Persistence of invariant tori in generalized Hamiltonian systems

Yong Li* and Yingfei Yi†

Abstract

We present some results of KAM type, comparable to the KAM theory for volume-preserving maps and flows, for generalized Hamiltonian systems which may admit a distinct number of action and angle variables. In particular, systems under consideration can be odd dimensional. Applications to the perturbation of three dimensional steady Euler fluid particle path flows are considered with respect to the existence problem of barriers to fluid transport and mixing.

Keywords. Euler flows, KAM tori, generalized Hamiltonian systems

AMS(MOS). Mathematics Subject Classification. 58F05, 58F27, 58F30.

1 Introduction and main results

The celebrated KAM theory (Kolmogorov [17], Arnold [1], Moser [20]) asserts that with respect to the standard symplectic form on a $2n$ dimensional smooth manifold, most invariant n-tori of an integrable Hamiltonian system under certain non-degenerate condition will survive after small perturbations. The same was shown by Parayuk [24], Herman [15],[16](also see a survey by Yoccoz in [33]), Moser [22] in Hamiltonian coisotropic context, i.e., nearly integrable Hamiltonian systems defined on a symplectic manifold $(R^l \times T^n, \omega^2)$, $l < n$, $l + n$ is even. The symplectic forms above were assumed to have constant coefficients. The non-constant-coefficient case was treated in a recent work of Cong and Li [9].

Due to important technical reasons, the development of KAM theory for "odd dimensional" systems has been considered as a challenging problem, as pointed out in de la Llave [11], Mezić and Wiggins [19], Sevryuk [29]. Particularly including "odd dimensional" systems, KAM type of theory has been developed for volume preserving flows (see Broer, Huitema and Takens [3], Broer, Huitema and Sevryuk [4],[5], Delshams and R. de la Llave [12]) and for diffeomorphisms which are either volume-preserving with one action variable (see Cheng and Sun [8], Delshams and R. de la Llave [12], Herman

*Department of Mathematics, Jilin University, Changchun, 130012, P. R. China. e-mail: ylimd@mail.jl.cn

†School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA, and, Department of Computational Science, National University of Singapore, Singapore 119260. e-mail: yi@math.gatech.edu

[14], Xia [30], Yoccoz [33]) or satisfying the so-called intersection property - a relaxed version of volume preservation (see Cong, Li and Huang [10], Xia [31]).

In the spirit of works in volume preserving context, the aim of the present paper is to prove some KAM type of results for generalized, nearly integrable Hamiltonian systems defined on Poisson manifolds (i.e., the underlying 2-forms are not necessary non-degenerate) which particularly include systems in odd dimensions and systems which admit more action than angle variables. The systems to be considered need not satisfy the intersection property (hence not necessary volume preserving) but they need to preserve a prescribed Poisson structure. The development of KAM theory in the generalized setting is important especially when considering applications arising in the perturbation of three dimensional incompressible fluid flows (see [6],[23],[19]). This is in fact one of the main motivation of the present work.

A so-called generalized Hamiltonian system is defined on a Poisson manifold which can be odd dimensional and structurally degenerate. Consider the manifold $G \times T^n$, where $G \subset R^l$ is a bounded, connected, and closed region, T^n is the standard n-torus, l, n are positive integers. Let $I = (A_{ij}) : G \times T^n \to R^{(l+n)\times(l+n)}$ be a real analytic, anti-symmetric, matrix valued function, called *structure matrix*, which satisfies rank$I > 0$ and the Jacobi identity:

$$\sum_{m=1}^{l+n}(A_{im}\frac{\partial A_{jk}}{\partial z_m} + A_{jm}\frac{\partial A_{ki}}{\partial z_m} + A_{km}\frac{\partial A_{ij}}{\partial z_m}) = 0 \tag{1.1}$$

for all $z = (y, x) \in G \times T^n$ and $i, j, k = 1, 2, \cdots, l + n$. Such a structure matrix defines a Poisson structure or a 2-form ω^2: $\omega^2(\cdot, I\omega^1) = \omega^1(\cdot)$, for all 1-form ω^1 defined on $G \times T^n$; which can be also determined in the following way:

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

for all smooth functions f_1 and f_2 defined on $G \times T^n$, where $\{\cdot, \cdot\}$ denotes the Poisson bracket and ∇ denotes the standard Euclidean gradient on $R^l \times T^n$. The 2-form ω^2 is required to be invariant relative to T^n. Hence its coefficients, or equivalently, the structure matrix I, is independent of $x \in T^n$, i.e., $I = I(y)$, $y \in G$.

