h}_0|_\Lambda(I_0,e_0) + \partial_I^2(h_0(I^*|_{\hat{\Lambda}}))(I-I_0,I-I_0)\end{aligned}$$

for some $I^* \in [I_0, I]$. Using the bounds $\|\partial_I h_0\| < M_1$ and $\|\partial_I^2 h_0\| < M_2$ it follows from the Taylor expansion that

$$|e-e_0| \quad \leq \quad \Delta t(M_1\|I-I_0\| + M_2\|I-I_0\|^2) = \Delta t m^*\|I-I_0\|,$$

for sufficiently small $\|I-I_0\|$. From this bound we have

$$\begin{aligned}&\max_{0\leq\eta\leq\xi^*} \min_{\substack{(I,e)\in\Lambda\\ \|(I,e)-(I_0,e_0)\|=\eta}} \|\hat{P}(\partial_{I,e}\hat{h}_0)\|\\
=\;& \max_{0\leq\eta\leq\xi} \min_{\substack{(I,e)\in\Lambda\\ \|I-I_0\|^2+|e-e_0|^2=\eta^2}} \|\hat{P}(\partial_{I,e}\hat{h}_0)\|\\
\geq\;& \max_{0\leq\eta\leq\xi} \min_{\substack{(I,e)\in\Lambda\\ \eta^2\geq\|I-I_0\|^2\geq\eta^2-(\Delta t m^*)^2\|I-I_0\|^2}} \|\hat{P}(\partial_{I,e}\hat{h}_0)\|\\
=\;& \max_{0\leq\eta\leq\xi} \min_{\substack{(I,e)\in\Lambda\\ \eta^2\geq\|I-I_0\|^2\geq\eta^2/(1+(\Delta t m^*)^2)}} \|\hat{P}(\partial_{I,e}\hat{h}_0)\|\\
\geq\;& \max_{0\leq\eta\leq\frac{\xi}{\sqrt{1+(\Delta t m^*)^2}}} \min_{\substack{(I,e)\in\Lambda\\ \|I-I_0\|=\eta}} \|\hat{P}(\partial_{I,e}\hat{h}_0)\|\\
\geq\;& \frac{C_i}{(1+(\Delta t m^*)^2)^{p_i/2}}\xi^{p_i}.\end{aligned}$$

Therefore $\hat{h}_0$ is steep with indicies $\hat{p}_i = p_i$, $1 \leq i < n$ and $\hat{p}_n = 1$ and constants $\hat{C}_i = \frac{C_i}{(1+(\Delta t m^*)^2)^{p_i/2}}$, where $C_n = c_1 M$. □

Corollary 4 *Let $h_0(I)$ be steep, and suppose the flow of the perturbed system $h = h_0(I)+\epsilon r_1(I,\theta)$ is discretized by a consistent symplectic integration method, $(I_{n+1},\theta_{n+1}) = \Psi_{\Delta t,h}(I_n,\theta_n)$. Then for sufficiently small Δt, and ϵr_1 the variation in the action variables is bounded over exponentially long time-intervals.*

$$\|I_n - I_0\| \leq R(\epsilon, \Delta t), \text{ for } |n| \leq \frac{1}{\Delta t}T,$$

where $R(\|\epsilon\hat{r}_1\|) = \mathcal{O}(\|\epsilon\hat{r}_1\|^{\hat{b}})$ and $T = \mathcal{O}(\exp(\hat{C}/\|\epsilon\hat{r}_1\|^{\hat{a}}))$ for a positive constant $\hat{C}$ and

$$\hat{a} = \hat{b} = \frac{1}{(2n+1)p_1\cdots p_{n-1}+1}.$$

In order to prove the corollary a small modification of Niedermans proof is necessary. Niederman assumes that $\partial_I^2 h_0^{-1}$ is bounded, while for $\hat{h}_0$ we assume that the isoenergetic non-degeneracy holds with the corresonding bound.

It was noticed by Pöschel that for initial values $I(0)$ close to resonant orbits the stability exponents a, b can be improved. In particular for periodic orbits in convex and quasi-convex systems one may take $a = b = 1/2$ [36, 7], however there are further improvements in the neighbourhood of strongly non-resonant orbits.

3.0.3 Super-exponential estimates

Although steepness and non-degeneracy as required by the Nekhoroshev and KAM theorems are independent concepts, the intersection of the set of steep functions and e.g. Rüssmann non-degenerate functions is not empty. It is clear, for example, that convexity implies Kolmogorov non-degeneracy, and it is then natural to ask if improvements in the Nekhoroshev estimates can be made if non-resonance conditions are taken into account.

