NEKHOROSHEV AND KAM STABILITIES IN GENERALIZED HAMILTONIAN SYSTEMS

YONG LI AND YINGFEI YI

Abstract. We present some Nekhoroshev stability results for nearly integrable, generalized Hamiltonian systems which can be odd dimensional and admit a distinct number of action and angle variables. Using a simultaneous approximation technique due to Lochak, Nekhoroshev stabilities are shown for various cases of quasi-convex generalized Hamiltonian systems along with concrete estimates on stability exponents. Discussions on KAM metric stability of generalized Hamiltonian systems are also made.

1. Introduction

In this paper, *generalized Hamiltonian systems* (or *Poisson-Hamilton systems*) are referred to as those defined on a Poisson manifold, in contrast to standard Hamiltonian systems defining on a symplectic manifold. To be more specific, let (M, ω^2) be a finite dimensional Poisson manifold, that is, a smooth manifold M endowed with a differential 2-form ω^2, or equivalently, a Poisson bracket $\{\cdot,\cdot\}$, which can be uniquely determined by a bundle map $\mathcal{I} : T^*M \to TM$ such that $\omega^2(\cdot, \mathcal{I}\omega^1) = \omega^1(\cdot)$, for all 1-form ω^1 on M, or equivalently, under a given local coordinate system $z = (z_i)$ on M,

$$\omega^2(\mathcal{I}\mathrm{d}f_1, \mathcal{I}\mathrm{d}f_2) = df_2(\mathcal{I}df_1) = \{f_1, f_2\} = \langle \nabla f_1, I\nabla f_2\rangle,$$

for all smooth functions f_1 and f_2 on M, where $I = I(z) = (I_{ij})$ is the matrix representation of $\mathcal{I}$ under the given coordinate, called *structure matrix*, which is a skew symmetric, smooth, matrix valued function on M satisfying the Jacobi identity:

$$\sum_m (I_{im}\frac{\partial I_{jk}}{\partial z_m} + I_{jm}\frac{\partial I_{ki}}{\partial z_m} + I_{km}\frac{\partial I_{ij}}{\partial z_m}) = 0,\ z \in M,\ \forall\ i, j, k. \tag{1.1}$$

Under the given coordinate z on M, the generalized Hamiltonian system with a given Hamiltonian function $H : M \to R$ then has the form

$$\dot{z} = I(z)\nabla H(z). \tag{1.2}$$

In the context of Darboux's theorem, an even dimensional Poisson manifold with non-degenerate Poisson structure is actually symplectic, that is, there is a local symplectic coordinate system on M, consisting of conjugate pairs $(p, q) = (p_i, q_i)$, under which the 2-form becomes $dp \wedge dq = \Sigma dp_i \wedge dq_i$ with the standard symplectic

---

1991 *Mathematics Subject Classification.* Primary 37J40.

Key words and phrases. Effective stability, Generalized Hamiltonian systems, KAM stability, stability exponents.

The first author is partially supported by NSFC grants 10531050, 10225107, National 973 project of China, and 985 program of Jilin University. The second author was partially supported by NSF grant DMS0204119 and NSFC grant 10428101.

matrix J as its structure matrix, and moreover, (1.2) becomes a standard Hamiltonian system . In general, extended Darboux's theorem holds in the sense that a Poisson manifold (M, ω^2) admits a foliation into Poisson submanifolds of constant ranks, and, each such Poisson submanifold is further foliated into symplectic submanifolds characterized by common level manifolds of the Casmir functions (see [46, 47, 49, 51, 58, 69] and a recent review article [48] for details).

Generalized Hamiltonian systems, especially those associated with multiple and degenerate Poisson structures, arise naturally in problems of celestial mechanics, fluid dynamics, plasma physics, mean field theory, chemical and biological population, optics, etc (see [14, 34, 50, 53, 58, 61] and references therein). Similar to standard Hamiltonian systems, global instability for generalized Hamiltonian systems should be generally expected and non-global stability phenomena become an important subject to study (as an example, one can certainly 'embed' the Arnold's example [3] of instability into a generalized Hamiltonian system using the extended Darboux's theorem).

The aim of the present paper is to study the stability of motions for a nearly integrable, real analytic, generalized Hamiltonian system in "action-angle" variables. Here, the *complete integrability* of a Hamiltonian N on a d-dimensional Poisson manifold (M, ω^2) is defined in the extended Liouville's sense, that is, M is locally diffeomorphic to $G \times T^n$, where $G \subset R^l$ is a bounded, connected, and closed region, $l + n = d$, and, there is a local coordinate system $(y, x) \in G \times T^n$, referred as the *action-angle* variables, with respect to which the components of $y = (y_1, y_2, \cdots, y_l)^\top \in G$ are first integrals in involution and the structure matrix is independent of x. Hence, if N is completely integrable in the above sense, then in term of the action-angle variables, it is independent of the angle variables, and moreover, the structure matrix $I(y)$ must have the form

$$\begin{pmatrix} O & B \\ -B^\top & C \end{pmatrix} \tag{1.3}$$

where $O = O_{l,l}$ is the zero matrix, $B = B(y) = B_{l,n}$, $C = C(y) = C_{n,n}$ with $C^\top = -C$. Hence, the associated integrable generalized Hamiltonian system (flow) reads

$$\begin{pmatrix} \dot{y} \\ \dot{x} \end{pmatrix} = I(y)\nabla N(y), \tag{1.4}$$

or equivalently,

$$\begin{cases} \dot{y} = 0, \\ \dot{x} = \omega(y), \end{cases}$$

where

$$\omega(y) = -B^\top(y)\frac{\partial N}{\partial y}(y).$$

Consequently, the phase space $G \times T^n$ is foliated into invariant unperturbed n-tori $\{T_y = \{y\} \times T^n : y \in G\}$ carrying parallel flows. The notion of integrability above is slightly stronger than the *integrability in broad sense* defined in [14]. Let N be a Hamiltonian on a Poisson manifold (M^d, ω^2) which is integrable in the broad sense of [14], that is, a) there are l functionally independent first integrals f_i, $i = 1, \cdots, l$; b) there is a set of $n = d - l$ Hamiltonian symmetry functions s_j, $j = 1, \cdots, n$, $\{s_j, s_{j'}\}$ =constants, $\{s_j, f_i\} = 0$, for all $i = 1, \cdots, l$, $j, j' = 1, \cdots, n$, whose Hamiltonian vector fields are linearly independent on generic points of M^d.

With the notion of [14], the set of $\{s_j\}$ corresponds to an abelian Lie algebra of Hamiltonian symmetries which preserves the set of first integrals $\{f_i\}$. Now, using the standard proof of the classical Liouville-Arnold theorem (e.g., [4]), it is clear that if the level sets $M_c = \{f_i = c_i, i = 1, \cdots, l\}$, for $c = (c_1, c_2, \cdots, c_l)$ lying in a bounded closed region $G \subset R^l$, are compact connected, then they must be n dimensional tori carrying parallel flows with respect to a global coordinate $x \in T^n$. Hence by letting $y = c \in R^l$ both the Hamiltonian N and the structure matrix I under the action-angle coordinate $(y, x) \in G \times T^n$ are independent of x. Now, if we further assume that the first integrals $\{f_i\}$ are in involution (which is not required in [14]) then the two sets $\{f_i\}, \{s_j\}$ can well be overlapped and the structure matrix $I(y)$ has the form (1.9), hence the Hamiltonian N is integrable in the extended Liouville sense defined above. We refer the readers to [13, 14, 22, 36] for more discussions on action-angle variables in Hamiltonian systems.

In this paper, we restrict our attention to a flat Poisson manifold $(G \times T^n, \omega^2)$ and consider a completely integrable Hamiltonian $N(y)$ with respect to a set of action-angle variables $(y, x) \in G \times T^n$, where $G \subset R^l$ is a bounded, connected, and closed region, l, n are positive integers. Let $I(y)$ be the structure matrix associated with ω^2 under the action-angle variables. Adding a perturbation $P(y, x)$ to $N(y)$, one is led to the following nearly integrable generalized Hamiltonian

$$H(y, x) = N(y) + \varepsilon P(y, x), \tag{1.5}$$

whose corresponding nearly integrable generalized Hamiltonian system reads

$$\begin{pmatrix} \dot{y} \\ \dot{x} \end{pmatrix} = I(y)\nabla(N(y) + \varepsilon P(y, x)), \tag{1.6}$$

where $\varepsilon > 0$ is a small parameter. Throughout of the paper, we assume that $I(y), N(y), P(y, x)$ are real analytic in a complex neighborhood $G \times T^n + \delta$ of $G \times T^n$ for a fixed $\delta > 0$.

Clearly, when $n = l$ and $I \equiv J$ - the standard symplectic matrix, $(G \times T^n, \omega^2)$ becomes the usual symplectic manifold and (1.6) becomes a standard nearly integrable Hamiltonian system:

$$\begin{pmatrix} \dot{y} \\ \dot{x} \end{pmatrix} = J\nabla(N(y) + \varepsilon P(y, x)). \tag{1.7}$$

In general, if $l > n$ or $n + l$ is odd, then the structure matrix I becomes singular on G, or the associated 2-form ω^2 becomes degenerate (hence not symplectic in the usual sense).

It is well known that a nearly integrable Hamiltonian system of form (1.7) with higher degree of freedom (that is, $n > 2$) can be globally unstable in the sense that in an arbitrary open subset E of G there is a phase orbit $(y(t), x(t))$ with $y(0) \in E$ such that $|y(T) - y(0)| = O(1)$ at some time T. In fact, as observed by Arnold in [3], in the resonance zone of a nearly integrable Hamiltonian system of higher degree of freedom, the slow variables of the motion can be drifted arbitrarily far away from their initial positions and randomly wander between invariant tori, resulting in a rather complicated slow diffusion process known as the *Arnold Diffusion* (see also [12, 24, 25, 43]). Nevertheless, a nearly integrable Hamiltonian system can enjoy two type of non-global stabilities: the *KAM metric stability* (or almost everywhere perpetual stability) and the *Nekhoroshev stability* (or exponential stability). On one hand, the classical KAM theory due to Kolmogorov ([37]), Arnold ([2]), and Moser

([54]) asserts under certain non-degenerate conditions on N that the Hamiltonian (1.7) is metrically stable in the sense that there is a family of Cantor-like sets $G_\varepsilon \subset G$ with $|G \setminus G_\varepsilon| \to 0$ as $\varepsilon \to 0$ such that $\sup_{t\in R} |y(t) - y(0)| \to 0$ as $\varepsilon \to 0$ uniformly for $y(0) \in G_\varepsilon$. On the other hand, the celebrated work of Nekhoroshev ([56]) says that if N is a "steep" function then (1.7) is effectively stable over exponentially long time, that is, there are positive constants a, b, c, depending on n, G, N, P and the steepness indices, with $a, b \to 0$ as $n \to \infty$, such that for all $y(0) \in G$,

$$|y(t) - y(0)| \le c\varepsilon^b \quad \text{as} \quad |t| \le \exp(c\varepsilon^{-a}).$$

Hence, in a Nekhoroshev stable system the slow variables after perturbation evolute like those in the unperturbed system over exponentially long time beyond the time scale in the classical perturbation theory. Originated from the study of stability for Solar system, the theory of Nekhoroshev stability has led to deep understandings of physical insight to Hamiltonian systems or Hamiltonian networks arising in classical and celestial mechanics, bio-science, and physics. We refer the readers to [5, 6, 7, 10, 11, 16, 26, 27] and references therein for recent progress in the subject.

In connection with the study of Arnold diffusions in generic Hamiltonian systems of higher degree of freedom, the estimate of stability radius and time (or equivalently, the *stability exponents* a, b) become evidently important. In particular, the estimate of stability time provides an upper bound for the diffusion time at which possible Arnold diffusion would take place, and, the estimate of the exponent a characterizes the *diffusion rate* (speed of possible Arnold diffusion) at which a stochastic orbit travels in the resonance zone (see [18] for more discussions). The estimate in the steepest case was originally considered by Benettin, Gallavotti, Galgani and Giorgilli in [8, 9] for (1.7) in which the unperturbed Hamiltonian $N(y)$ is strictly convex in a bounded closed region; and, the estimate $a \approx 1/(n^2 + n)$ was obtained. This estimate was improved by Lochak ([41, 42]) to $a \approx 1/(2n + 2)$ by introducing a novel technique involving the simultaneous approximation and the Dirichlet's approximation theorem, and later further improved by Lochak and Neishtadt ([44]), Pöschel ([64]) independently to

$$a = b = \frac{1}{2n}. \tag{1.8}$$

In fact, (1.8) was shown to be true in [44, 64] for the quasi-convex case, that is, the unperturbed Hamiltonian $N(y)$ in (1.7) is strictly convex on any energy surface within a bounded closed region of G. Certain heuristic arguments and numerical estimates of the Arnold diffusion rates seem to suggest that the estimate (1.8) is optimal ([44]) in the convex and quasi-convex cases. A rather general estimate of the stability exponents is made in a recent work of Niederman ([57]) who considered standard, steep, nearly integrable Hamiltonian systems of form (1.7) and obtained a concrete estimate $a = b = 1/(2np_1 \cdots p_n)$, where $\{p_1, \cdots, p_n\}$ are steepness indices of $N(y)$ in (1.7).