On the Poisson manifold $(G \times T^n, \omega^2)$, we consider the following Hamiltonian

$$H(y, x) = N(y) + \varepsilon P(y, x), \tag{1.2}$$

where N and P are real analytic functions, $\varepsilon > 0$ is a small parameter. Then the equation of motions of (1.2) associated to the 2-form ω^2 reads

$$\begin{pmatrix} \dot{y} \\ \dot{x} \end{pmatrix} = I(y)\nabla(N(y) + \varepsilon P(y, x)). \tag{1.3}$$

We further require that the unperturbed system associated to (1.2) or (1.3) is completely integrable, i.e., $y = (y_1, y_2, \cdots, y_l)^\top \in G$ need to satisfy the involution conditions:

$$\{y_i, y_j\} = 0, \quad i, j = 1, 2, \cdots, l.$$

Hence the structure matrix I must have the following form

$$\begin{pmatrix} O & B \\ -B^\top & C \end{pmatrix} \tag{1.4}$$

where $O = O_{l,l}$, $B = B_{l,n}$, $C = C_{n,n}$ with $C^\top = -C$.

We refer the system (1.3) as a *generalized Hamiltonian system.* Obviously, when $n = l$ and $I \equiv J$ - the standard symplectic matrix, $(G \times T^n, \omega^2)$ becomes the usual symplectic manifold and (1.3) becomes a standard Hamiltonian system. But in a generalized system I is clearly singular on G if $l > n$ or $n + l$ is odd, in which cases, the 2-form ω^2 becomes degenerate (hence not symplectic in the usual sense). This kind of singularity demonstrates an essential difference between a generalized Hamiltonian system and a standard one.

Let $\varepsilon = 0$ in (1.3). Then the unperturbed system associated to $N(y)$ reads

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

where

$$(\underbrace{0, \cdots, 0}_{l}, \omega_1(y), \cdots, \omega_n(y))^\top = I(y) \nabla N(y). \tag{1.5}$$

Hence, the phase space $G \times T^n$ is foliated into invariant n-tori $\{T_y : y \in G\}$ carrying parallel flows.

Similar to the KAM theory for standard Hamiltonian systems, the persistence of these invariant n-tori for the generalized system (1.3) will be subjected to certain non-degenerate conditions of the associated unperturbed system, by taking into account of the associated Poisson structure. We thus assume the following Rüssmann like non-degenerate condition:

R) $\max_{y \in G} \operatorname{rank} \left\{ \frac{\partial^i \omega}{\partial y^i} : |i| \leq n - 1 \right\} = n,$ *where* $i \in Z_+^n$ *and* $|i| = \sum_{j=1}^n |i_j|$.

As shown in [32] for standard Hamiltonian systems, the condition **R)** is equivalent to the following Rüssmann non-degenerate condition ([27]): the frequencies $\{\omega(y) : y \in G\}$ do not lie in any hyperplane in R^n. The Rüssmann condition is known to be the weakest non-degenerate condition for KAM tori to persist in standard Hamiltonian systems, as shown in works of Cheng and Sun ([7]), Sevryuk ([28]), and Xu, You and Qiu ([32]).

Our main result states as the following.