A complete picture of what happens to convex Hamiltonian systems was given by Morbidelli and Giorgilli [32]. By studying the proof of the Nekhoroshev theorem a drastic improvement of the stability in the neighbourhood of points $I \in \mathcal{D}_\infty$ was discovered. An alternative proof for the quasi-convex case which is of interest to us was given in [7]. Let δ denote the distance in $\mathcal{D}$ from a torus satisfying (5), then Delshams and Gutiérrez state that for quasi-convex $\hat{h}_0$

$$\begin{aligned} \|I(t) - I(0)\| &= \mathcal{O}(\|\epsilon\hat{r}_1\| \exp\left(-\delta^{-1/(\zeta+1)}\right)^{1/2n}), \quad |t| < T \\ \text{where } T &= \mathcal{O}(\exp\left(\frac{C_1}{\|\epsilon\hat{r}_1\|}\exp(C_2\delta^{-1/(\zeta+1)})\right)^{1/2n}), \end{aligned} \tag{11}$$

for some positive constants C_1 and C_2.

Note 4 *The bound (11) emphasizes Note 3 on numerical integration of convex Hamiltonians even more. In particular when simulating convex Hamiltonians the tori $\mathcal{D}_\infty$ represent "very sticky" sets [51] and it is this stickiness that gives rise to the observations from computer simulations.*

3.0.4 Discretizations of perturbed linear systems

If no structural assumptions are made on the perturbations, perturbed systems of Harmonic oscillators $h =< \omega, I > +\epsilon r_1(I, \theta)$ are degenerate in the sense of Rüssmann. In this situation the best estimates one can obtain are for strongly non-resonant perturbed linear systems giving exponential stability, $T = \mathcal{O}(\exp(C/\|\epsilon\hat{r}_1\|^{1/(\zeta+1)}))$, see e.g. [51, 18, 48, 11, 30]. Indeed Katok [22] showed that for systems with $\omega_i > 0$ arbitrary small perturbations exist so that the motions are ergodic on the energy surface, and hence the I are not necessarily bounded as $t \to n$. See Lochak [27] for a discussion of how the nonlinearity of $\epsilon\hat{r}_1$ can restore the non-degeneracy for systems of Harmonic oscillators, which may explain why numerical simulations of resonant systems of oscillators still show good preservation of invariants. For linear systems of resonant oscillators other mechanisms stabilizing the error growth might explain the good preservation of invariants in numerical experiments[18], although it is unlikely that the variation in all the action variables(integrals) can be bounded in such cases, which is necessary for good error propagation.

3.0.5 Singularities in action-angle coordinates

We have relied on action angle coordinates and the assumption that h_0 and ϵr_1 are analytic along the trajectory. One important case when the analyticity does not hold is at elliptic equilibrium points. One can still deduce stability results in annular neighbourhoods of the equilibrium point but not on domains that include it. Cartesian coordinates avoid this singularity, and special versions of the Nekhoroshev theorems [13, 37] and KAM theorems have been developed for this case. The version of the Nekhoroshev theorem proved in [13] can be adapted to the quasi-convex case, and we may therefore apply it to $\hat{h}_0$. In contrast to the standard Nekhoroshev theorem, a weak non-resonance condition is needed in this case. Assuming that the equilibrium point is at $(p, q) = 0$ and forming the Taylor expansion $h(p, q) = h_2 + h_3 + h_4 + + \cdots$, where h_j is homogeneous of order j in p and q the first condition of [13] is that $< \omega, m > \neq 0$ where $h_2 = \sum_j \omega_j (p_j^2 + q_j^2)$ for $|m| \leq 4$ while the second is that the terms $h_3 + h_4$ makes $h_2 + h_3 + h_4$ convex (or quasi-convex). The stability for the SI schemes follows by the same argument as for the standard Nekhoroshev result, and we may therefore conclude that also for elliptic equilibrium points a Nekhoroshev type theorem holds for SI schemes provided the step size Δt belongs to some open set of non-resonant steps up to order 4.

4 Implications for the global error growth

Under different conditions several authors have given proofs bounding the error growth of SI schemes [10, 5, 6, 17, 48, 18, 30]. In the earliest results quite degenerate systems with periodic orbits were studied, while some of the latter gave results for strongly non-resonant orbits. We are now able to give an error bound with even weaker assumptions on the Hamiltonian h, in particular we will see that the Nekhoroshev bound gives linear error bounds for exponentially long intervals.

We assume that h_0 satisfies non-degeneracy, steepness or strong non-resonance conditions such as those required by the Nekhoroshev or KAM theorems so that Corollaries 2 and 4 hold with the corresponding bound R on the variations of $I(t)$ for $t < T$.