With respect to the generalized Hamiltonian system (1.6), KAM theory has been developed for (1.6) in [17, 31, 32, 40, 55, 60], under various non-degenerate conditions, which asserts the persistence of the majority of the unperturbed n-tori hence the KAM metric stability of (1.5) as ε small. Even with respect to the standard symplectic case, such KAM theory has its own rights because it actually gives a way to show the persistence of higher dimensional tori in standard Hamiltonian systems.

Indeed, some of degenerate generalized Hamiltonian systems are absolutely not generated by extending classical Hamiltonian systems. This can be illustrated by (1.6) with $y \in R^1, x \in T^3$ and

$$I(y) = \begin{pmatrix} 0 & \alpha & \beta & \gamma \\ -\alpha & 0 & 0 & 0 \\ -\beta & 0 & 0 & 0 \\ -\gamma & 0 & 0 & 0 \end{pmatrix}, \tag{1.9}$$

where constants α, β, γ are rationally independent.

Based on the simultaneous approximation technique developed in in [41, 42, 44], we will show in this work under certain quasi-convexity of the unperturbed Hamiltonian and certain non-degeneracy of the frequency map that (1.6) is also Nekhoroshev stable and an estimate similar to (1.8) holds. Such a result is already significant even for the perturbation of a standard integrable Hamiltonian, simply because it allows non-standard Hamiltonian perturbations as long as they preserve a Poisson structure that is invariant to the unperturbed standard integrable Hamiltonian system ([13, 21, 59]). In general, it would be interesting and also important to be able to characterize steepness indices for generalized Hamiltonians and obtain estimate of stability exponents similar to that [57].

We note that Dirichlet's theorem (Lemma 2.1 in section 2 below) plays a key role in applying the simultaneous approximation technique but it does not holds on general manifolds even those characterized by the Casmir functions of a generalized Hamiltonian system.

Effective stabilities in the generalized setting present some unique features, because not only do they depend on the steepness such as convexity and quasi-convexity of a unperturbed Hamiltonian, but also on the nature of its associated structure matrix, and further on certain non-degeneracy of the resulting frequency map, which, unlike the standard Hamiltonian case, is not a direct consequence of the non-degeneracy of the unperturbed Hamiltonian. Indeed, arbitrary high degeneracy of the frequency map caused by the degeneracy of the corresponding Poisson structure in a quasi-convex, nearly integrable, generalized Hamiltonian system is another obstacle to its Nekhoroshev stability, as demonstrated by a non-autonomous example in Section 2.

Also differing from the standard Hamiltonian case, a quasi-convex, Nekhoroshev stable generalized Hamiltonian system can be non-KAM stable (see Section 3). We conjecture that the converse is also true, that is, one can also have a KAM stable generalized Hamiltonian system which is not Nekhoroshev stable. In fact, in the case $l \neq n$, the two stabilities seems to lie in different domains in general. Taking the case $l < n$ for example, the KAM stability of a generalized Hamiltonian system requires the non-resonancy of all components of the frequency map which is usually ensured by non-constancy of the matrix B, while the Nekhoroshev stability requires the resonancy of any $l + 1$ frequency components which typically occurs when the matrix B is constant with integer entries. We will discuss these issues in the paper along with some reviews on the KAM stability of generalized Hamiltonian systems.

The paper is organized as follows. Section 2 deals with the Nekhoroshev stabilities of (1.6) for the case that the unperturbed Hamiltonian is quasi-convex, while the frequency map can be either non-degenerate or degenerate. Our arguments are based on the Dirichlet theorem, the simultaneous approximation lemma, and a reduction lemma near the resonance. In Section 3, we review a KAM metric stability result from [40] and also present an iso-energetic KAM stability result for nearly integrable generalized Hamiltonian systems. Using these results, generalized Hamiltonians with either $l = 1$ or $n = 1$ are shown to be perpetually stable. Perpetual stability is also discussed for the case $l = 2$ for iso-energetic KAM stable systems. As an application of both KAM and Nekhoroshev stability results, we consider the perturbation problems of three dimensional divergence free vector fields arising naturally as the Euler fluid particle path flows in which perturbations need not preserve the volume preserving symmetry. Finally, we give some discussions on the compatibility of the two stability types for quasi-convex generalized Hamiltonian systems.

2. Effective stability

We say that (1.6) is *Nekhoroshev stable* if it is effectively stable over exponential long time, that is, there are positive constants c_0, a, b independent of ε, where a, b are referred to as *stability exponents*, such that as $\varepsilon > 0$ sufficiently small every solutions $(y(t), x(t))$ of (1.6) with $y(0) \in G$ satisfy the estimate

$$|y(t) - y(0)| \leq c_0 \varepsilon^b,$$

as $|t| \leq \exp\left(c_0 \varepsilon^{-a}\right)$,

Similar to the standard Hamiltonian case ([44, 64]), we assume that the unperturbed Hamiltonian N in (1.6) is *quasi-convex* in the sense that

QC) there is a $\rho = \rho(\delta, G, N) > 0$ such that

$$|\omega(y)| \geq \rho, \ \ |\langle \xi, \frac{\partial^2 N}{\partial y^2}(y)\xi \rangle| > \rho|\xi|^2,$$

for all $y \in \text{Re}(G + \delta)$ and all $\xi \in R^l$ with $\langle \partial N / \partial y(y), \xi \rangle = 0$.

Due to the general nature of the generalized Hamiltonian systems in particular the fact that non-degeneracy of the frequency maps needs not follow from that of the Hamiltonian functions, we assume that the unperturbed frequency map in (1.6) is *quasi-convex reducible* in the sense that

QR) $\omega(y) = \tilde{B}\tilde{\omega}(y)$, where $\tilde{B}$ is an $n \times l$ constant matrix, $\tilde{\omega} : G + \delta \to R^l$ is such that there are $\rho' = \rho'(\delta, G, N) > 0$ and $r' = r'(\delta, G, N) > 0$ for which

$$|\tilde{\omega}(y) - \tilde{\omega}(y_0)| > \rho' |y - y_0|,$$

whenever $y, y_0 \in \text{Re}(G + \delta)$, $0 < |y - y_0| \leq r'$, and $\langle \partial N / \partial y(y_0), y - y_0 \rangle = 0$.

We say that a matrix is *essentially integral* if it is a constant multiple of a matrix with integer entries.

Our main results states as follows.

Theorem 1. *Assume* QC), QR) *and that B is a constant matrix. Then the following holds.*

1) *If $\tilde{B}$ is essentially integral, then* (1.6) *is Nekhoroshev stable with the stability exponents*

$$a = b = \frac{1}{2l}.$$

2) *If* $\text{rank}\tilde{B} = n$, *then* (1.6) *is Nekhoroshev stable with the stability exponents*

$$a = b = \frac{1}{2n}.$$

In the case that B is a constant matrix, we let $\tilde{B} = -B^\top$ and $\tilde{\omega}(y) = \partial N/\partial y(y)$. Then an immediate consequence of Theorem 1 is the following.

Corollary 1. *Assume* QC) *and that B is a constant matrix.*

1) *If B is essentially integral, then* (1.6) *is Nekhoroshev stable with stability exponents*

$$a = b = \frac{1}{2l}.$$

2) *If* $\mathrm{rank} B = n$*, then* (1.6) *is Nekhoroshev stable with stability exponents*

$$a = b = \frac{1}{2n}.$$

In the case of standard Hamiltonian systems, $l = n$ and $-B$ =identity, hence results in both parts of the above corollary coincide with those of [44, 64] for standard Hamiltonian systems with quasi-convex unperturbed Hamiltonians.

We note that both parts 1) of the above theorem and corollary have no restriction on l and n but do require either $\tilde{B}$ or B to be essentially integral. Such an essentially integral condition, which we believe to be necessary in general, indicates a kind of sensitivity of Nekhoroshev stability of a near integrable Hamiltonian on its Poisson structure - a phenomenon may only be seen in non-standard Hamiltonian systems.

The conditions in parts 2) of the above theorem and corollary automatically imply that $l \geq n$. In fact, the constancy of B can be relaxed in those cases.

Corollary 2 ($l \geq n$). *Assume* QC) *and that* $\mathrm{rank} B \equiv n$ *on G. Then* (1.6) *is Nekhoroshev stable with stability exponents*

$$a = b = \frac{1}{2n}.$$

Proof. Let $B(y)$ have constant rank n_0 over G. Then an argument similar to the extended Darboux theorem ([58, 69]) yields that near any $(y_0, x_0) \in G \times T^n$ there are local coordinates on $G \times T^n$, again denote by $(y, x) = (y_1, \cdots, y_l, x_1, \cdots, x_n)$, such that $\{y_i, y_j\} = \{y_i, x_k\} = \{x_k, x_{k'}\} = 0, \{y_k, x_k\} = 1$, for all $i, j = 1, \cdots, l$, $k, k' = 1, \cdots, n_0$, and $i \neq k$. Furthermore, the commutativity and periodicity of the Hamiltonian flows associated with $\{y_k\}_{k=1}^{n_0}$ allow us to extend the local coordinates $(x_1, \cdots, x_{n_0})$ globally on T^{n_0} so that the above in involution properties remain unchanged (see [22, 45]). Thus, if $n_0 = n$ (hence $l \geq n$ is necessary) then near each $y_0 \in G$ there are local coordinates $y \in G$ and global coordinates $x \in T^n$ under which the matrix B becomes $(I_n, 0)^\top$, where I_n is the $n \times n$ identity matrix. Since the quasi-convexity of N is independent of coordinates, the corollary immediately follows from Corollary 1 2). □

We now discuss some consequences of Theorem 1 in the case $l < n$. We observe that if QR) holds, then the frequency map $\omega(y)$ is always degenerate (not one to one) when $l > n$ but can well be non-degenerate when $l \leq n$. Also, the condition for $\tilde{B}$ being essentially integral implies that any $n_* + 1$, where $n_* = \mathrm{rank}\tilde{B}$, components of the frequency map $\omega(y)$, as functions on $G + \delta$, are rationally dependent, leading to the everywhere resonancy of the frequency map. We thus require the frequency map to be non-degenerate but resonant in the sense that

RD) any $l+1$ components of the frequency map $\omega(y)$, as functions on $G+\delta$, are rationally dependent;

ND) there are l rationally independent components $\{\omega_{i_1}(y), \cdots, \omega_{i_l}(y)\}$ of $\omega(y)$ such that $\tilde{\omega}(y) = (\omega_{i_1}(y), \cdots, \omega_{i_l}(y))^\top$ is non-degenerate in the sense of QR).

In the case that B is a constant matrix, we have the following.

Corollary 3 ($l < n$). *Assume* QC), RD), ND) *and that B is a constant matrix. Then* (1.6) *is Nekhoroshev stable with stability exponents*

$$a = b = \frac{1}{2l}.$$

Proof. Let $\tilde{\omega}(y)$ be as in ND). Then it is clear that $\omega(y) = \tilde{B}\tilde{\omega}(y)$ for some constant, essentially integral matrix $\tilde{B}$. Hence the corollary follows from Theorem 1 1). □

We can also assume the following non-degeneracy:

ND') there is a $\rho' = \rho'(\delta, G, N) > 0$ such that

$$|\omega(y) - \omega(y_0)| > \rho'|y - y_0|,$$

for any $y, y_0 \in \mathrm{Re}(G+\delta)$, $y \neq y_0$.

Corollary 4 ($l < n$). *Assume* QC), RD), ND') *and that B is a constant matrix. Then* (1.6) *is Nekhoroshev stable with the stability exponents*

$$a = b = \frac{1}{2l}.$$

Proof. The resonant condition RD) is equivalent to that there is a rank $n-l$ subgroup g of Z^n such that

$$\langle k, \omega(y)\rangle = 0, \ \ \forall\, k \in g.$$

It follows that there is an $n \times n$ unimodular matrix

$$K_0 = (K_1, K_2),$$

with integer entries, where K_1 is $n \times l$ and K_2 is $n \times (n-l)$, such that $K_2^\top \omega(y) = 0$. Let $\tilde{\omega}(y) = K_1^\top \omega(y)$, $\tilde{B}$ be the left $n \times l$ block of $(K_0^\top)^{-1}$. Then $\tilde{B}$ has integer entries and $\omega(y) = \tilde{B}\tilde{\omega}(y)$. Thus if ND') holds, then $\tilde{\omega}(y)$ satisfies QR). The corollary again follows from Theorem 1 1). □

The non-degenerate conditions QC), ND) and ND') of the frequency map are less restrictive than those needed for KAM stabilities of generalized Hamiltonian systems. Hence a perturbed generalized Hamiltonian system can be both KAM and Nekhoroshev stable even when $l \neq n$. In particular, when $l < n$ this leads to the persistence of higher dimensional tori (even on symplectic manifolds) with fewer parameters. However, in the case $l < n$, with the additional condition QR) or RD), a generic nearly integrable generalized Hamiltonian system is Nekhoroshev stable but not KAM stable, due to the everywhere resonancy of the frequency map. This demonstrates a kind of non-compatibility of the two type of stabilities in the generalized Hamiltonian setting especially when $l < n$. We refer the readers to Section 4 for more discussions in this regard.