Theorem 1 *Consider* (1.3) *and assume the non-generate condition* **R)**. *Then there is an* $\varepsilon_0 > 0$ (*depending on* l, n, I, H, *a complex neighborhood of* $G \times T^n$, *and a Diophantine constant* τ *to be specified below*) *and a family of Cantor sets* $G_\varepsilon \subset G$, $0 < \varepsilon \le \varepsilon_0$, *for which the following holds.*

1) *For any* $y \in G_\varepsilon$, *the unperturbed torus* T_y *persists and gives rise to an analytic, Diophantine, invariant* n*-torus of the perturbed system whose toral frequency* $\omega_\varepsilon(y)$ *is of the Diophantine type* (γ, τ) *for some* $0 < \gamma \le \varepsilon^{\frac{1}{8n+12}}$ *and a fixed* $\tau > \max\{0, l(l-1)-1, n(n-1)-1\}$. *Moreover, the perturbed tori form a Whitney smooth family.*

2) *The Lebesgue measure* $|G \setminus G_\varepsilon| = O(\varepsilon^{\frac{1}{4(2n+3)(l_*-1)}}) \to 0$ *as* $\varepsilon \to 0$, *where*

$$l_* = \begin{cases} 2, & \text{if } n = 1, \\ \max\{l, n\}, & \text{if } n > 1. \end{cases}$$

3) *If* I *is a constant matrix and if the Hessian matrix* $\dfrac{\partial^2 N}{\partial y^2}$ *is non-singular on* G, *then all unperturbed Diophantine tori of the Diophantine types* (γ, τ), *with* $0 < \gamma \le \varepsilon^{\frac{1}{8n+12}}$ *and a fixed* $\tau > n-1$, *will persist and give rise to perturbed tori which preserve their corresponding unperturbed toral frequencies.*

In the above (and also below), a toral frequency $\omega \in R^n$ or its corresponding torus is said to be Diophantine of type (γ, τ) if

$$|\langle k, \omega \rangle| > \frac{\gamma}{|k|^\tau}, \qquad k \in Z^n \setminus \{0\}.$$

We note that when I is invertible and $n + l$ is even, parts 1), 2) of Theorem 1 are main results of [9].

Due to the possible lack of a sufficient number of action variables, the non-degenerate condition **R)** above can fail in applications (see Example 5.1 in Section 5)). For a KAM type of result to hold in this situation, additional deformation parameters are needed so that co-nondegeneracy with respect to both parameters and action variables can be considered (this is also necessary in some cases of standard Hamiltonian systems, see [5]). We thus consider the following generalized Hamiltonian system:

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

where ξ is a parameter lying in a bounded closed region Ξ of a Euclidean space R^p, $x \in T^n, y \in G \subset R^l, \varepsilon, I, N, P$ are as in (1.3), except that I, N, P now depend on the parameter ξ in the real analytic fashion. Instead of **R)**, we now assume the following coupling non-degenerate condition:

R1) $\max_{y\in G,\xi\in\Xi} \operatorname{rank}\left\{\frac{\partial^i \omega}{\partial(y,\xi)^i} : |i| \le n-1\right\} = n$, *where*

$$(\underbrace{0,\cdots,0}_{l},\omega_1(y,\xi),\cdots,\omega_n(y,\xi))^\top = I(y,\xi)\nabla N(y,\xi).$$

We have the following result similar to Theorem 1.

Theorem 2 *Consider* (1.6) *and assume the non-degenerate condition* **R1)**. *Then there is an* $\varepsilon_0 > 0$ (*depending on* l, n, p, I, H, *a complex neighborhood of* $G \times \Xi \times T^n$, *and a Diophantine constant* τ *to be specified below*) *and a family of Cantor sets* $G_\varepsilon \subset G \times \Xi$, $0 < \varepsilon \le \varepsilon_0$, *for which the following holds.*

1) *For any* $(y,\xi) \in G_\varepsilon$, *the unperturbed* n*-torus* $T_{y,\xi}$ *persists and gives rise to an analytic, Diophantine, invariant* n*-torus of the perturbed system whose toral frequency* $\omega_\varepsilon(y)$ *is of the Diophantine type* (γ,τ) *for some* $0 < \gamma \le \varepsilon^{\frac{1}{8n+12}}$ *and a fixed* $\tau > \max\{0, (l+p)((l+p)-1)-1, n(n-1)-1\}$. *Moreover, the perturbed tori form a Whitney smooth family.*