We will derive a bound on the error based on making two comparisons. First we compare the orbits $(I_0(t) = I_0, \theta_0(t) = \theta_0 + t\partial_I h_0(I_0))$ and $(I_1(t), \theta_1(t))$ corresponding to the flow of h_0 and $h_0 + \epsilon r_1$ respectively. Second we compare the orbit $(I_0(t), \theta_0(t))$ with that of $h_0 + \epsilon r_1 + \epsilon r_2 = h_0 + \epsilon \hat{r}_1$, given by $(I_2(t), \theta_2(t))$ all with the same initial values $I_0 \in \mathcal{D}$, $\theta_0 \in \mathbb{T}^n$. The triangle inequality then gives an upper bound for the global error.

We start by using the nonlinear variation of constants formula,

$$\begin{pmatrix} I_1(t) - I_0(t) \\ \theta_1(t) - \theta_0(t) \end{pmatrix} = \int_0^t \begin{bmatrix} 1\!\!1 & 0 \\ s\partial_I^2 h_0(I_1(s)) & 1\!\!1 \end{bmatrix} \begin{pmatrix} -\partial_\theta \epsilon r_1(I_1(s), \theta_1(s)) \\ \partial_I \epsilon r_1(I_1(s), \theta_1(s)) \end{pmatrix} ds,$$

hence for the angles this gives

$$\begin{aligned} \|\theta_1(t) - \theta_0(t)\| &\leq \left\| \int_0^t s\partial_I^2 h_0(I_1(s))\partial_\theta \epsilon r_1(I_1(s), \theta_1(s))ds \right\| + t \sup_{I \in \mathcal{D}, \theta \in \mathbb{T}^n} \|\partial_I \epsilon r_1(I, \theta)\| \\ &= \left\| t \int_0^t \partial_I^2 h_0(I_1(s))\partial_\theta \epsilon r_1 \, ds - \int_0^t \int_0^s \partial_I^2 h_0(I_1(s))\partial_\theta \epsilon r_1(I_1(s), \theta_1(s))ds \right\| \end{aligned}$$

$$
\begin{aligned}
&+ \quad t \sup_{I\in\mathcal{D},\theta\in\mathbb{T}^n} \|\partial_I \epsilon r_1(I,\theta)\| \\
&\le \quad t\left\|\int_0^t \frac{\partial}{\partial s}\partial_I h_0(I_1(s))ds\right\| + \left\|\int_0^t \partial_I h_0(I_1(s)) - \partial_I h_0(I_0))ds\right\| \\
&+ \quad t \sup_{I\in\mathcal{D},\theta\in\mathbb{T}^n} \|\partial_I \epsilon r_1(I,\theta)\| \\
&\le \quad 2tM_2R(\|\epsilon r_1\|) + t \sup_{I\in\mathcal{D},\theta\in\mathbb{T}^n} \|\partial_I \epsilon r_1(I,\theta)\|, \qquad (12)
\end{aligned}
$$

where $M_2 := \sup_{I\in\mathcal{D}} \|\partial_I^2 h_0(I)\|$.

Since ϵr_1 is bounded on $I \in \mathcal{D}$ and $R(\|\epsilon r_1\|)$ is bounded for $t < T$ it easily follows that $\|\theta_1(t) - \theta_0(t)\|$ is bounded by a function that grows linearly with time at most. The bound on $\|I_1(t) - I_0(t)\|$ for $t < T$ follows from the assumptions and after applying the corresponding stability theorem.

Because of Corollaries 2 and 4 exactly the same bounds hold for $(I_2(t), \theta_2(t))$ with the only modification that R and T is now a function of $\|\epsilon\hat{r}_1\|$ instead of ϵr_1. From these two bounds and the triangle inequality we have that the error in the numerical integration can be bounded by a linearly growing function:

$$
\begin{aligned}
\|I_2(t) - I_1(t)\| &\le C_1, \\
\|\theta_2(t) - \theta_1(t)\| &\le tC_2, \quad \text{for } t < T,
\end{aligned}
$$

with

$$
\begin{aligned}
C_1 &= R(\|\epsilon r_1\|) + R(\|\epsilon\hat{r}_1\|) \\
C_2 &= 2M_2(R(\|\epsilon r_1\|) + R(\|\epsilon\hat{r}_1\|)) \\
&+ \sup_{I\in\mathcal{D},\theta\in\mathbb{T}^n} \|\partial_I \epsilon r_1(I,\theta)\| + \sup_{I\in\mathcal{D},\theta\in\mathbb{T}^n} \|\partial_I \epsilon\hat{r}_1(I,\theta)\|.
\end{aligned}
$$