In all results above, the stability constants c_0 depend on $I, N, P, l, n, G, \delta, \rho, \rho', r'$.

2.1. **An example.** We give an example to illustrate the fact that if B is a non-constant matrix and QR) fails then high degeneracy of the frequency map can yield the the global instability of the corresponding nearly integrable Hamiltonian system.

Consider a non-autonomous Hamiltonian of form

$$H(y,x) = N(y) + \varepsilon P(y,x,\hat{\omega}t), \tag{2.1}$$

$$\begin{pmatrix} \dot{y} \\ \dot{x} \end{pmatrix} = I(y)\nabla(N(y) + \varepsilon P(y,x,\hat{\omega}t)), \tag{2.2}$$

where $(y,x) \in R^l \times T^n$, ε is sufficiently small, N, P are real analytic on $R^l \times T^n \times T^{n'}$, and $\hat{\omega} \in R^{n'}$ is a fixed n' dimensional non-zero frequency for some positive integer n'.

To construct the example, we let $l = 1$, $n = 2$,

$$B(y) = (\frac{\alpha}{y}, \frac{\beta}{y}),$$

for some constants $\alpha \neq \beta$, $C = 0$, $N(y) = (1/2)(y^2)$, and $P = \sin 2\pi(x_1 - x_2)\cos 2\pi(\alpha - \beta)t$. The matrix I defined through B, C above in the form (1.9) is easily seen to satisfy the Jacobi identity hence defines a structure matrix. We also let the domain G be an interval away from 0.

Now, the frequency map $\omega(y) \equiv -(\alpha,\beta)^\top$ is everywhere degenerate hence the condition QR) fails. By integrating the respective motions of the Hamiltonian, one finds that whenever $x_1(0) = x_2(0)$, then $y^2(t) - y^2(0) = 2\pi(\alpha-\beta)\varepsilon t + \varepsilon \sin 2\pi(\alpha - \beta)t\cos 2\pi(\alpha-\beta)t$. Hence the Hamiltonian system is non-Nekhoroshev.

We also note that no invariant tori of any dimension can exist in this example when $\varepsilon \neq 0$. Hence the Hamiltonian system is not KAM metric stable either.

2.2. **Proof of Theorem 1.** The rest of the section is devoted to the proof of Theorem 1, following the framework and techniques developed in [44]. As to be shown in Section 3 that all nearly integrable generalized Hamiltonian systems with either $l = 1$ or $n = 1$ will be perpetually stable, we assume without loss of generality that $l > 1, n > 1$ for the rest of this section.

We first prepare some technical lemmas to be needed in the proof.

For $\alpha \in R^k$, we denote

$$\langle \alpha \rangle \equiv \inf_{\xi \in Z^k} |\alpha - \xi|_\infty = \inf_{\xi \in Z^k} \{ \max_{i=1,\cdots,k} |\alpha_i - \xi_i| \}.$$

Lemma 2.1. (Dirichlet's theorem) *Let $\omega \in R^k$ and let $Q > 1$ be a real number. Then there is an integer q with $1 \leq q < Q$ such that*

$$\langle q\omega \rangle \leq Q^{-\frac{1}{k}}.$$

Proof. See [68]. □

In the sequel, all $c_1 - c_5$ are positive constants depending only on $I, N, P, l, n, G, \delta, \rho, \rho', r'$. We also use c to denote any intermediate constant which only depends on the above.

Define

$$Q_k = \varepsilon^{-\frac{k-1}{2k}}, \quad k > 1.$$

Lemma 2.2.

a) *Assume the conditions of* Theorem 1 1). *Then there are constants* $c_1 > 0$, $c_2 \geq 1$ *such that for any given* $y_0 \in G$ *there exist a* T *with* $c_1^{-1} \leq T < c_1 Q_l$, *an* $\omega_* \in R^n$ *with* $\omega_* T \in Z^n$, *and a* $y_* \in \mathrm{Re}(G+\delta)$, *satisfying*

$$|y_0 - y_*| < \frac{c_2}{T(Q_l)^{\frac{1}{l-1}}},$$
$$\omega_* = \omega(y_*).$$

b) *Assume the conditions of* Theorem 1 2). *Then there are constants* $c_1 > 0$, $c_2 \geq 1$ *such that for any given* $y_0 \in G$, *there exist a* T *with* $c_1^{-1} \leq T < c_1 Q_n$, *an* $\omega_* \in R^n$ *with* $\omega_* T \in Z^n$, *and a* $y_* \in \mathrm{Re}(G+\delta)$, *satisfying*

$$|y_0 - y_*| < \frac{c_2}{T(Q_n)^{\frac{1}{n-1}}},$$
$$\omega_* = \omega(y_*).$$

Proof. To prove a), we assume without loss of generality that $\tilde{B}$ consists of integer entries. Let $c_1 = \max\{(\min_{1\leq i\leq l}|\tilde{\omega}_i|_G)^{-1}, \max_{1\leq i\leq l}|\tilde{\omega}_i|_G\}$, where $\tilde{\omega}_i(y)$, $i = 1, 2, \cdots, l$, are components of $\tilde{\omega}(y)$. For given $y_0 \in G$, we denote

$$\tilde{\omega}_0 = \tilde{\omega}(y_0) = (\omega_{01}, \omega_{02}, \cdots, \omega_{0l})^\top.$$

Since $\tilde{\omega}_0 \neq 0$, we assume without loss of generality that $0 < \omega_{01} = \max_{1\leq i\leq l}\{|\omega_{0i}|\}$. Then an application of the Dirichlet theorem to $\hat{\omega}_0 = (1/\omega_{01})(\omega_{02}, \cdots, \omega_{0l})$ yields that there is an $\hat{\omega}_* \in R^l$ and an integer $1 \leq q < Q_l$ such that

$$|\hat{\omega}_* - \hat{\omega}_0| \leq \frac{c\sqrt{l-1}}{q(Q_l)^{\frac{1}{l-1}}},$$

for some constant c. Let $T = q/\omega_{01}$, $\tilde{\omega}_* = \omega_{01}(1, \hat{\omega}_*)^\top$. Then $c_1^{-1} \leq T \leq c_1 Q_l$, $\tilde{\omega}_* T \in Z^l$, and,

$$|\tilde{\omega}_* - \tilde{\omega}_0| \leq \frac{c\sqrt{l-1}}{T(Q_l)^{\frac{1}{l-1}}},$$

for some constant c. By QR), $\tilde{\omega}$ is an open mapping on the cone $V = \{y : |\langle \partial N/\partial y(y_0), y - y_0\rangle| < \eta'|y - y_0| < \eta' r'\}$. Hence as ε sufficiently small, there is a $y_* \in V$ such that $\tilde{\omega}_* = \tilde{\omega}(y_*)$. It also follows from QR) that

$$\rho'|y_* - y_0| \leq |\tilde{\omega}_* - \tilde{\omega}_0| \leq \frac{c\sqrt{l-1}}{T(Q_l)^{\frac{1}{l-1}}}.$$

Let $\omega_* = \tilde{B}\tilde{\omega}_*$. Then $\omega(y_*) = \tilde{B}\tilde{\omega}(y_*) = \omega_*$ and $T\omega_* \in Z^n$.

In the case of b), a similar application of the Dirichlet theorem as above yields the existence of a $c_1 > 0$, an $\omega_* \in R^n$, a T with $c_1^{-1} \leq T < c_1 Q_n$, such that $\omega_* T \in Z^n$, and,

$$|\omega_* - \omega(y_0)| \leq \frac{c\sqrt{n-1}}{T(Q_n)^{\frac{1}{n-1}}}, \tag{2.3}$$

for some constant c. Note that $\tilde{B}$ is right invertible in this case. Hence

$$\tilde{\omega}(y_0) = \tilde{B}^{-1}\omega(y_0) + v_0,$$

for some $v_0 \in \ker \tilde{B}$, where $\tilde{B}^{-1}$ denotes the right inverse of $\tilde{B}$. By QR) and the argument in the proof of a), there is a cone V near y_0, and a $y_* \in V$, such that

$$\tilde{\omega}(y_*) = \tilde{B}^{-1}\omega_* + v_0.$$

Then

$$\omega(y_*) = \tilde{B}\tilde{\omega}(y_*) = \omega_*,$$

and it also follows from QR) that

$$|y_* - y_0| \leq \frac{1}{\rho'}|\tilde{\omega}(y_*) - \tilde{\omega}(y_0)| \leq \frac{1}{\rho'}|\tilde{B}^{-1}||\omega_* - \omega(y_0)| \leq \frac{c\sqrt{n-1}}{T(Q_n)^{\frac{1}{n-1}}}.$$

□

Lemma 2.3. *Assume* QC) *and let y_0, y_* be such that $|y_0 - y_*| < \mu$ for a sufficiently small $\mu > 0$. Then there is a constant $c_3 \geq 1$ such that the set of the intersection of the plane $\langle \frac{\partial N}{\partial y}(y_*), y - y_0\rangle = 0$ and the surface $\{N(y) = N(y_0)\}$ is contained in a ball of radius $c_3\mu$ centered at y_*.*

Proof. Let y lie in the intersection of the plane $\langle \partial N/\partial y(y_*), y - y_0\rangle = 0$ and the surface $\{N(y) = N(y_0)\}$. By Taylor's formula, we have that

$$N(y) = N(y_0) + \langle y - y_0, \frac{\partial N}{\partial y}(y_0)\rangle + \frac{1}{2}\langle y - y_0, \frac{\partial^2 N}{\partial y^2}(y_1)(y - y_0)\rangle,$$

where $y_1 = (1-s)y + sy_0$ for some $0 < s < 1$. Hence

$$\begin{aligned}\frac{1}{2}\langle y - y_0, \frac{\partial^2 N}{\partial y^2}(y_1)(y - y_0)\rangle &= -\langle \frac{\partial N}{\partial y}(y_0), y - y_0\rangle = \langle \frac{\partial N}{\partial y}(y_*) - \frac{\partial N}{\partial y}(y_0), y - y_0\rangle \\ &= \langle \frac{\partial^2 N}{\partial y^2}(y_2)(y_* - y_0), y - y_0\rangle, \qquad (2.4)\end{aligned}$$

for some $y_2 = (1-t)y_0 + ty_*$. Since $|y - y_0|$ is small as long as μ is, we have by letting μ sufficiently small that

$$|\langle \frac{\partial N}{\partial y}(y_1), y - y_0\rangle| = |\langle \frac{\partial N}{\partial y}(y_1) - \frac{\partial N}{\partial y}(y_*), y - y_0\rangle|.$$

It follows from (2.4) and QC) that

$$\frac{1}{2}\rho|y - y_0|^2 \leq M|y - y_0||y_0 - y_*|,$$

where

$$M = |\frac{\partial^2 N}{\partial y^2}|_G.$$

Hence

$$|y - y_0| < \frac{2M}{\rho}\mu,$$

and consequently,

$$|y - y_*| \leq |y - y_0| + |y_* - y_0| < (\frac{2M}{\rho} + 1)\mu.$$

□

Let y_0, y_*, ω_*, T be as in Lemma 2.2. To unify the notations, we let

$$Q = \varepsilon^{-\frac{n_*-1}{2n_*}}, \quad R = \frac{3c_2c_3}{TQ^{\frac{1}{n_*-1}}},$$
$$\omega_* = \omega(y_*) = -B^\top \frac{\partial N}{\partial y}(y_*),$$

where

$$n_* = \begin{cases} l, & \text{in the case of Theorem 1 1)}, \\ n, & \text{in the cases of Theorem 1 2)}. \end{cases}$$

With these notations, we have $c_1^{-1} \le T < c_1 Q$, $\omega_* T \in Z^n$, and

$$|y_* - y_0| < R.$$

Below, for a given function $F : G \times T^n \to R^1$, we let $\bar{F}(y, x)$ be the time average of $F(y, x + \omega_* t)$ over the interval $[0, T]$, that is,

$$\bar{F}(y, x) = \frac{1}{T} \int_0^T F(y, x + \omega_* t) \mathrm{d}t.$$

We also denote $\tilde{F}$ as $F - \bar{F}$.

By suspending the constant term $N(y_*)$ we can rewrite (1.5) as

$$H = N_0(y) + \varepsilon \bar{P}(y, x) + \varepsilon \tilde{P}(y, x), \tag{2.5}$$

where

$$N_0(y) = N(y) - N(y_*) = \langle \frac{\partial N}{\partial y}(y_*), y - y_* \rangle + h(y - y_*),$$
$$h(y - y_*) = O(|y - y_*|^2).$$

The following lemma makes use of the arguments in [44] and the iterative method introduced in [40] for generalized Hamiltonian systems.