2) *The Lebesgue measure* $|G \setminus G_\varepsilon| = O(\varepsilon^{\frac{1}{4(2n+3)(l_*-1)}}) \to 0$ *as* $\varepsilon \to 0$, *where*

$$l_* = \begin{cases} 2, & \text{if } n = 1, \\ \max\{l+p, n\}, & \text{if } n > 1. \end{cases}$$

3) *If* I *is a constant matrix and if the Hessian matrix* $\frac{\partial^2 N}{\partial(y,\xi)^2}$ *is non-singular on* $G \times \Xi$, *then all unperturbed Diophantine tori of the Diophantine types* (γ,τ), *with* $0 < \gamma \le \varepsilon^{\frac{1}{8n+12}}$ *and a fixed* $\tau > n-1$, *will persist and give rise to perturbed tori which preserve their corresponding unperturbed toral frequencies.*

Remark 1.1 1) *The analyticity assumed in the perturbations in both theorems above can be weakened to sufficient smoothness, as in the standard KAM case* ([25]). *But then the matrices in the nondegenerate conditions* **R)** *and* **R1)** *should be assumed to have the constant rather than maximal rank* n *over their domains of definition.*

2) *Both theorems above resemble closely to the KAM type of results under the Lie algebra framework which originated in Moser* [21] *and was later extensively explored in Broer, Huitema and Takens* [3]. *The Lie algebra framework allows the consideration of*

the persistence of invariant tori for a class of integrable, structural preserving vector fields which is admissible, non-degenerate on the invariant tori in a definite sense (*see* [3]), *and forms a closed sub-algebra of the general Lie algebra of vector fields on a fixed phase space. Using this framework, several KAM type of results were obtained in* [3] *for a broad class of integrable vector fields under certain structural preserving perturbations, which particularly include volume preserving flows* (*with one action variable*) *and flows with symplectic structures. Comparing with these results,* Theorems 1,2 *above are both weaker and stronger in the following sense. On one hand, it is easy to see that with a fixed constant structure matrix* I *the unperturbed part of the generalized Hamiltonian vector field* (1.3) *is admissible in the sense of* [3]. *Thus, if* I *is a non-degenerate constant structure matrix, then* (1.3) *can well be considered under the Lie algebra framework along with an implicit function theorem argument similar to* [21]. *On the other hand, not only is the structure matrix in* (1.3) *allowed to be action-variable-dependent in our results, but also our results allow non-volume-preservation, more action than angle variables, odd dimensionality, and more importantly, the weak Rüssmann like non-resonance conditions* **R)** *or* **R1)**. *Yet, based on* [3] *and the present work, it is possible to have a unified persistence result of KAM type, similar to* Theorems 1,2 *above, for general structural preserving vector fields under the Lie algebra framework, by weakening the admissible class defined in* [3] *and assuming a Rüssmann like nondegenerate conditions on the invariant tori.*

One of the significant applications of the theorems above is the the study of fluid particle paths for certain three dimensional steady fluid flows under small viscosity. The existence of KAM type of invariant tori for such systems provides a mechanism and theoretical justification for the possible barriers to fluid transport and mixing. We will discuss this issue in details in Section 5 along with some examples. In Section 5, we will also apply Theorem 2 above to construct a Herman type of counterexample for the invalidity of general closing lemma on Poisson manifolds.

The proof of our results uses the standard KAM procedure. However, for the generalized systems like (1.3) or (1.6), not only does the y dependence of a structure matrix need to be taking into consideration in all KAM steps, but also iterative sequences need to be carefully selected to overcome difficulties due to the possible lack of parameters in systems like (1.3) and the degeneracy of the structure matrix in general. This will be made possible by introducing a linear iterative scheme which modifies the ones typically used in standard KAM theory (see Section 2 for details), by using the idea of adjusting frequencies in [13], and by applying a measure estimate from [32]. By assuming the non-degenerate condition **R)** or **R1)**, our work has shown that neither the oddness of dimension nor the degeneracy of the structure matrix is essential to the problem of the persistence of invariant tori in generalized Hamiltonian systems.