We can summarize the stability results in the following table which is ordered by decreasing stability times

Degeneracy	Resonance	$\Delta t \in$	$T =$	$R =$
Rüssmann	non-resonance (5)	$\mathcal{NR}$	∞	$\mathcal{O}(\|\epsilon\hat{r}_1\|^{1/2})$
convex	non-resonance	$\mathcal{NR}$	super exponential (11)	$\mathcal{O}(\|\epsilon\hat{r}_1\|^{1/2})$
degenerate	non-resonance	$\mathcal{NR}$	$\mathcal{O}(\exp(C/\|\epsilon\hat{r}_1\|^{1/(\hat{\varsigma}+1)}))$	$\mathcal{O}(\|\epsilon\hat{r}_1\|)$
steep	any	any	$\mathcal{O}(\exp(C/\|\epsilon\hat{r}_1\|^{\hat{a}}))$	$\mathcal{O}(\|\epsilon\hat{r}_1\|^{\hat{b}})$
degenerate	any	any	$\mathcal{O}(\|\epsilon\hat{r}_1\|^{-1})$	1

From the proof above it is apparent that for a Hamiltonian system for which one can establish boundedness in the variation of the actions, I for $t < T$ the numerical approximation produced by a SI scheme will have a global error that can be bounded by a linearly growing function in t up to $t < T$.

Example 1 *The two degrees of freedom Hamiltonian*

$$
h = \frac{1}{2}(I_1^2 - I_2^2) + \epsilon\sin(\theta_1 - \theta_2)
$$

admits the solution $I(t) = (-\epsilon t, \epsilon t)$ $\theta(t) = -\frac{1}{2}(\epsilon t^2, \epsilon t^2)$, *thus the actions are not bounded for* $t \in \mathbb{R}$ *for this particular solution. This Hamiltonian is Kolmogorov* *non-degenerate, but not steep and we may expect some initial values to be on, or close to, invariant tori, to further improve the chances of this we choose the small value* $\epsilon = 10^{-7}$ *for the parameter. A numerical simulation of this trajectory shows a quadratically growing error in* θ_1*, while a linearly growing error in* I_1*.*

Changing the initial values to $(I_1, I_2) = (1, \gamma)$, $\gamma = (\sqrt{5}+1)/2$ *and* $(I_1, I_2) = (1/10, \gamma/10)$ *(*$\theta_1 = \theta_2 = 0$*) ensures that the strong non-resonance condition is satisfied and the solution together with the numerical approximation behaves very differently. We see, for this case, in Figure 1 a piecewise linear growth in the global error in* θ_1*. This can be explained by the stickiness of invariant tori. Once a trajectory comes close to an invariant torus the actions,* I*, remain close to that torus for a relatively long time, and in this way forcing the angles to evolve linearly. Then through the slow drift away from the torus the stickiness is rapidly lost and the actions quite quickly evolve until they hit the next torus and the angles evolve linearly again et cetera. Figure 2 shows the error in the action* I_1*, which is linearly growing for* $(I_1, I_2) = (0, 0)$*.*

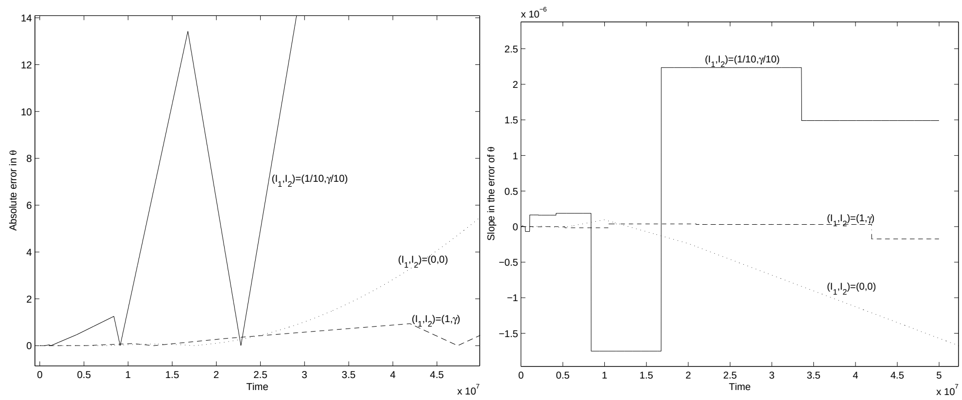

Figure 1: Graphs of global error in θ_1 against time shows a quadratically growing error in the angles for the initial value $(I_1, I_2) = (0, 0)$, while for the other two initial values give a linearly bounded error growth. The integration was done using the Störmer/Verlet scheme with a step size of $\Delta t = \sqrt{3}/100$ which should ensure that $\Delta t \in \mathcal{NR}(\partial_I h_0)$.