Lemma 2.4. *Let*

$$D = \{(y, x)) \in G \times T^n + \delta : |y - y_*| < \frac{R}{3}, \ |\mathrm{Im} x| < \frac{\delta}{3}\},$$
$$D_* = \{(Y, X) \in G \times T^n + \delta : |Y - y_*| < \frac{R}{2}, \ |\mathrm{Im} X| < \frac{\delta}{2}\}.$$

Then under the conditions of Lemma 2.2 *there exists a real analytic, canonical, near identity change of variables* $\Psi : D \to D_* : (y, x) \mapsto (Y, X)$ *which transforms* (2.5) *into the form*

$$H_* = \langle \frac{\partial N}{\partial y}(y_*), Y - y_* \rangle + h(Y - y_*) + \varepsilon \bar{P}_*(Y, X) + \varepsilon \tilde{P}_*(Y, X), \ \ (Y, X) \in D_*,$$

where

$$|\varepsilon \frac{\partial \tilde{P}_*}{\partial X}|_{D_*} < \exp(-c_4 \varepsilon^{-\frac{1}{2n_*}}),$$
$$|\varepsilon P_*|_{D_*} \le c_4 \varepsilon T,$$
$$|\varepsilon \frac{\partial P_*}{\partial Y}|_{D_*} < c_4 R, \ \ |\varepsilon \frac{\partial P_*}{\partial X}|_{D_*} < c_4 R^2,$$

for some constant $c_4 > 0$. *Moreover, there is a constant* $c_5 > 0$ *such that for any* $(y, x) \in D$, $(Y, X) = \Psi(y, x)$,

$$|Y - y| < c_5 \varepsilon T,$$
$$|X - x| < c_5 \frac{\varepsilon}{R}.$$

Proof. We assume without loss of generality that $\delta < 1$. Let ϵ be sufficiently small. Then $R \ll 1$ and

$$\varepsilon \leq R^2. \tag{2.6}$$

Define

$$\begin{aligned}
M &= \max\{1, |\partial^{\imath} I|_{G+\delta}, |\partial^{\imath}\omega|_{G+\delta}, 2|\partial^{\jmath} P|_{G\times T^n+\delta} : |\imath|,\ |\jmath| \leq 1\},\\
M_* &= \frac{600M^3}{\delta},\\
K &= \frac{1}{[M_* T R]},\\
\gamma &= \frac{R}{10K},\\
\sigma &= \frac{\delta}{10K},\\
D_j &= \mathcal{D}_j \times \mathcal{O}_j,\\
D'_j &= \mathcal{D}'_j \times \mathcal{O}'_j,\\
\mathcal{D}_j &= \{y \in G + \delta : |y - y_*| < \frac{4}{5}R - j\gamma\},\\
\mathcal{D}'_j &= \{y \in G + \delta : |y - y_*| < \frac{4}{5}R - (j + \frac{1}{2})\gamma\},\\
\mathcal{O}_j &= \{x \in T^n + \delta : |\mathrm{Im} x| < \frac{4}{5}\delta - j\sigma\},\\
\mathcal{O}'_j &= \{x \in T^n + \delta : |\mathrm{Im} x| < \frac{4}{5}\delta - (j + \frac{1}{2})\sigma\},\\
\zeta_j &= 5M^3 \varepsilon T R \left(\frac{1}{2}\right)^j,
\end{aligned}$$

for $j = 0, 1, \cdots, K - 1$.

The proof amounts to the construction of K canonical transformations Φ_i, $i = 0, 1, \cdots, K-1$, which successively transform the Hamiltonian (2.5) into the desired form. We first show such a construction for a typical step. Suppose at some ith step, where $i = 0, 1, \cdots, K - 1$, the Hamiltonian has the form

$$H_i = N_0(y) + \varepsilon P_i(y, x) = \langle \frac{\partial N}{\partial y}(y_*), y - y_* \rangle + h(y - y_*) + \varepsilon \bar{P}_i(y, x) + \varepsilon \tilde{P}_i(y, x). \tag{2.7}$$

To unify the notation, we let $H_0 = H, P_0 = P$.

Let

$$S_i(y, x) = \frac{1}{T} \int_0^T t \tilde{P}_i(y, x + \omega_* t) \mathrm{d}t. \tag{2.8}$$

Since $\omega_* T \in Z^n$, $P_i(y, x + \omega_* t)$ is T-periodic in t and S_i is easily seen to satisfy the homological equation

$$\langle \omega_*, \frac{\partial S_i}{\partial x} \rangle - \tilde{P}_i = 0. \tag{2.9}$$

Let $\phi_i^t(y, x)$ be the flow generated by the vector field $\varepsilon I(y) \nabla S_i(y, x)$, that is,

$$\phi_i^t(y, x) = \begin{pmatrix} y \\ x \end{pmatrix} + \varepsilon \int_0^t (I \nabla S_i) \circ \phi_i^s(y, x) \mathrm{d}s. \tag{2.10}$$

Then the transformation

$$\begin{pmatrix} y \\ x \end{pmatrix} = \Phi_i(Y, X) =: \phi_i^1(Y, X) = \begin{pmatrix} Y \\ X \end{pmatrix} + \varepsilon \int_0^1 (I \nabla S_i) \circ \phi_i^t(Y, X) \mathrm{d}t$$

is canonical in the sense that

$$\nabla \Phi_i(Y, X)^\top I(Y) \nabla \Phi_i(Y, X) = I(y),$$

hence it preserves the Poisson structure defined by I. Let

$$\begin{aligned} P_{i+1}(Y, X) = \bar{P}_i(y, x) &+ \int_0^1 (\tilde{P}_i \circ \phi_i^1(Y, X) - \tilde{P}_i \circ \phi_i^t(Y, X)) \mathrm{d}t \\ &+ \int_0^1 \langle \omega(y_*) - \omega(y), \frac{\partial S_i}{\partial x} \rangle \circ \phi_i^t(Y, X) \mathrm{d}t. \end{aligned} \tag{2.11}$$

It follows from (2.7), (2.9) that

$$\begin{aligned} &H_{i+1}(Y, X) =: H_i \circ \Phi_i(Y, X) = N_0(Y) + \varepsilon \int_0^1 \{N_0, S_i\} \circ \phi_i^t(Y, X) \mathrm{d}t \\ &+\varepsilon P_i \circ \Phi_i(Y, X) = N_0(Y) - \varepsilon \int_0^1 \langle \omega(y), \frac{\partial S_i}{\partial x} \rangle \circ \phi_i^t(Y, X) \mathrm{d}t + \varepsilon \bar{P}_i \circ \Phi_i(y, x) \\ &+\varepsilon \tilde{P}_i \circ \Phi_i(Y, X) = N_0(Y) - \varepsilon \int_0^1 (\langle \omega_*, \frac{\partial S_i}{\partial x} \rangle - \tilde{P}_i) \circ \phi_i^t(Y, X) \mathrm{d}t + \varepsilon P_{i+1}(Y, X) \\ &= N_0(Y) + \varepsilon P_{i+1}(Y, X). \end{aligned}$$

Let

$$\mathcal{P}_{i+1}(Y, X) = P_{i+1}(Y, X) - \bar{P}_i(Y, X). \tag{2.12}$$

Then it is clear that

$$\tilde{\mathcal{P}}_{i+1} = \tilde{P}_{i+1}. \tag{2.13}$$

We now show by induction that

$$\Phi_i : D_i' \to D_i, \tag{2.14}$$

$$|\varepsilon \tilde{P}_{i+1}|_{D_i'} \le 2 |\varepsilon \mathcal{P}_{i+1}|_{D_i'} < \zeta_i, \tag{2.15}$$

$$|\varepsilon \frac{\partial P_{i+1}}{\partial Y}|_{D_{i+1}} < 2MR, \ |\varepsilon \frac{\partial P_{i+1}}{\partial X}|_{D_{i+1}} < 2MR^2, \tag{2.16}$$

for all $i = 0, 1, \cdots, K - 1$.

Starting from $i = 0$, we have by (2.8) that

$$|\varepsilon \partial^\jmath S_0|_{G \times T^n + \delta} \le T |\varepsilon \partial^\jmath P_0|_{G \times T^n + \delta} \le \frac{\varepsilon M T}{2}, \tag{2.17}$$

for all $\jmath = (\jmath_1, \jmath_2) \in Z_+^l \times Z_+^n, |\jmath| = |\jmath_1| + |\jmath_2| \le 1$.

Let $(Y,X)\in D_0'$, $t^*=\sup\{t\in[0,1]:\Phi_0(Y,X)\in D_0\}$. Then it follows from (2.6), (2.10), (2.17) that, as $0\le t\le t^*$, $(y,x)=\Phi_0(Y,X)$ satisfies

$$\begin{aligned}\big||y-y_*|-|Y-y_*|\big| &\le |y-Y|\le |B^\top|_{D_0}|\varepsilon\frac{\partial S_0}{\partial x}|_{D_0}\le\frac{\varepsilon M^2T}{2}\\ &\le (10M^2TRK)\frac{\varepsilon}{R^2}\frac{\gamma}{2}<(M^*TRK)\frac{\gamma}{2}\le\frac{\gamma}{2},\end{aligned}\tag{2.18}$$

$$\begin{aligned}\big||\mathrm{Im}x|-|\mathrm{Im}X|\big| &\le |x-X|\le |B|_{D_0}|\varepsilon\frac{\partial S_0}{\partial x}|_{D_0}+|C|_{D_0}|\varepsilon\frac{\partial S_0}{\partial y}|_{D_0}\le\varepsilon M^2T\\ &\le \frac{20M^2TRK}{\delta}\frac{\varepsilon}{R}\frac{\sigma}{2}<(M^*TRK)\frac{\sigma}{2}\le\frac{\sigma}{2}.\end{aligned}\tag{2.19}$$

It follows that $t^*=1$ and

$$\phi_0^t: D_0'\to D_0,\ 0\le t\le 1.\tag{2.20}$$

In particular, (2.14) holds for $i=0$.

Denote ϕ_{i1}^t,ϕ_{i2}^t as the components of ϕ_i^t with respect to $G+\delta, T^n+\delta$ respectively, for all $i=0,1,\cdots,K$. By (2.11)-(2.13), (2.17), (2.20), we have that

$$\begin{aligned}&|\varepsilon\tilde P_1|_{D_0'}=|\varepsilon\tilde{\mathcal P}_1|_{D_0'}\le 2|\varepsilon\mathcal P_1|_{D_0'}\le 2\varepsilon|\frac{\partial\bar P_0}{\partial y}|_{D_0}|y-Y|_{D_0'}+2\varepsilon|\frac{\partial\bar P_0}{\partial x}|_{D_0}|x-X|_{D_0'}\\ +&\ 2\varepsilon|\frac{\partial\tilde P_0}{\partial y}|_{D_0}\sup_{0\le t\le1}|\phi_{01}^1(Y,X)-\phi_{01}^t(Y,X)|_{D_0'}\\ +&\ 2\varepsilon|\frac{\partial\tilde P_0}{\partial x}|_{D_0}\sup_{0\le t\le1}|\phi_{02}^1(Y,X)-\phi_{02}^t(Y,X)|_{D_0'}\\ +&\ 2|\frac{\partial\omega}{\partial y}|_{D_0'}|y-y_*|_{D_0}|\varepsilon\frac{\partial S_0}{\partial x}|_{D_0'}\\ \le&\ 6\varepsilon|\frac{\partial P_0}{\partial y}|_{D_0}\sup_{0\le t\le1}|Y-\phi_{01}^t(Y,X)|_{D_0'}+6\varepsilon|\frac{\partial P_0}{\partial x}|_{D_0}\sup_{0\le t\le1}|X-\phi_{02}^t(Y,X)|_{D_0'}\\ +&\ 2(|\frac{\partial\omega}{\partial y}|_{D_0'}|y-y_*|_{D_0})\le(6\varepsilon M)(\frac{\varepsilon M^2T}{2})\\ +&\ (6\varepsilon M)(\varepsilon M^2T)+(4\varepsilon MR)(\frac{MT}{2})=(\frac{9\varepsilon}{R}+4)\varepsilon M^3RT<5\varepsilon M^3RT=\zeta_0.\end{aligned}$$

Hence (2.15) holds for $i=0$. Moreover, we also have

$$|\varepsilon P_1|_{D_0'}\le|\varepsilon\mathcal P_1|_{D_0'}+|\varepsilon P_0|_{D_0'}\le\zeta_0+\varepsilon M\le 6M^3\varepsilon T.\tag{2.21}$$

Since by Cauchy estimate,

$$\begin{aligned}\varepsilon|\frac{\partial\mathcal P_1}{\partial Y}|_{D_1}&\le\frac{\zeta_0}{\gamma}=100M^3RTK\frac{\varepsilon}{R^2}\frac{R}{2}\le(M^*RTK)\frac{R}{2}\le\frac{R}{2},\\ \varepsilon|\frac{\partial\mathcal P_1}{\partial X}|_{D_1}&\le\frac{\zeta_0}{\sigma}=\frac{100M^3RTK}{\delta}\frac{\varepsilon}{R^2}\frac{R^2}{2}\le(M^*RTK)\frac{R^2}{2}\le\frac{R^2}{2},\end{aligned}$$

we have that

$$\varepsilon|\frac{\partial P_1}{\partial Y}|_{D_1}\le\varepsilon(|\frac{\partial\mathcal P_1}{\partial Y}|_{D_1}+|\frac{\partial P_0}{\partial Y}|_{D_1})\le\frac{R}{2}+\varepsilon M<MR,\tag{2.22}$$

$$\varepsilon|\frac{\partial P_1}{\partial X}|_{D_1}\le\varepsilon(|\frac{\partial\mathcal P_1}{\partial X}|_{D_1}+|\frac{\partial P_0}{\partial X}|_{D_1})\le\frac{R^2}{2}+\varepsilon M<\frac{3MR^2}{2}.\tag{2.23}$$

Thus, (2.16) also holds for $i=0$.