The next three sections of the present paper are mainly devoted to the proof of parts 1) 2) of Theorem 1. For simplicity, we will omit the details for the proof of part 3) of Theorem 1 and only outline the main ideas of it at the end of Section 4. Since the proof for Theorems 2 is almost identical to that for Theorem 1, we will only outline the main differences between the two at the end of Section 4.

Throughout the rest of sections, we will use the same symbol $|\cdot|$ to denote the Euclidean norm of vectors and the Lebesgue measure of sets, use $|\cdot|_D$ to denote the sup-norm of a function over a domain D, and use the symbol $\langle\cdot,\cdot\rangle$ to denote the usual inner product in Euclidean spaces. For given $r, s > 0$, we let

$$D(s,r) = \{(y,x) : |y| < s^2, |\mathrm{Im}x| < r\}, \quad D(s) = \{y : |y| < s^2\}$$

be the (s^2, r), s^2 complex neighborhoods of $G \times T^n$, G respectively.

Acknowledgment. The first author is partially supported by the National 973 Project: Nonlinearity in China, the NSFC grant 19971042 and the outstanding youth project of Ministry of Education of China, and the second author is partially supported by NSF grant DMS9803581. We are grateful to the referee for valuable comments in particular for pointing out the possible use of the Lie algebra approach in developing a KAM theory for general structural preserving vector fields.

2 KAM Step

In this section, we will show detailed construction and estimates for one KAM cycle in the proof of Theorem 1.

Let $y_0 \in G$ be arbitrarily given. The Taylor expansion of the Hamiltonian (1.2) about y_0 reads

$$H(y, y_0, x) = e_0(y_0) + \langle \Omega(y_0), y - y_0\rangle + \bar{h}(y - y_0) + \varepsilon P(y, x),$$

where $e_0(y_0) = N(y_0)$, $\Omega_0(y_0) = \dfrac{\partial N(y_0)}{\partial y}$, and $\bar{h}(y - y_0) = O(|y - y_0|^2)$. Using the transformation $(y - y_0) \to y$ in the above, we have

$$H(y, y_0, x) = e_0(y_0) + \langle \Omega(y_0), y\rangle + \bar{h}(y) + \varepsilon P(y + y_0, x).$$

Write

$$\bar{h}(y) = h_0(y) + h_*(y),$$

where $h_0(y)$ is the truncation of the Taylor series of $\bar{h}(y)$ up to order $2(n+1)+2$. Then we have the following normal form

$$H_0 = H(y, y_0, x) = N_0 + P_0, \tag{2.1}$$

$$N_0 = N_0(y, y_0) = e_0(y_0) + \langle \Omega_0(y_0), y \rangle + h_0(y),$$
$$P_0 = P_0(y, y_0, x) = h_*(y) + \varepsilon P(y + y_0, x).$$

Thus, as $\varepsilon = 0$, for each $y_0 \in G$, the invariant n-torus T_{y_0} associated to (1.2) or (1.3) corresponds to $\mathcal{T}_{y_0} = \{0\} \times T^n$ of (2.1). To show the persistence of the majority of these tori, we will use a KAM type of approach to eliminate the x-dependent terms in the perturbation, via a sequence of inductively constructed, *canonical* transformations (i.e., transformations which preserve the 2-form ω^2) defined on nested domains.

Define $m = 2(n+1) + 1$ and $\gamma = \varepsilon^{\frac{1}{4m}}$. We let a_0, b, σ, d be chosen such that $0 < a_0 < b \ll \sigma \ll 1, 0 < d \ll 1$, and,

$$\begin{aligned} &\frac{\sigma}{b+\sigma} - (b+\sigma) > 2a_0, \\ &2 - m(b+\sigma) - \sigma > \frac{3}{2}, \\ &\delta(1+b+\sigma) > 1, \end{aligned} \tag{2.2}$$

where $\delta = 1 - d$. These constants can be easily shown to exist by first setting $a_0 = b = d = 0$ in the above then use perturbation arguments.