Note 5 *It is clear from the above error bound that the linear growth in the error originates from phase errors in the angle variables. In a version of the proof of KAM theorems the perturbed system* $h = h_0 + \epsilon r_1$ *is put into a normal form which only depends on the actions,* $h(\Phi(\tilde{I}, \tilde{\theta})) = \tilde{h}_0(\tilde{I}) = h_0(\tilde{I}) + \epsilon n_1(\tilde{I}, \epsilon)$*, while for the numerical scheme* $\hat{h}$ *is put into the form* $h_0 + \epsilon n_1 + \epsilon n_2$*. The phase errors then behave as* $t\partial_I n_2(I_0)$*. However, by the non-degeneracy condition on* h_0 *there exists a modified initial value* I_0^* *so that* $\partial_I h_0(I_0) + \partial_I \epsilon n_1(I_0) = \partial_I h_0(I_0^*) + \partial_I n_1(I_0^*) + \partial_I n_2(I_0^*)$*, and therefore the frequencies of the orbits produced by the numerical scheme can be made equal to those of the exact solution by a perturbation in the initial values. If that is achieved then the numerical integration error will be bounded by a constant. This idea was partly carried out by Saha and Tremaine [42] who create a warm-up scheme that in effect modifies the initial values so that all the terms of order* ϵ *are removed from the phase errors. This represents an alternative to the coordinate transformation used by the very accurate method of Wisdom and Holman [52].*

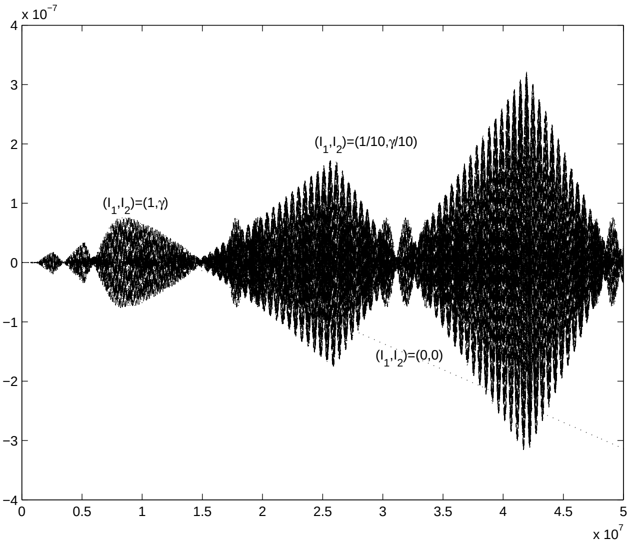

Figure 2: Graphs of global error in I_1 against time shows a linearly growing error for the initial value $(I_1, I_2) = (0, 0)$, while the error remain bounded, or slowly growing for the two other initial values(whose graphs overlap).

4.1 Non-symplectic numerical integration schemes

So far we have only considered symplectic flows discretized by SI schemes. We will now focus on two general cases of vector fields that often share similar stability results as Hamiltonian vector fields. They are the vector fields of Lie subalgebras of the Lie algebra of smooth vector fields in $\mathbb{R}^n$(In addition to Hamiltonian vector fields there are divergence free vector fields and contact vector fields [29].) and vector fields with a time reversing involution $\mathcal{I} : \mathbb{R}^n \mapsto \mathbb{R}^n$. A mapping $\mathcal{I}$ is called an involution if $\mathcal{I} \circ \mathcal{I}(x) = x$, and the vector field $g(y,t)$ is said to be reversible with respect to $\mathcal{I}$ provided $g(y,t) = -d\mathcal{I}g(\mathcal{I}(y), -t)$, where $d\mathcal{I}$ is the Jacobian of $\mathcal{I}$ in $\mathbb{R}^n$. Pronin and Treschev [49] gave a theorem which can be formulated for $\Psi_{\Delta t,h}$ as follows

Theorem 9 *Let $\Psi_{\Delta t,h}$ be an analytic map on $\mathbb{R}^n$ isotropic to the identity. Then there exists a vector field $g(y,t)$ analytic and Δt-periodic in t so that its flow coincides with $\Psi_{\Delta t,h}$. Furthermore if $\Psi_{\Delta t,f}$ belongs to some Lie subgroup of the group of diffeomorphisms on $\mathbb{R}^n$, then g can be constructed so that it belongs to the corresponding Lie subalgebra of smooth vector fields. Or if $\Psi_{\Delta t,f}$ is reversible, i.e. $\mathcal{I} \circ \Psi_{\Delta t,f} \circ \mathcal{I} = \Psi^{-1}_{\Delta t,f}$, then g can be made $\mathcal{I}$-reversible.*

Theorem 9 can be a starting point for the analysis of more general geometric integrators that e.g. preserve reversing symmetries, volume or contact structure for which KAM theorems do exist, while less is known regarding the existence of Nekhoroshev type results. For systems with reversing symmetries and volume preserving integrators results for the case of perturbed linear oscillators were given in [30, 18], while KAM results for reversible systems can e.g. be found in [45] and in the volume preserving case in [53]. Numerical experiments indicate that both these cases give rise to the same linear error growth as we have shown for Hamiltonian systems [18, 38].