Now suppose that for some $j = 1, 2, \cdots, K-1$, (2.14)-(2.16) hold for all $i = 0, 1, \cdots, j-1$. Then by (2.8), (2.15),

$$|\varepsilon S_j|_{D_j} \le \frac{T}{2}|\varepsilon \tilde{P}_j|_{D'_{j-1}} \le \frac{T}{2}\zeta_{j-1}.$$

It follows from Cauchy estimate that

$$\left|\varepsilon \frac{\partial S_j}{\partial y}\right|_{D'_j} \le \frac{T}{\gamma}\zeta_{j-1}, \tag{2.24}$$

$$\left|\varepsilon \frac{\partial S_j}{\partial x}\right|_{D'_j} \le \frac{T}{\sigma}\zeta_{j-1}. \tag{2.25}$$

Using (2.24), (2.25) and arguments similar to that for (2.18), (2.19), we have that

$$\begin{aligned}
&\big||y-y_*| - |Y-y_*|\big| \le |y-Y| \le |B^\top|_{D_j}|\varepsilon \frac{\partial S_j}{\partial x}|_{D_j} \le \frac{MT\zeta_{j-1}}{\sigma} \\
&\le \frac{200MTK^2}{\delta R}\zeta_0\frac{\gamma}{2} = \frac{1000M^4T^2R^2K^2}{\delta}\frac{\varepsilon}{R^2}\frac{\gamma}{2} < (M^*TRK)^2\frac{\gamma}{2} \le \frac{\gamma}{2}, \qquad (2.26)\\
&\big||\mathrm{Im}x| - |\mathrm{Im}X|\big| \le |x-X| \le |B|_{D_j}|\varepsilon\frac{\partial S_j}{\partial x}|_{D_j} + |C|_{D_j}|\varepsilon\frac{\partial S_j}{\partial y}|_{D_j} \\
&\le MT\zeta_{j-1}(\frac{1}{\gamma}+\frac{1}{\sigma}) \le \frac{200MTK^2}{\delta}(\frac{1}{R}+\frac{1}{\delta})\zeta_0\frac{\sigma}{2} \\
&= (\frac{1000M^4T^2R^2K^2}{\delta}\frac{\varepsilon}{R^2} + \frac{1000M^4T^2R^2K^2}{\delta^2}\frac{\varepsilon}{R})\frac{\sigma}{2} \\
&\le \frac{2000M^4T^2R^2K^2}{\delta^2} < (M^*TRK)^2\frac{\sigma}{2} \le \frac{\sigma}{2}, \qquad (2.27)
\end{aligned}$$

for all $(Y, X) \in D'_j$, $0 \le t \le 1$, and $(y, x) = \phi_j^t(Y, X)$. In particular, (2.14) holds for $i = j$.

By (2.11)-(2.13), (2.24)-(2.27) and the induction hypothesis, we have that

$$\begin{aligned}
&|\varepsilon\tilde{P}_{j+1}|_{D'_j} = |\varepsilon\tilde{\mathcal{P}}_{j+1}|_{D'_j} \le 2|\varepsilon\mathcal{P}_{j+1}|_{D'_j} \le 2\varepsilon|\frac{\partial\bar{P}_j}{\partial y}|_{D_j}|y-Y|_{D'_j} \\
&+2\varepsilon|\frac{\partial\bar{P}_j}{\partial x}|_{D_j}|x-X|_{D'_j} + 2\varepsilon|\frac{\partial\tilde{P}_j}{\partial y}|_{D_j}\sup_{0\le t\le 1}|\phi^1_{j1}(Y,X) - \phi^t_{j1}(Y,X)|_{D'_j} \\
&+2\varepsilon|\frac{\partial\tilde{P}_j}{\partial x}|_{D_j}\sup_{0\le t\le 1}|\phi^1_{j2}(Y,X) - \phi^t_{j2}(Y,X)|_{D'_j} + 2|\frac{\partial\omega}{\partial y}|_{D'_j}|y-y_*|_{D_j}|\varepsilon\frac{\partial S_j}{\partial x}|_{D'_j} \\
&\le 6\varepsilon|\frac{\partial P_j}{\partial y}|_{D_j}\sup_{0\le t\le 1}|Y - \phi^t_{j1}(Y,X)|_{D'_j} \\
&+6\varepsilon|\frac{\partial P_j}{\partial x}|_{D_j}\sup_{0\le t\le 1}|X - \phi^t_{j2}(Y,X)|_{D'_j} + 2(|\frac{\partial\omega}{\partial y}|_{D'_j}|y-y_*|_{D_j} \\
&\le \frac{12RMT\zeta_{j-1}}{\sigma} + \frac{12R^2MT\zeta_{j-1}}{\gamma} + \frac{4MRT\zeta_{j-1}}{\sigma} \\
&\le \frac{16RMT\zeta_{j-1}}{\sigma} + \frac{12R^2MT\zeta_{j-1}}{\gamma} = (\frac{160RMTK}{\delta} + 120RM^2TK)\zeta_{j-1} \\
&\le \frac{560RM^2TK}{\delta}\zeta_j < (M^*RTK)\zeta_j \le \zeta_j, \qquad (2.28)
\end{aligned}$$

that is, (2.15) holds for $i=j$.

Since by the induction hypothesis and (2.28)

$$|\varepsilon P_{i+1}-\varepsilon P_i|_{D'_i}=|\varepsilon\mathcal{P}_{i+1}-\varepsilon\tilde{P}_i|_{D'_i}\le|\varepsilon\mathcal{P}_{i+1}|_{D'_i}+\varepsilon|\tilde{P}_i|_{D'_i}\le 2\zeta_{i-1}=4\zeta_i,$$

for all $i=0,1,\cdots,j$, we have that

$$|\varepsilon P_{j+1}-\varepsilon P_1|_{D'_j}\le\sum_{i=1}^{j}|\varepsilon P_{i+1}-\varepsilon P_i|_{D'_i}\le 4\sum_{i=1}^{j}\zeta_i\le 4\zeta_0. \tag{2.29}$$

Hence by (2.22), (2.23) and Cauchy estimate we have that

$$\begin{aligned}
&|\varepsilon\frac{\partial P_{j+1}}{\partial Y}|_{D_{j+1}}\le|\frac{\partial}{\partial Y}(\varepsilon P_{j+1}(Y,X)-\varepsilon P_1(Y,X))|_{D_{j+1}}+|\varepsilon\frac{\partial P_1(Y,X)}{\partial Y}|_{D_1}\\
&\le\frac{8\zeta_0}{\gamma}+MR=(400M^3RTK)\frac{\varepsilon}{R^2}R+MR\le(M^*TKR)MR+MR\le 2MR,\\
&|\varepsilon\frac{\partial P_{j+1}}{\partial X}|_{D_{j+1}}\le|\frac{\partial}{\partial X}(\varepsilon P_{j+1}(Y,X)-\varepsilon P_1(Y,X))|_{D_{j+1}}+|\varepsilon\frac{\partial P_1(Y,X)}{\partial X}|_{D_{j+1}}\\
&\le\frac{8\zeta_0}{\sigma}+\frac{3MR^2}{2}=\frac{400M^3RTK}{\delta}\frac{\varepsilon}{\delta R^2}MR^2+\frac{3MR^2}{2}\\
&\le(M^*TKR)\frac{MR^2}{2}+\frac{3MR^2}{2}\le 2MR^2,
\end{aligned}$$

that is, (2.16) also holds for $i=j$. This proves (2.14)-(2.16) for all $i=0,1,\cdots,K-1$.

Now let $\Phi=\Phi_0\circ\Phi_1\circ\cdots\circ\Phi_{K-1}$, $H_*=H_K$, $P_*=P_K$. By (2.16), (2.21) and (2.29), we first have that

$$\begin{aligned}
&|\varepsilon P_*|_{D_K}\le|P_K-P_1|_{D_K}+|P_1|_{D_1}\le 20M^3\varepsilon TR+6M^3\varepsilon T\le 26M^3\varepsilon T,\\
&|\varepsilon\frac{\partial P_*}{\partial Y}|_{D_K}<2MR,\\
&|\varepsilon\frac{\partial P_*}{\partial X}|_{D_K}<2MR^2.
\end{aligned}$$

Since

$$K\ge\frac{1}{M_*c_2}Q^{\frac{1}{n_*-1}}=\frac{1}{M_*c_2}\varepsilon^{-\frac{1}{2n_*}},$$

we have by (2.15) that there is a constant $c>0$ such that

$$\begin{aligned}
|\varepsilon\tilde{P}_*|_{D'_{K-1}}&\le\zeta_{K-1}=10M^3\varepsilon TR\left(\frac{1}{2}\right)^K\\
&\le(M^*TRK)\varepsilon\left(\frac{1}{2}\right)^K\le\varepsilon\left(\frac{1}{2}\right)^K<\exp(-c^{-1}\varepsilon^{-\frac{1}{2n_*}}).
\end{aligned}$$

Performing Cauchy estimate for $\varepsilon\tilde{P}_*$ on D_K, we also have

$$\begin{aligned}
|\varepsilon\frac{\partial\tilde{P}_*}{\partial X}|_{D_K}&\le\frac{200M^3\varepsilon TR}{\delta}\left(\frac{1}{2}\right)^K\le(M^*TRK)\varepsilon\left(\frac{1}{2}\right)^K\\
&\le\varepsilon\left(\frac{1}{2}\right)^K<\exp(-c\varepsilon^{-\frac{1}{2n_*}}),
\end{aligned}$$

for some constant $c>0$.

Let

$$D_* = \{(Y, X) \in G \times T^n + \delta : |Y - y_*| < \frac{R}{2},\ |\mathrm{Im}X| < \frac{\delta}{2}\}.$$

Then $D_* \subset D'_K \subset D_K$. Hence, all above estimates hold on D_*, and, $\Psi =: \Phi^{-1} : \Phi(D_*) \to D_*$ is well defined and transforms H to H_*. Now, for any $(Y, X) \in D_*$, $(y, x) = \Psi^{-1}(Y, X)$, we have by a successive application of (2.18), (2.19), (2.26), (2.27) that

$$\begin{aligned}
|y - Y| &\le \varepsilon M^2 T + \frac{MT}{\sigma}\sum_{i=1}^{K-1}\zeta_i \le \varepsilon M^2 T + \frac{MT}{\sigma}5M^3\varepsilon TR \\
&= M^2\varepsilon T + \frac{50M^4TKR}{\delta}\varepsilon T \le M^2\varepsilon T + M(M^*TKR)\varepsilon T \le 2M^2\varepsilon T, \\
|x - X| &\le 2\varepsilon M^2 T + MT(\frac{1}{\gamma} + \frac{1}{\sigma})\sigma\sum_{i=1}^{K-1}\zeta_i \\
&\le 2\varepsilon M^2 T + MT(\frac{1}{\gamma} + \frac{1}{\sigma})5M^3\varepsilon TR \\
&\le 3\varepsilon M^2 T + \frac{MT}{\delta}5M^3\varepsilon TR \\
&= 3M^2\varepsilon T + \frac{50M^4TKR}{R}\varepsilon T \le 3M^2\varepsilon T + \frac{M}{R}(M^*TKR)\varepsilon T \le 4M^2\frac{\varepsilon T}{R}.
\end{aligned}$$

The fact $D \subset \Psi^{-1}(D_*)$ follows from the above and the fact that $\varepsilon T = o(R)$. This completes the proof. □

Proof of Theorem 1. We first consider the autonomous Hamiltonian (1.6). Fix $(y(0), x(0)) \in G \times T^n$. Then there is a $y_* \in G + \delta$ satisfying the properties described in Lemma 2.2. In particular,

$$|y(0) - y_*| < \frac{R}{3c_3}. \tag{2.30}$$

For this y_*, it follows from Lemma 2.4 that there is a canonical, real analytic defeomorphism $\Psi : D \to D_* : (y, x) \mapsto (Y, X)$, where D, D_* are as in Lemma 2.4, which transforms (2.5) to

$$H_* = \langle \frac{\partial N}{\partial y}(y_*), Y - y_* \rangle + h(Y - y_*) + \varepsilon \bar{P}_*(Y, X) + \varepsilon \tilde{P}_*(Y, X) \tag{2.31}$$

satisfying all properties described in Lemma 2.4. Let $(Y(t), X(t))$ be the orbit of motion associated to (2.31) with the initial value $(Y(0), X(0)) = \Psi(y(0), x(0))$. Then by Lemma 2.4 and the fact that $\varepsilon T = o(R)$, we have

$$|Y(0) - y_*| \le |y(0) - y_*| + |Y(0) - y(0)| < \frac{R}{3c_3} + o(R) < \frac{R}{2c_3}.$$

Note that

$$\langle \omega_*, \frac{\partial \bar{P}_*}{\partial X}(Y, X) \rangle = \frac{1}{T}\int_0^T \frac{dP_*}{dt}(Y, X + \omega_* t)dt = 0.$$