To begin with the induction, we initially set $\mathcal{O}_0 = G$, $r_0 = \delta$, $\gamma_0 = 4\gamma$, $\beta_0 = s_0$, $s_0 = \varepsilon^{\frac{1}{2m}}$, $\mu_0 = \frac{1}{4}\varepsilon^{\frac{1}{m+1}}$. Without loss of generality, we assume that $0 < r_0, \beta_0, \gamma_0, \mu_0, s_0 \le 1$.

Clearly, as ε small, we have

$$|P_0|_{D(s_0, r_0) \times \mathcal{O}_0} \le \gamma_0 s_0^m \mu_0.$$

Suppose at a KAM step, say the ν-th step, we have arrived at a Hamiltonian

$$\begin{aligned} H &= H_\nu = N + P \\ N &= N_\nu = e(y_0) + \langle \Omega(y_0), y \rangle + h(y), \end{aligned} \tag{2.3}$$

where $(y, x) \in D = D_\nu = D(r, s)$, $r = r_\nu \le r_0, s = s_\nu \le s_0$, $y_0 \in \mathcal{O}$, $e(y_0) = e_\nu(y_0)$, $\Omega(y_0) = \Omega_\nu(y_0)$ are real analytic on $\mathcal{O}$, $h(y) = h_\nu(y, y_0) = O(y^2)$ is a polynomial in y of order less or equal to $m+1$, and, $h = h_\nu(y, y_0)$ and $P = P_\nu(y, y_0, x)$ are analytic in $(y, x) \in D$ and $y_0 \in \mathcal{O}$, and moreover,

$$|P|_{D \times \mathcal{O}} \le \gamma s^m \mu, \tag{2.4}$$

for some $0 < \mu = \mu_\nu \le \nu_0$, $0 < \gamma = \gamma_\nu \le \gamma_0$.

We will construct a canonical transformation $\Phi = \Phi_{\nu+1}$ which transforms the Hamiltonian (1.2), in smaller phase and frequency domains, to the desired Hamiltonian into the next KAM cycle (the $(\nu+1)$-th KAM step). Below, we show the details for one

KAM cycle by constructing the canonical transformation and estimating the transformed Hamiltonian, etc. For simplicity, quantities (domains, normal form, perturbation, etc.) in the next KAM cycle will be simply indexed by $+$ ($=\nu+1$) and we will not specify the dependence of P, P_+ etc. on their arguments. All constants $c_1 - c_5$ below are positive and independent of the iteration process. For simplicity, we will also use c to denote any intermediate positive constant which is independent of the iteration process.

Let $\tau > \max\{0, l(l-1)-1, n(n-1)-1\}$ be fixed and define

$$
\begin{aligned}
r_+ &= \delta r - d(1-\frac{\delta^2}{2})r_0,\\
\gamma_+ &= \frac{\gamma_0}{4} + \frac{\gamma}{2},\\
s_+ &= s^{1+b+\sigma},\\
K_+ &= ([\log\frac{1}{s}]+1)^3,\\
\Gamma(u) &= \sum_{0<|k|\le K_+} |k|^{n+\tau+2m+8} e^{-\frac{u}{8}},\\
\Delta_+ &= (\gamma(s_+^{m+1}\mu + \frac{s_+^{m+1}}{s}\mu) + s^{2m}\mu^2 + s^{m+2}\mu)\Gamma^3(r-r_+),\\
D_+ &= D(s_+, r_+),\\
\tilde{D} &= D(\beta_0, r_+ + \frac{5}{8}(r-r_+)),\\
D_* &= D(\frac{s}{2}, r_+ + \frac{6}{8}(r-r_+)),\\
D_{**} &= D(s, r_+ + \frac{7}{8}(r-r_+)),\\
D_i &= D(is_+, r_+ + \frac{i-1}{8}(r-r_+)), \quad i=1,2,\cdots,8.
\end{aligned}
$$

2.1 Truncation

Consider the Taylor-Fourier series of P:

$$P = \sum_{i\in Z_+^n, k\in Z^n} p_{ki} y^i e^{\sqrt{-1}\langle k,x\rangle},$$

and let

$$R = \sum_{|i|\le m+1, |k|\le K_+} p_{ki} y^i e^{\sqrt{-1}\langle k,x\rangle} \tag{2.5}$$

be the truncation of P to the order of K_+ in x and the order of $m+1$ in y.