4.2 BEA vs. na-BEA

In the rigorous derivation of results in BEA, few assumptions are made on the structure of the flow of h, and is naturally restricted by the worst case scenario (resonance). This is in contrast to the systems for which a good understanding of stability exists such as the close to integrable systems with non-degenerate unperturbed part h_0 with an appropriate resonance condition. Applying BEA to systems for which h_0 is non-degenerate and non-resonant is therefore heavy handed, while na-BEA leads to a Hamiltonian $\hat{h}$ which is sufficient for applying the stability results which were valid for h. If one is willing to accept suboptimal estimates such as (1) BEA gives the advantage that $\overline{h}^* = h_0 + \epsilon r_1 + \mathcal{O}(\epsilon)$, hence the non-degeneracy and steepness conditions are automatically satisfied for the perturbed Hamiltonian $\overline{h}^*$ and most of the discussion in this paper would have been avoided, at the cost of optimal stability times, T.

5 Conclusions

The qualitative behavior seen for symplectic integrators has been explained here using well established results from Hamiltonian perturbation theory in as general a setting as possible. We have not aimed at quantitative bounds on e.g. $\epsilon\hat{r}_1$, but believe that the results as we have stated them give a complete picture of what happens when close to integrable systems are discretized by SI schemes. The main finding is that within the class of Rüssmann non-degenerate Hamiltonians the transformation $h_0 \mapsto \hat{h}_0 := h_0 + e/\Delta t$ does not introduce degeneracy and the same holds for steep systems. This fact together with Theorem 1 then opens up the whole theory available to Hamiltonian differential equations for use in the analysis of symplectic numerical integration schemes. The classical stability theorems such as the KAM and Nekhoroshev results are then applied to $\hat{h}$ and a general linear error bound for SI schemes is derived. Indeed it follows from the analysis that for SI schemes it is the boundedness in the variation in the actions that leads to a linear bound on the error.
Acknowledgements: The author thanks the Australian Reserach Council for its financial support, and R. Quispel for reading through a draft of this paper and discussions on symplectic integrators.

References

[1] V.I. Arnold, V.I., V.V. Kozlov, and A.I. Neishtadt, “Mathematical aspects of classical and celestial mechanics, in: Dynamical Systems III ”, V.I. Arnold, ed. Springer-Verlag 1988.

[2] G. Benettin, L. Galgani and A. Giorgilli, “A proof of Nekhoroshev’s theorem for the stability time in nearly integrable Hamiltonian systems” *Celestial Mech. Dyn. Astr.* **37** (1985) no. 1, 1–25.

[3] G. Benettin and A. Giorgilli, “On the Hamiltonian interpolation of near-to-the-identity symplectic mappings with application to symplectic integration algorithms” J. Statist. Phys. 74 (1994), no. 5-6, 1117–1143.

[4] O.I. Bogoyavlenskij, “Necessary conditions for existence of non-degenerate Hamiltonian structures” Comm. Math. Phys. 182 (1996), no. 2, 253–289.

[5] B. Cano and J.M. Sanz-Serna, "Error growth in the numerical integration of periodic orbits, with application to Hamiltonian and reversible systems" SIAM J. Numer. Anal. 34 (1997), no. 4, 1391–1417

[6] M.P. Calvo and E. Hairer, "Accurate long-term integration of dynamical systems" Seventh Conference on the Numerical Treatment of Differential Equations (Halle, 1994). Appl. Numer. Math. 18 (1995), no. 1-3, 95–105.

[7] A. Delshams and P. Gutiérrez, "Effective stability and KAM theory" J. Differential Equations 128 (1996), no. 2, 415–490.

[8] R. Douday, "Une démonstration directe de l'équivalance des théorèmes de tores invariants pour difféomorphismes et champs de vecteurs", C.R. Acad. Sci Paris Ser. I Math. vol. 295 no 2. (1982), 201-204.