We have

$$\begin{aligned}\langle\frac{\partial N}{\partial y}(y_*),\dot{Y}\rangle &= \langle B^\top\frac{\partial N}{\partial y}(y_*),\varepsilon\frac{\partial P_*}{\partial X}\rangle = -\langle\omega_*,\varepsilon\frac{\partial\tilde{P}_*}{\partial X}\rangle\\ &= O(\exp(c_4\varepsilon^{-\frac{1}{2n_*}})).\end{aligned}\tag{2.32}$$

Define

$$T_\varepsilon = \exp(\frac{c_4}{2}\varepsilon^{-\frac{1}{2n_*}}),$$

$$T_\varepsilon^0 = \sup\{0<|t|<T_\varepsilon : |Y(t)-y_*|<\frac{R}{2}\}.$$

It follows from (2.32) that

$$\langle\frac{\partial N}{\partial y}(y_*),Y(t)-Y(0)\rangle = o(R),\tag{2.33}$$

as $|t|\le T_\varepsilon^0$. By Lemma 2.4 and virtue of conservation of energy, we also have

$$\begin{aligned}N(Y(t)) &= N(Y(0))+\varepsilon P_*(Y(0),X(0))-\varepsilon P_*(Y(t),X(t))\\ &= N(Y(0))+O(\varepsilon T) = N(Y(0))+o(R),\end{aligned}\tag{2.34}$$

as $|t|\le T_\varepsilon^0$. Now, as $|t|\le T_\varepsilon^0$, (2.33), (2.34) imply that $Y(t)$ lies in an $o(R)$ neighborhood of the intersection of the plane $\langle\frac{\partial N}{\partial y}(y_*),Y-Y(0)\rangle=0$ with the surface $N(Y)=N(Y(0))$, which, by Lemma 2.3, is contained in an open ball of radius $c_3|Y(0)-y_*|<R/2$ centered at y_*. Hence

$$|Y(t)-y_*|<\frac{R}{2},\tag{2.35}$$

as $|t|\le T_\varepsilon^0$. This shows that $T_\varepsilon^0=T_\varepsilon$ and (2.35) holds for all $|t|\le T_\varepsilon$.

Let $(y(t),x(t))$ be the orbit of motion associated to (2.5) with the initial value $(y(0),x(0))$. We have by Lemma 2.4, (2.30) and (2.35) that

$$\begin{aligned}|y(t)-y(0)| &\le |y(t)-Y(t)|+|Y(t)-y_*|+|y_*-y(0)| = O(\varepsilon T)+O(R)\\ &= O(R)=O(\varepsilon^{\frac{1}{2n_*}}),\end{aligned}$$

as $|t|\le T_\varepsilon$. This completes the proof of Theorem 1.

3. KAM stability

KAM stability, particularly implying metric stability, for nearly integrable generalized Hamiltonian systems was shown in [40] under a nondegenerate condition of Rüssmann type - known as the weakest nondegenerate condition for the existence of KAM tori in general. More precisely, the following was proved in [40].

Theorem 2. *Consider* (1.6). *The following holds as* ε *sufficiently small.*

1) *If there is a constant* $C>0$ *such that*

R) $\operatorname{rank}\left\{\frac{\partial^i\omega}{\partial y^i} : i\in Z_+^n,\ |i|\le C\right\}\equiv n$ *over* G,

then the majority of unperturbed n*-tori on* G *will persist in the sense that there is a smooth family of analytic, canonical transformations* $\Phi_\varepsilon : G\times T^n\to G\times T^n$ *uniformly close to the identity, and a family of Cantor-like sets* $G_\varepsilon\subset G$ *with* $|G\setminus G_\varepsilon|\to 0$ *as* $\varepsilon\to 0$ *such that* $\Phi_\varepsilon(G_\varepsilon\times T^n)$ *is a Whitney smooth family of analytic, Diophantine, invariant* n*-tori of the perturbed system.*

2) *If I is a constant matrix and if the Hessian matrix $\partial^2 N/\partial y^2$ is non-singular on G, then for any given (γ,τ), with $\gamma > \varepsilon^{\frac{1}{8n+12}}$, $\tau > n-1$, all unperturbed Diophantine tori whose frequencies $\omega(y)$ are of the Diophantine types (γ,τ) will persist and give rise to analytic, Diophantine, invariant n-tori of the perturbed system with the same frequencies.*

Combining proofs of [19, 40], the following iso-energetic KAM theorem can be obtained for generalized Hamiltonian systems.

Theorem 3. *Consider* (1.6). *The following holds as ε sufficiently small.*

1) *Let $M = \{y \in G : N(y) = E\}$ be a fixed energy surface of the unperturbed Hamiltonian and λ be a local coordinate on M. If there is a constant $C > 0$ such that*

R1) $\operatorname{rank}\left\{\frac{\partial^\alpha \omega}{\partial \lambda^\alpha} : |\alpha| \le C\right\} \equiv n$ *over* M,

then the majority of unperturbed n-tori on M will persist in the sense that there is a smooth family of analytic, canonical transformations $\Phi_{\varepsilon,E} : M \times T^n \to M \times T^n$ uniformly close to the identity, and a family of Cantor-like sets $M_\varepsilon \subset M$ with $|M \setminus M_\varepsilon| \to 0$ as $\varepsilon \to 0$ such that $\Phi_{\varepsilon,E}(M_\varepsilon \times T^n)$ is a Whitney smooth family of analytic, Diophantine, invariant n-tori of on the perturbed energy surface $\{H(y,x) = E\}$.

2) *If I is a constant matrix,* R) *holds, and if*

$$\det \begin{pmatrix} \frac{\partial^2 N}{\partial y^2}(y) & \frac{\partial N}{\partial y}(y) \\ \frac{\partial N}{\partial y}(y)^\top & 0 \end{pmatrix} \neq 0 \quad \text{on } G,$$

then the conclusion of 1) *holds on any energy surface and each perturbed n-torus preserves the ratio of the components of the corresponding unperturbed frequency.*

Theorems 2, 3 can be easily extended to the non-autonomous system (2.2) as follows.

Theorem 4. *Assume $\hat\omega$ satisfies the Diophantine condition*

$$|\langle k, \hat\omega\rangle| > \frac{\gamma'}{|k|^{\tau'}} \quad \forall k \in Z^{n'} \setminus \{0\},$$

for fixed $\gamma' > 0, \tau' > n' - 1$. The following holds as ε sufficiently small.

1) *If the condition of* Theorem 2 1) *holds, then there is a smooth family of analytic, canonical transformations $\Phi_\varepsilon : G \times T^n \to G \times T^n$ uniformly close to the identity, and a family of Cantor-like sets $G_\varepsilon \subset G$ with $|G \setminus G_\varepsilon| \to 0$ as $\varepsilon \to 0$ such that $\Phi_\varepsilon(G_\varepsilon \times T^n) \times T^{n'}$ is a Whitney smooth family of analytic, Diophantine, invariant $n + n'$-tori of the perturbed system whose last n' toral frequency equals $\hat\omega$.*

2) *If the conditions of* Theorem 2 2) *holds, then for any $y \in G$ such that $(\omega(y), \hat\omega)$ is of the Diophantine types (γ,τ) for fixed $\gamma > \varepsilon^{\frac{1}{8(n+n')+12}}$, $\tau > \max\{\tau', n + n' - 1\}$, $\Phi_\varepsilon(\{y\} \times T^n) \times T^{n'}$ is an analytic, Diophantine, invariant $n+n'$-tori of the perturbed system with the same frequencies $(\omega(y), \hat\omega)$.*

3) *If the condition of* Theorem 3 1) *holds on the unperturbed energy surface M with energy E, then there is a smooth family of analytic, canonical transformations $\Phi_{\varepsilon,E} : M \times T^n \to M \times T^n$ uniformly close to the identity, and a family of Cantor-like sets $M_\varepsilon \subset M$ with $|M \setminus M_\varepsilon| \to 0$ as $\varepsilon \to 0$ such that $\Phi_{\varepsilon,E}(M_\varepsilon \times T^n) \times T^{n'}$ is a Whitney smooth family of analytic, Diophantine, invariant $n + n'$-tori of on the perturbed energy surface $\{H(y, x, \hat x) = E\}$, whose last n' toral frequency equals $\hat\omega$.*

4) *If the conditions of* Theorem 3 2) *hold, then the conclusion in* 3) *holds on any energy surface and each perturbed* $n+n'$*-torus preserves the ratio of the first* n *components of the corresponding unperturbed frequency.*

Remark. *The Rüssmann non-degenerate condition was first given by Rüssmann* ([65])*. Let* $\omega : G \subset R^n \to R^n$ *be a real analytic frequency map. Geometrically, the Rüssmann non-degenerate condition says that* $\omega(G)$ *should not lie in any* $d-1$ *dimensional subspace. KAM theorems under the Rüssmann non-degenerate condition were shown in* [67, 70] *under analytic conditions similar to* R) *by taking* $C = n-1$*. Extensions to the case of sub-manifolds and also to the case of generalized Hamiltonian systems were made by the authors in* [19] *and* [40] *respectively.*

It is recently remarked by Rüssmann ([66]) *that only the condition R) with a general* C *is equivalent to the Rüssmann non-degenerate condition and the one with* $C = n-1$ *is much stronger. However, we remark that results in* [19, 40] (*also in* [38, 39, 67, 70]) *do hold with a general* C *in place of* $n-1$ *in the respective conditions* R) *or* R1) (*the proofs go through except that the measure estimates depend on* C).

We now consider some particular cases of these theorems in which the Nekhoroshev stability results stated in the previous section are of minor interest, as the corresponding generalized Hamiltonian systems will be perpetually stable.

3.1. $l = 1$ **or** $n = 1$.

Proposition 3.1. *Let* $l = 1$ *and* $N(y) \not\equiv 0$ *on* G.

1) (1.6) *is perpetually stable, hence metrically and Nekhoroshev stable.*

2) (1.6) *is KAM stable if both* N' *and the Wronskian of* B *are non-zero on* G*. Moreover, if* I *is a constant matrix,* B *is a Diophantine vector, and* N'' *is non-zero on* G*, then all unperturbed* n*-tori persist.*

Proof. The proof of 1) is straightforward since each energy surface consists of finitely many invariant n-tori.

To prove 2), we note that $W(\omega) = (-1)^n W(B)(N')^2$, where W denotes the Wronskian of a vector valued function. Hence the non-degenerate condition R) holds and the KAM stability follows from part 1) of Theorem 2. In the case that I is a constant matrix, B is a Diophantine vector, and N'' is non-zero on G, all unperturbed n-tori are Diophantine, hence part 2) of Theorem 2 is applicable. □

The invariant n-tori on an energy surface are not necessary quasi-periodic, because if B is a resonant constant vector then $\omega = -B^\top N'$ becomes everywhere resonant, hence no quasi-periodic n-tori need to existence under generic perturbations. In the case that conditions in part 2) of the proposition holds, the resulting KAM stability is significant with fewer parameters (in fact only one) in the frequency map. We note from the conditions of part 2) of the proposition that the non-constancy of the structure matrix is generally needed to ensure such KAM stability.

Proposition 3.2. *Let* $n = 1$ *and* $\omega(y) \neq 0$ *on* G*. Then all unperturbed* 1*-tori persist, that is,* (1.6) *is both KAM and perpetually stable, hence also metrically and Nekhoroshev stable.*

Proof. We note that $\text{rank}\{\omega(y)\} \equiv 1$ on G. It follows from part 1) of Theorem 2 that (1.6) is KAM stable. In fact, since the one dimensional frequencies $\omega(y)$ are

Diophantine, no measure estimate is needed in the proof of KAM stability in this case (see the proof of [40]). It follows that all unperturbed 1-tori persist - leading to the desired perpetual stability. □

3.2. $l = 2$. In a standard nearly integrable Hamiltonian system of two degrees of freedom, it is well known that if the unperturbed Hamiltonian is iso-energetic non-degenerate (in particular quasi-convex), then the system is perpetually stable. The same holds in the generalized setting if $l = 2$.

Proposition 3.3. *Consider* (1.6) *with* $l = 2$*. If the iso-energetic non-degenerate condition* R1) *holds, then* (1.6) *is both KAM and perpetually stable, hence also metrically and Nekhoroshev stable.*

Proof. In this case, (1.6) is iso-energetically KAM stable in the sense of Theorem 3. Hence each energy surface is almost foliated into invariant, co-dimension 1 tori. Therefore, any orbit on the energy surface is forever trapped between these tori - leading to the desired perpetual stability. □

We note that if $l = n = 2$ then QR) together with rank$\tilde{B} = 2$ implies the condition R), hence R1) (for example, let $N(y)$ be quasi-convex in the sense of QC), B be any invertible 2×2 constant matrix, and C be any 2×2 skew symmetric matrix which can be y dependent). In this case, although it follows from Theorem 1 2) that the respective perturbed generalized Hamiltonian system is Nekhoroshev stable, it is actually KAM and perpetual stable by Proposition 3.3.

Therefore, non-perpetual, Nekhoroshev stability phenomena of autonomous generalized Hamiltonian systems become significant only when $l > 3, n \geq 2$ or $l = 2, n > 2$.