We have the following estimate.

Lemma 2.1 *Assume that*

H1) $s_+ \le \frac{s}{16}$;

H2) $\int_{K_+}^{\infty} \lambda^n e^{-\lambda \frac{r-r_+}{16}} \mathrm{d}\lambda \le s^{(m+1)(1+b+\sigma)}$.

Then there is a constant c_1 such that

$$|P - R|_{D_8} \le c_1 \gamma (s^{(m+1)(1+b+\sigma)} + \frac{s_+^{m+1}}{s}) \mu.$$

Proof: Denote

$$\begin{aligned} I' &= \sum_{|k|>K_+} p_{ki} y^i e^{\sqrt{-1}\langle k,x\rangle}, \\ II' &= \sum_{|k|\le K_+, |i|>m+1} p_{ki} y^i e^{\sqrt{-1}\langle k,x\rangle} = \int \frac{\partial^p}{\partial y^p} \sum_{|k|\le K_+, |i|>m+1} p_{ki} e^{\sqrt{-1}\langle k,x\rangle} y^i \mathrm{d}y, \end{aligned}$$

where $\int$ is the pth order anti-derivative of $\frac{\partial^p}{\partial y^p}$ for $|p| = m+1$. Clearly,

$$P - R = I' + II'.$$

To estimate I', we note by the Cauchy estimate that

$$|\sum_{i\in Z_+^n} p_{ki} y^i| \le |P|_{D(s,r)} e^{-|k|r} \le \gamma s^m \mu e^{-|k|r}.$$

This together with H2) yields that

$$\begin{aligned} |I'|_{D_{**}} &\le \sum_{|k|\ge K_+} \gamma s^m \mu e^{-|k|r} e^{|k|(r_+ + \frac{7}{8}(r-r_+))} \\ &\le \gamma s^m \mu \sum_{|u|\ge K_+} |u|^n e^{-|u|\frac{r-r_+}{8}} \le \gamma s^m \mu \int_{K_+}^{\infty} \lambda^n e^{-\lambda\frac{r-r_+}{16}} \mathrm{d}\lambda \\ &\le \gamma s^{(m+1)(1+b+\sigma)} \mu. \end{aligned}$$

Hence

$$|P - I'|_{D_{**}} \le |P|_{D(s,r)} + |I'|_{D_{**}} \le 2\gamma s^m \mu.$$

Note that

$$\begin{aligned} & II' = P - I' - R, \\ & \frac{\partial^i (P - I')}{\partial y^i} = \frac{\partial^i II'}{\partial y^i}, \quad |i| > m+1, \end{aligned}$$

and by H1) that $D_8 \subset D_{**}$. It follows from the Cauchy estimate on D_{**} that

$$\begin{aligned}
|II'|_{D_8} &\le |\int \frac{\partial^i}{\partial y^i} \sum_{|k|\le K_+, |u|>m+1} p_{kuq} e^{\sqrt{-1}\langle k,x\rangle} y^u \mathrm{d}y|_{D_8} \\
&\le |\int |\frac{\partial^i}{\partial y^i}(P-I')|\mathrm{d}y|_{D_8} \\
&\le 2\left(\frac{1}{s-8s_+}\right)^{m+1} \gamma s^m \mu |\int \mathrm{d}y|_{D_8} \le 2^{m+1}\gamma\mu\frac{s_+^{m+1}}{s},
\end{aligned}$$

where $\int$ denotes the $|i|$-th order anti-derivative for $|i| = m+1$.