[9] D.J.D. Earn and S. Tremaine, "Exact numerical studies of Hamiltonian maps: iterating without roundoff error" Phys. D 56 (1992), no. 1, 1–22.

[10] D.J. Estep, A.M. Stuart, "The rate of error growth in Hamiltonian-conserving integrators" Z. Angew. Math. Phys. 46 (1995), no. 3, 407–418.

[11] F. Fasso, "Lie series method for vector fields and Hamiltonian perturbation theory" ZAMP 41(1990) 843-864.

[12] B. Fiedler and J. Scheurle, "Discretization of homoclinic orbits, rapid forcing and "invisible" chaos", Mem. Amer. Math. Sov. 119, no. 570, 1996.

[13] M. Guzzo, F. Fassó and G. Benettin, " On the stability of elliptic equilibria" Math. Phys. Electron. J. 4 (1998), Paper 1, 16 pp.

[14] G. Benettin, F. Fassó and M. Guzzo, "Nekhoroshev-stability of L4 and L5 in the spatial restircted three-body problem" Regular and Chaotic dynamics 3(1998) 56-72 .

[15] H.W. Broer, G.B. Huitema and M.B. Sevryuk, "Families of quasi-periodic motions in dynamical systems depending on parameters" Nonlinear dynamical systems and chaos (Groningen, 1995), 171–211, Progr. Nonlinear Differential Equations Appl., 19, Birkhuser, Basel, 1996.

[16] O. Gonzalez, D.J. Higham and A.M. Stuart, "Qualitative properties of modified equations" IMA J. Numer. Anal. 19 (1999), no. 2, 169–190.

[17] E. Hairer and Ch. Lubich, "The life-span of backward error analysis for numerical integrators" Numer. Math. 76 (1997), no. 4, 441–462.

[18] E. Hairer, Ch. Lubich and G. Wanner, "Geometric numerical integration. Structure-preserving algorithms for ordinary differential equations." Springer Series in Computational Mathematics, 31. Springer-Verlag, Berlin, 2002.

[19] E. Hairer, Ch. Lubich and G. Wanner, "Geometric numerical integration illustrated by the Störmer/Verlet method" to appear in Acta Numerical 2003, Cambridge University Press.

[20] A. Haro, "Interpolation of an exact symplectomorphism by a Hamiltonian flow", Mathematical Physics Preprint Archive 2000:http://www.ma.utexas.edu/mp

[21] Yu.S. Il'yashenko, " A criterion of steepness for analytic functions" Uspekhi Mat. Nauk 41 (1986), no. 1(247), 193–194.

[22] A.B. Katok, "Ergodic perturbations of degenerate integrable Hamiltonian systems", Math. USSR. Izv. 7, (1973), 535-572.

[23] A.N. Kolmogorov,. "On conservation of conditionally periodic motions for a small change in Hamilton's function" (Russian) Dokl. Akad. Nauk SSSR (N.S.) 98, (1954). 527–530. V.I. Arnol'd, "The stability of the equilibrium position of a Hamiltonian system of ordinary differential equations in the general elliptic case" (Russian) Dokl. Akad. Nauk SSSR 137 1961 255–257. J. Moser, "On invariant curves of area-preserving mappings of an annulus" Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II 1962 1962 1–20.

[24] S. Kuksin, "On the inclusion of an almost integrable analytic symplectomorphism into a Hamiltonian flow" Russian J. Math. Phys. 1 (1993), no. 2, 191–207

[25] S. Kuksin and J. Pöschel, "On the inclusion of analytic symplectic maps in Hamiltonian flows and its applications" Seminar on Dynamical Systems (St. Petersburg, 1991), 96–116, Progr. Nonlinear Differential Equations Appl., 12, Birkhäuser, Basel, 1994.

[26] A.H. Lichtenberg and M.A. Lieberman, "Regular and Stochastic Motion" AMS 38, Springer-Verlag 1983.

[27] P. Lochak, "Stability of Hamiltonian systems over exponentially long times: the near-linear case." Hamiltonian dynamical systems (Cincinnati, OH, 1992), 221–229, IMA Vol. Math. Appl., 63, Springer, New York, 1995. ibid. "Hamiltonian perturbation theory: periodic orbits, resonances and intermittency" Nonlinearity 6 (1993), no. 6, 885–904. ibid. "Canonical perturbation theory: an approach based on joint approximations" Russian Math. Surveys 47 (1992), no. 6, 57–133.

[28] R.I. McLachlan and P. Atela, "The accuracy of symplectic integrators" Nonlinearity 5 (1992), no. 2, 541–562.

[29] R.I. McLachlan and G.R.W. Quispel, "What kinds of dynamics are there? Lie pseudogroups, dynamical systems and geometric integration" Nonlinearity 14 (2001), no. 6, 1689–1705.