3.3. $l+n = 3$. Due to their natural presence in many physical systems, generalized Hamiltonian systems in 3D, corresponding to the simplest non-trivial cases where Poisson structures do not imply the symplectic ones, have received a considerable amount of attention in recent years and a rich class of Poisson structures in 3D has been characterized (see [33, 34] and references therein).

As an application of our results, we now revisit the perturbation problem of the 3D steady Euler (inviscid and incompressible) fluid particle path flows considered in [40, 52]. As shown in [52], the 3D steady Euler fluid particle path flow admits a one-parameter, spatial, volume preserving symmetry group ([52, 58]), and hence under suitable coordinates it can be described by a three dimensional volume preserving flow of the following form:

$$\begin{cases} \dot{z}_1 = \dfrac{\partial H(z_1, z_2)}{\partial z_2} \\ \dot{z}_2 = -\dfrac{\partial H(z_1, z_2)}{\partial z_1} \\ \dot{z}_3 = h(z_1, z_2), \end{cases} \tag{3.1}$$

where the right hand side describes the velocity field v_0 of the steady Euler fluid flow (under the present coordinate) with H being a first integral. We assume that *the steady Euler flow is real analytic and admits a family of elliptic vortex lines* in a bounded closed domain of the (z_1, z_2)-plane. Then by introducing the standard action-angle coordinates $(\mathcal{I}, \theta)$ in the (z_1, z_2) domain, (3.1) can be transformed into the form

$$\begin{cases} \dot{\mathcal{I}} = 0 \\ \dot{\theta} = \omega_1(\mathcal{I}) \\ \dot{z}_3 = h(\mathcal{I}, \theta). \end{cases} \tag{3.2}$$

By [52], if $\omega_1(\mathcal{I}) \neq 0$ on $\mathcal{D}$, then the volume preserving transformation $(\mathcal{I}, \theta, z_3) \to (\mathcal{I}, \theta, \phi)$:

$$\phi = z_3 + \frac{\theta}{2\pi} \int_0^{2\pi} \frac{h(\mathcal{I}, \theta)}{\omega_1(\mathcal{I})} \mathrm{d}\theta - \int \frac{h(\mathcal{I}, \theta)}{\omega_1(\mathcal{I})} \mathrm{d}\theta$$

will transform (3.2) to the system

$$\begin{cases} \dot{\mathcal{I}} = 0 \\ \dot{\theta} = \omega_1(\mathcal{I}) \\ \dot{\phi} = \omega_2(\mathcal{I}), \end{cases} \tag{3.3}$$

where

$$\omega_2(\mathcal{I}) = \frac{1}{2\pi} \int_0^{2\pi} h(\mathcal{I}, \theta) \mathrm{d}\theta,$$

and, $\phi \in S^1$ or R^1, depending on whether the symmetry group acting on the flow is S^1 or R^1. Thus, under the above conditions, the particle phase space R^3 of the steady Euler flow is foliated into either invariant cylinders if $\omega_2(\mathcal{I}) \equiv 0$ on $\mathcal{D}$, or invariant 2-tori if $\omega_2(\mathcal{I}) \not\equiv 0$ on $\mathcal{D}$, carrying either *action-action-angle* or *action-angle-angle variables*, respectively.

In the *action-action-angle* case, we let $\omega_1(\mathcal{I}) \neq 0$, $\omega_2(\mathcal{I}) \equiv 0$ on $\mathcal{D}$ and define $y_1 = \mathcal{I}$, $y_2 = \phi$, $y = (y_1, y_2)^\top$, $x = \theta \in T^1$, $\omega(y) = \omega_1(y_1)$. Let $\alpha(y), \beta(y), N(y)$ be such that $-(\alpha(y), \beta(y))\nabla_y N(y) \equiv \omega(y)$. Then the right hand side of (3.3) can be written as

$$I\nabla N(y) = \begin{pmatrix} 0 & 0 & \alpha(y) \\ 0 & 0 & \beta(y) \\ -\alpha(y) & -\beta(y) & 0 \end{pmatrix} \nabla N(y),$$

where the matrix I is easily seen to satisfy the Jacobi identity hence defines a structure matrix. With the structure preserving, time-periodic perturbation $\varepsilon P(y, x, t)$, the perturbed system of (3.3) becomes

$$\begin{pmatrix} \dot{y} \\ \dot{x} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \alpha b(y) \\ 0 & 0 & \beta b(y) \\ -\alpha b(y) & -\beta b(y) & 0 \end{pmatrix} \nabla(N(y) + \varepsilon P(y, x, t)). \tag{3.4}$$

In the *action-angle-angle* case, we let $y = \mathcal{I} \in \mathcal{D}$, $x = (x_1, x_2)^\top = (\theta, \phi)^\top \in T^2$, $\omega(y) = (\omega_1(y), \omega_2(y))^\top$, α, β, γ be constants such that $|\alpha| + |\beta| \neq 0$, $b(y), N(y) \neq 0$ be real analytic functions. Then it is easy to see that

$$I(y) = \begin{pmatrix} 0 & \alpha b(y) & \beta b(y) \\ -\alpha b(y) & 0 & -\gamma \\ -\beta b(y) & \gamma & 0 \end{pmatrix}$$

satisfies the Jacobi identity hence determines a structure matrix. Adding a structure preserving, time-periodic perturbation $\varepsilon P(y, x, t)$ to the unperturbed Hamiltonian N, the perturbed system of (3.3) becomes

$$\begin{pmatrix} \dot{y} \\ \dot{x} \end{pmatrix} = \begin{pmatrix} 0 & \alpha b(y) & \beta b(y) \\ -\alpha b(y) & 0 & -\gamma \\ -\beta b(y) & \gamma & 0 \end{pmatrix} \nabla(N(y) + \varepsilon P(y, x, t)). \tag{3.5}$$

We note that the perturbations in (3.4), (3.5) are allowed to break the volume preserving symmetry (corresponding to compressible fluid perturbations). Indeed, it is easy to see that the perturbations $I(y)\nabla P(y,x,t)$ in (3.4), (3.5), respectively, are divergence free if and only if

$$(\alpha_{y_1}+\beta_{y_2})P_x \equiv 0,$$

and,

$$b'(y)(\alpha P_{x_1}+\beta P_{x_2}) \equiv 0,$$

respectively.

In the cases that the perturbations in (3.4), (3.5) are autonomous, it follows from Propositions 3.2 that (3.4) is both perpetually and KAM stable, and, from Propositions 3.1 that (3.5) is perpetually stable. Moreover, by part 2) of Propositions 3.2, (3.5) becomes KAM stable in a domain of interest if I is a constant matrix (that is, $b(y)\equiv$constant), $(\alpha,\beta)^\top$ is Diophantine of Diophantine type (γ,τ) for fixed $\gamma>0,\tau>1$, and N'' is everywhere non-zero in the domain (for example, $N(y)=y^2/2$). We note that the Wronskian of B in (3.5) is identically zero.

With non-autonomous perturbations, (3.4), (3.5) are no longer perpetually stable in general. Let G denote a domain in the y - space. Applying Corollary 2 and Theorem 4 1), we have the following.

Proposition 3.4. *Consider* (3.4).

1) *If* $\omega=\alpha N_{y_1}+\beta N_{y_2}\neq 0$ *on* G, *then* (3.4) *is KAM stable on* G.

2) *If* $|\alpha|+|\beta|\neq 0$ *and if* N *is quasi-convex on* G, *then* (3.4) *is Nekhoroshev stable on* G.

Applying Theorem 1 1) and Theorem 4 2), we also have the following.

Proposition 3.5. *Consider* (3.5). *Let* b *be a non-zero constant and* $N''\neq 0$ *on* G.

1) *If* $(\alpha,\beta)^\top$ *is Diophantine of Diophantine type* (γ,τ) *for fixed* $\gamma>0,\tau>2$, *then* (3.5) *is KAM stable on* G.

2) *If* α,β *are integer multiple of each other, then* (3.5) *is Nekhoroshev stable on* G.

Proof. 2) follows immediately from Theorem 1 1).

To prove 1), we let $\hat\omega$ be the frequency of P in t, $\omega(y)=bN'(y)(\alpha,\beta)^\top$, $\Omega(y)=(\omega(y),\hat\omega)^\top$, and

$$G_\varepsilon=\{y\in G:\ |\langle k,\Omega(y)\rangle|>\frac{\varepsilon^{\frac{1}{36}}}{|k|^\tau},\ \ \forall k\in Z^3\setminus\{0\}\}.$$

For each $k=(k_1,k_2,k_3)\in Z^3\setminus\{0\}$, consider the set

$$S_k=\{y\in G:\ |\alpha k_1+\beta k_2+a(y)k_3|<\frac{\varepsilon^{\frac{1}{36}}}{|k|^\tau}\},$$

where $a(y)=\hat\omega/(bN'(y))$. Since $a(y)$ is never zero on G, a standard measure estimate shows that there exists a constant $c>0$ independent of k,ε such that

$$|S_k|\le\frac{c\varepsilon^{\frac{1}{36}}}{|k|^\tau},$$

if $k_3\neq 0$. By noting that $S_k=\emptyset$ if $k_3=0$, we have that

$$|\cup_k S_k|=O(\varepsilon^{\frac{1}{36}}).$$

Hence

$$|G \setminus G_\varepsilon| = O(\varepsilon^{\frac{1}{36}}).$$

By Theorem 4 2), we have that for any $y \in G_\varepsilon$, there exists an invariant 3-torus of (3.5) with the frequency $\Omega(y)$. □

Linking the above results to the fluid flow kinematics in 3D physical space, one can conclude from the perpetual stability that fluid transport and mixing are not possible in steady fluid particle path flows like (3.4), (3.5) near elliptic vortex lines under structure preserving, steady fluid perturbations which can be viscid and compressible. As shown in [52], in the case that (3.1) admits a family of hyperbolic vortex lines in the domain of interest in the (z_1, z_2)-plane, fluid transport and mixing is possible as a result of chaos, under a suitable steady perturbation of the Euler fluid particle path flow. In fact, the resulting chaotic dynamics provide enhancement to the fluid transport and mixing.

However, fluid transport and mixing are possible in fluid particle path flows like (3.4), (3.5) near elliptic vortex lines under (possibly viscid and compressible) non-steady fluid perturbations. Under the conditions of Propositions 3.4, 3.5, the existence of KAM tori provides barriers to fluid transport and mixing, and, the Nekhoroshev stability exhibits the exponentially slow diffusion of fluid flows before possible turbulence takes place.

3.4. **Nekhoroshev stability** *vs* **KAM stability.** A standard, quasi-convex, integrable Hamiltonian system simultaneously enjoys KAM and Nekhoroshev stabilities under small perturbations. In fact, a deep connection between the KAM and Nekhoroshev stability for a convex standard Hamiltonian system is discovered in a work of Giorgilli and Morbidelli ([28]) in the sense that KAM tori can be constructed by iterating the Nekhoroshev stability estimates on nested domains $\{D_r : 0 \le r \le \infty\}$, characterized by exponential stability times $\{T_r : r \ge 0\}$ with $T_r \to \infty$ exponentially as $r \to \infty$, such that $D_0 = G$ and $D_\infty = G_\varepsilon$ - the KAM Cantor-like set.

For quasi-convex generalized Hamiltonian systems, due to the appearance of general Poisson structures which are not necessarily in compatible with the standard symplectic ones, we conjecture that KAM and Nekhoroshev stabilities can be completely independent of each other, that is, one can have a quasi-convex generalized Hamiltonian system which is Nekhoroshev stable but not KAM stable especially when $l < n$, and vice versa especially when $l > n$. In fact, as already shown in Proposition 3.5 for 3D time-periodic fluid particle path flow (3.5) $(l = 1, n = 2)$ with b being a non-zero constant that KAM stability requires Diophantine property of (α, β), while Nekhoroshev stability crucially depends on the resonancy of (α, β).

The following example shows that a Nekhoroshev but non-KAM stable generalized Hamiltonian system can be easily obtained when the number of non-resonant components of frequency map $\omega : \Omega \subset R^l \to R^n$ is straightly less then n.

Example 3.1. *Let $l, n \ge 2$ be arbitrary, $N(y)$ be convex in the sense of* QC), *B be a $l \times n$ constant matrix with integer entries and of rank $< n$, and C be any $n \times n$, skew symmetric, matrix valued function. Then the conditions of* Corollary 1 1) *are satisfied hence the Nekhoroshev stability of the respective perturbed generalized Hamiltonian systems follows. We note that neither* R) *nor* R1) *is satisfied for $\omega(y) = -B^\top \partial N / \partial y(y)$. In fact, it is easy to see that the frequency map ω is* *everywhere resonant. Hence all unperturbed n-tori can be destroyed via generic perturbations.*

It is clear that when $l < n$ and B is any $l \times n$ constant matrix consisting of integer entries, the above example holds for any quasi-convex Hamiltonian $N(y)$. Hence the phenomenon of non-KAM, Nekhoroshev stability is typical when $l < n$ and B is a constant matrix. Indeed, in the case $l < n$ and B is a constant matrix, the frequency map will never be non-degenerate in the sense of R) or R1) even when B admits irrational entries, hence KAM stability is not generally expected.

References

[1] R. Abraham and J. Marsden, Foundations of Mechanics, 2nd ed., Benjamin, 1978

[2] V. I. Arnold, Proof of a theorem by A. N. Kolmogorov on the preservation of quasi-periodic motions under small perturbations of the Hamiltonian, *Usp. Mat. Nauk.* **18** (1963), 13-40

[3] V. I. Arnold, Instability of dynamical systems with several degrees of freedom, *Sov. Mat. Dokl.* **5** (1966), 581-585

[4] V. I. Arnold, Mathematical methods of classical mechanics (translated from Russian by K. Vogtmann and A. Weinstein), Springer-Verlag, New-York Heidelberg Berlin, 1978

[5] D. Bambusi, A Nekhoroshev-type theorem for the Pauli-Fierz model of classical electrodynamics, *Ann. Inst. H. Poincaré Phys. Théor.* **60** (1994), 339-371

[6] G. Benettin, Applications of the Nekhoroshev theorem to celestial mechanics (Italian), *Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat.* **(8) 4** (2001), 71-95

[7] G. Benettin, F. Fasso and M. Guzzo, Nekhoroshev-stability of $L4$ and $L5$ in the spatial restricted three-body problem, J. Moser at 70 (Russian), *Regul. Chaotic Dyn.* **3** (1998), 56-72

[8] G. Benettin, L. Galgani and A. Giorgilli, A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems, *Celestial Mech.* **37** (1985), 1-25

[9] G. Benettin and G. Gallavotti, Stability of motion near resonances in quasi-integrable Hamiltonian systems, *J. Stat. Phys.* **44** (1986), 293-338

[10] G. Benettin, L. Galgani and A. Giorgilli, The dynamical foundations of classical statistical mechanics and the Boltzmann-Jeans conjecture, Seminar on Dynamical Systems (St. Petersburg, 1991), 3-14, *Progr. Nonlinear Differential Equations Appl.,* 12, Birkhäuser, Basel, 1994

[11] G. Benettin, P. Hjorth and P. Sempio, Exponentially long equilibrium times in a one-dimensional collisional model of classical gas, *J. Statist. Phys.* **94** (1999), 871-891

[12] U. Bessi, L. Chierchia and E. Valdinoci, Upper bounds on Arnold diffusion times via Mather theory, *J. Math. Pures Appl.* **(9) 80** (2001), 105-129

[13] O. I. Bogoyavlenskij, Theory of tensor invariants of integrable Hamiltonian systems: I. Incompatible Poisson structures, *Comm. Math. Phys. bf 180* (1996), 529-586 : II. Theorem on symmetries and its applications, *Comm. Math. Phys. bf 184* (1997), 301-365

[14] O. I. Bogoyavlenskij, Extended integrability and bi-Hamiltonian systems, Comm. Math. Phys. bf 186 (1998), 19-51

[15] J. R. Cary and R. G. Littlejohn, Hamiltonian mechanics and its application to magnetic field line flow, *Ann. Phys.* **151** (1982), 1-34

[16] A. Celletti and L. Ferrara, An application of the Nekhoroshev theorem to the restricted three-body problem, *Celestial Mech. Dynam. Astronom.* **64** (1996), 261-272

[17] F. Z. Cong and Y. Li, Existence of higher dimensional invariant tori for Hamiltonian systems, *J. Math. Anal. Appl.* **222** (1998), 255-267

[18] L. Chierchia and G. Gallavotti, Drift and diffusion in phase space, *Ann. Inst. H. Poincaré B* **60** (1994), 1-144

[19] S.-N. Chow, Y. Li and Y. Yi, Persistence of invariant tori on sub-manifolds in Hamiltonian systems, in *J. Nonl. Sci.* **12** (2002), 585-617

[20] A. Delshams and P. Gutiérrez, Effective stability and KAM theory, *J. Differential Equations* **128** (1996), 415-490

[21] I. Y. Dorfman, Dirac structures and integrability of nonlinear evolution equations, Wiley and Sons, New York, 1993

[22] J. J. Duistermaat, On global action-angle coordinates, *CMAP* **33** (1980), 687-706

[23] F. Fasso, M. Guzzo and G. Benettin, Nekhoroshev-stability of elliptic equilibria of Hamiltonian systems, *Comm. Math. Phys.* **197** (1998), 347-360

[24] E. Fontich and P. Martin, Arnold diffusion in perturbations of analytic integrable Hamiltonian systems, *Discrete Contin. Dynam. Systems* **7** (2001), 61-84

[25] G. Gallavotti, Arnold's diffusion in isochronous systems, *Math. Phys. Anal. Geom.* **1** (1998/99), 295-312

[26] A. Giorgilli, Energy equipartition and Nekhoroshev-type estimates for large systems, Hamiltonian dynamical systems (Cincinnati, OH, 1992), 147-161, IMA Vol. Math. Appl., 63, Springer, New York, 1995

[27] A. Giorgilli, A. Delshams, E. Fontich, L. Galgani and C. Simó, Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three-body problem, *J. Differential Equations* **77** (1989), 167-198.

[28] A. Giorgilli and A. Morbidelli, Invariant KAM tori and global stability for Hamiltonian systems, *Z. angew. Math. Phys.* **48** (1997), 102-134

[29] A. Giorgilli and E. Zehnder, Exponential stability for time dependent potentials, *ZAMP* **43** (1992), 827-855

[30] M. Guzzo, Nekhoroshev stability of quasi-integrable degenerate Hamiltonian systems, *Regul. Chaotic Dyn.* **4** (1999), 78-102

[31] M. R. Herman, Différéntiabilité optimal et contreexemples à la fermeture en topologie C^∞ des orbites récurrentes de flots hamiltoniens, *C. R. Acad. Sci. Paris. Sér. I Math.* **313** (1991), 49-51

[32] M. R. Herman, Exemples de flots Hamiltoniens dont aucune perturbation en topologie C^∞ n'a d'orbites périodiques sur un ouvert de surfaces d'energies, *C. R. Acad. Sci. Paris. Sér. I Math.* **312** (1991), 989-994

[33] B. Hernández-Bermejo, One solution of the 3D Jacobi identities allows determining an infinity of them, *Phy. Lett. A* **287** (2001), 371-378

[34] B. Hernández-Bermejo, New solutions of the Jacobi equations for three-dimensional Poisson structures, *J. Math. Phy.* **42** (2001), 4984-4996

[35] D. D. Holm and K. B. Wolf, Lie-Poisson description of Hamiltonian ray optics, *Phys. D* **51** (1991), 189–199

[36] T. Kappeler and M. Makarov, On action-angle variables for the second Poisson bracket of Kdv, *Comm. Math. Phys.* **214** (2000), 651-677

[37] A. N. Kolmogorov, On the conservation of conditionally periodic motions for a small change in Hamilton's function, *Dokl. Akad. Nauk. SSSR* **98** (1954), 525-530

[38] Y. Li and Y. Yi, Persistence of hyperbolic tori in Hamiltonian systems, *J. Differential Equations* **208** (2005), 344-387

[39] Y. Li and Y. Yi, On Poincaré-Treshchev tori in Hamiltonian systems, *Proc. Equadiff 2003*, Dumortier et al (Ed.), World Scientific, 2005, 136-151

[40] Y. Li and Y. Yi, Persistence of invariant tori in generalized Hamiltonian systems, *Erg. Th. Dyn. Sys.* **22** (2002), 1233-1261

[41] P. Lochak, Canonical perturbation theory via simultaneous approximation, *Russian Math. Surveys* **47** (1992), 57-133

[42] P. Lochak, Hamiltonian perturbation theory: periodic orbits, resonances and intermittency, *Nonlinearity* **6** (1993), 885-904

[43] P. Lochak, Arnold diffusion: a compendium of remarks and questions, Hamiltonian systems with three or more degree of freedom (S'Agaró, 1995), 168-183, NATO Adv. Sci. Inst. Ser.C Math. Phys. Sci., 533, Kluwer Acad. Publ., Dordrecht, 1999

[44] P. Lochak and A. I. Neishtadt, Estimates of stability times for nearly integrable systems with quasiconvex Hamiltonian, *Chaos* **2** (1992), 495-499

[45] L. Markus and K. R. Meyer, Generic Hamiltonian systems are neither integrable nor ergodic, *Mem. Amer. Math. Soc.* **144** 1974

[46] J. Marsden and T. Ratiu, Introduction to Mechanics and Symmetry, Texts in Applied Mathematics, vol. 17, Springer-Verlag, 2nd Ed., 1999

[47] J. Marsden, T. Ratiu and A. Weinstein, Semi-direct products and reduction in mechanics, *Trans. Amer. Math. Soc.,* **281** (1984), 147-177.

[48] J. Marsden and A. Weinstein, Comments on the history, theory, and applications of symplectic reduction, in *Quantization of Singular Symplectic Quotients* (Landsman et al ed.) Birkhuser Boston, 2001, 1-20

[49] J. Marsden and A. Weinstein, Reduction of symplectic manifolds with symmetry, *Rep. Math. Phys.* **5** (1974), 121-130

[50] C. Martin, The Poisson structure of the Mean-Field equations in the ϕ^4 theory, *Ann. Phy.* **271** (1999), 294-305

[51] K. R. Meyer, Symmetries and integrals in mechanics, in *Dynamical systems* (Peixoto ed.), Academic Press, New York, 1973, 295-273

[52] I. Mezić and S. Wiggins, On the integrability and perturbations of three-dimensional fluid flows with symmetry, *J. Nonl. Sci.* **4** (1994), 157-194

[53] P. J. Morrison, Hamiltonian description of the ideal fluid, *Rev. Mod. Phy.* **70** (1998), 467-521

[54] J. Moser, On invariant curves of area preserving mappings of an annulus, *Nachr. Akad. Wiss. Gött. Math. Phys.* **K1** (1962), 1-20

[55] J. Moser, Old and new applications of KAM theory, Hamiltonian systems with three or more degree of freedom (S'Agaró, 1995), 184-192, NATO Adv. Sci. Inst. Ser.C Math. Phys. Sci., 533, Kluwer Acad. Publ., Dordrecht, 1999

[56] N. N. Nekhoroshev, An exponential estimate of the time of stability of nearly integrable Hamiltonian systems I, *Usp. Mat. Nauk,* **32** (1977), 5-66; *Russian Math. Surveys* **32** (1977), 1-65

[57] L. Niederman, Exponential stability for small perturbations of steep integrable Hamiltonian systems, to appear in *Erg. Th. Dyn. Sys.*

[58] P. J. Olver, Applications of Lie Groups to Differential Equations, Springer-Verlag, New York, 1993

[59] P. J. Olver, BiHamiltonian systems, Pitman research notes in Mathematics **157** (B. D. Sleeman and R. J. Jarvis, eds.) Longman Scientific and Technical, 1987, 176-193

[60] I. O. Parasyuk, On preservation of multidimensional invariant tori of Hamiltonian systems, *Ukrain Mat. Zh.* **36** (1984), 467-473

[61] G. Picard and T. W. Johnston, Instability cascades, Lotka-Volterra population equations, and Hamiltonian Chaos, *Phys. Rev. Lett. 48* (1982), 1610-1613

[62] G. Popov, Invariant tori, effective stability, and quasimodes with exponentially small error terms. I. Birkhoff normal forms, *Ann. Henri Poincaré* **1** (2000), 223-248

[63] G. Popov, Invariant tori, effective stability, and quasimodes with exponentially small error terms. II. Quantum Birkhoff normal forms, *Ann. Henri Poincaré* **1** (2000), 249-279

[64] J. Pöschel, Nekhoroshev estimates for quasiconvex Hamiltonian systems, *Math. Z.* **213** (1993), 495-499

[65] H. Rüssmann, Nondegeneracy in the perturbation theory of integrable dynamical systems, *Number theory and dynamical systems* (York, 1987), 5-18, London Math. Soc. Lecture Note Ser., 134, Cambridge Univ. Press, Cambridge, 1989, *Stochastics, algebra and analysis in classical and quantum dynamics* (Marseille, 1988), 211-223, Math. Appl., 59, Kluwer Acad. Publ., Dordrecht, 1990

[66] H. Rüssmann, Pravite communications, 2005.

[67] M. B. Sevryuk, KAM-stable Hamiltonians, *J. Dyn. Control Syst.* **1**(1995), 351-366

[68] W. M. Schmidt, Diophantine Approximation, Lecture Notes in Math. **785,** Springer, 1980

[69] A. Weinstein, The local structure of Poisson manifolds, *J. Differential Geometry* **18** (1983), 523-557

[70] J. X. Xu, J. G. You and Q. J. Qiu, Invariant tori for nearly integrable Hamiltonian systems with degeneracy, *Math. Z.* **226** (1997), 375-387

[71] J-C. Yoccoz, Travaux de Herman sur les tores invariants, *Asterisque* **206** (1992), 311-344

Y. Li: Department of Mathematics, Jilin University,, Changchun, 130012, P. R. China
E-mail address: `liyong@jlu.edu.cn`

Y. Yi: School of Mathematics, Georgia Institute of Technology,, Atlanta, GA 30332, USA
E-mail address: `yi@math.gatech.edu`