[30] P.C. Moan, "On backward error analysis and Nekhoroshev stability in the numerical analysis of conservative systems of ODEs" PhD thesis, University of Cambridge, 2002.

[31] P.C. Moan, "Symplectic integration of close to integrable Hamiltonian systems", in preparation.

[32] A. Morbidelli and A. Giorgilli, "Superexponential stability of KAM tori" J. Statist. Phys. 78 (1995), no. 5-6, 1607–1617. ibid. "Invariant KAM tori and global stability for Hamiltonian systems" *Z. Angew. Math. Phys.* **48** (1997) no. 1, 102–134.

[33] N.N. Nekhoroshev, "An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. II. (Russian) Trudy Sem. Petrovsk. No. 5, (1979), 5–50. An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. (Russian) Uspehi Mat. Nauk 32 (1977), no. 6(198), 5–66, 287.

[34] A.I. Neĭstadt, "The separation of motions in systems with rapidly rotating phase" J. Appl. Math. Mech. 48 (1984) no.2, 133-139.

[35] L. Niederman, "Exponential stability for small perturbations of steep integrable Hamiltonian systems", Mathematical Physics Preprint Archive 2000:http://www.ma.utexas.edu/mp

[36] J. Pöschel, "Nekhoroshev estimates for quasi-convex Hamiltonian systems" Math. Z. 213 (1993), no. 2, 187–216.

[37] J. Pöschel, "On Nekhoroshev's estimate at an elliptic equilibrium" Int. Math. Res. Not. No.4 (1999) 203-215.

[38] G.R.W. Quispel and C.P. Dyt "Volume-preserving integrators have linear error growth" Phys. Lett. A 242 (1998), no. 1-2, 25–30.

[39] S. Reich, "Backward error analysis for numerical integrators" SIAM J. Numer. Anal. 36 (1999), no. 5, 1549–1570.

[40] Y.-F. Tang, "Formal Energy of a Symplectic Scheme for Hamiltonian Systems and Its Applications I" Comp. Math. Appl. 27 (1994) no.7, 31–39.

[41] H. Rüssmann, "Nondegeneracy in the perturbation theory of integrable dynamical systems" Stochastics, algebra and analysis in classical and quantum dynamics (Marseille, 1988), 211–223, Math. Appl., 59, 1990.

[42] P. Saha and S. Tremaine, "Symplectic integrators for solar system dynamics" 1633–1640 ,Astron. Journal. 104 (4) 1992.

[43] J.M. Sanz-Serna, "Two Topics in Nonlinear Stability" In advances in Numerical Analysis Vol1, Will Light ed. Clarendon Press, Oxford, 1991.

[44] M.B. Sevryuk, "Invariant tori of Hamiltonian systems nondegenerate in the sense of Rüssmann", Dokl. Math. **53** (1996), 69-72.

[45] M.B. Sevryuk, " Reversible systems" Lecture Notes in Mathematics, 1211. Springer-Verlag, Berlin, 1986.

[46] M.B. Sevryuk, "The classical KAM theory at the dawn of the twenty-first century", preprint.

[47] Z.-j. Shang, "Resonant and Diophantine step sizes in computing invariant tori of Hamiltonian systems" Nonlinearity 13 (2000), no. 1, 299–308. "KAM theorem of symplectic algorithms for Hamiltonian systems" Numer. Math. 83 (1999), no. 3, 477–496.

[48] D. Stoffer, "On the qualitative behaviour of symplectic integrators. II. Integrable systems." J. Math. Anal. Appl. 217 (1998), no. 2, 501–520.

[49] A.V. Pronin and D.V. Treschev, "On the inclusion of analytic maps into analytic flows" Regul. Khaoticheskaya Din. 2 (1997), no. 2, 14–24.

[50] S.I. Trifonov, "Analytic diffeomorphisms as monodromy maps of analytic differential equations", Vestn. Mosk. Univ. (1986), no. 5, 70-72 (in Russian).

[51] A. Perry and S. Wiggins, " KAM tori are very sticky: rigorous lower bounds on the time to move away from an invariant Lagrangian torus with linear flow" Phys. D 71 (1994), no. 1-2, 102–121.

[52] J. Wisdom and M. Holman, "Symplectic maps for the n-body problem" Astron. J. (1991) 102, 1528-1538. ibid. "Symplectic maps for the n-body problem: stability analysis" Astron. J. (1992) 104, 2022-2029.

[53] Z. Xia, "Existence of invariant tori in volume-preserving diffeomorphisms" Ergodic Theory Dynam. Systems 12 (1992), no. 3, 621–631.