Thus,

$$|P-R|_{D_8} \le 2^{m+1}\gamma(s^{(m+1)(1+b+\sigma)} + \frac{s_+^{m+1}}{s})\mu.$$

■

2.2 Modified linear scheme

To transform (2.3) into the Hamiltonian in the next KAM cycle, we would like to find a canonical transformation Φ_+ to eliminate all resonant terms in R, i.e., all terms

$$p_{ki} y^i e^{\sqrt{-1}\langle k,x\rangle}, \ \ 0<|k|\le K_+, \ \ |i|\le m+1.$$

To do so, we first construct a generalized Hamiltonian F of the form

$$F = \sum_{0<|k|\le K_+, |i|\le m+1} f_{ki} y^i e^{\sqrt{-1}\langle k,x\rangle} \tag{2.6}$$

which satisfies the following linear equation:

$$\{N,F\} + R - [R] - Q = 0, \tag{2.7}$$

where

$$\begin{aligned}
[R] &= [R](y) = \frac{1}{(2\pi)^n}\int_{T^n} R(y,x)\mathrm{d}x, \\
Q &= \{h,F\} \\
&+ \sum_{0<|k|\le K_+, |i|\le m+1} \sqrt{-1}\langle k, (B^\top(y+y_0) - B^\top(y_0))\Omega(y_0)\rangle f_{ki} y^i e^{\sqrt{-1}\langle k,x\rangle}.
\end{aligned} \tag{2.8}$$

We note that, by subtracting the term Q which reflects the dependence of $\{N,F\}$ on I and h, (2.7) modifies the usual linear equation typically adopted in the KAM theory for standard Hamiltonian systems.

Substituting (2.5), (2.6) into (2.7) yields that

$$-\sum_{0<|k|\le K_+, |i|\le m+1} \sqrt{-1}\langle k, \omega(y_0)\rangle f_{ki} y^i e^{\sqrt{-1}\langle k,x\rangle} + \sum_{0<|k|\le K_+, |i|\le m+1} p_{ki} y^i e^{\sqrt{-1}\langle k,x\rangle} = 0,$$

where $\omega(y_0) = -B^\top(y_0)\Omega(y_0)$.

By comparing coefficients above, we obtain the following linear equations:

$$\sqrt{-1}\langle k, \omega(y_0)\rangle f_{ki} = p_{ki}, \quad 0 < |k| \le K_+, \quad |i| \le m+1. \tag{2.9}$$

Consider

$$\mathcal{O}_+ = \{y_0 \in \mathcal{O} : |\langle k, \omega(y_0)\rangle| > \frac{\gamma}{|k|^\tau} \quad 0 < |k| \le K_+\}. \tag{2.10}$$

It is clear that the equations (2.9) are uniquely solvable on $\mathcal{O}_+$. Moreover, all solutions f_{ki}, $0 < |k| \le K_+$, $|i| \le m+1$, are real analytic on $\mathcal{O}_+$. Thus we have found the desired generalized Hamiltonian F of form (2.6) which is real analytic in both $y_0 \in \mathcal{O}_+$ and $(y, x) \in D$.

Let $\Phi_+ = \phi_F^1$ be the time-1 map of the equation of motion associated to F, i.e.,

$$\begin{pmatrix} \dot{y} \\ \dot{x} \end{pmatrix} = I(y + y_0)\nabla F(y, x). \tag{2.11}$$

Then Φ_+ is a canonical transformation, and,

$$\begin{aligned}
H_+ &= H \circ \Phi_+ = H \circ \phi_F^1 = (N + R) \circ \phi_F^1 + (P - R) \circ \phi_F^1 \\
&= (N + R) + \{N, F\} + \int_0^1 \{R_t, F\} \circ \phi_F^t \mathrm{d}t + (P - R) \circ \phi_F^1 \\
&= (N + [R]) + (\{N, F\} + R - [R] - Q) + \int_0^1 \{R_t, F\} \circ \phi_F^t \mathrm{d}t \\
&+ (P - R) \circ \phi_F^1 + Q = (N + [R]) + \int_0^1 \{R_t, F\} \circ \phi_F^t \mathrm{d}t + (P - R) \circ \phi_F^1 + Q,
\end{aligned}$$

where $R_t = (1 - t)\{N, F\} + R$.

Let

$$\begin{aligned}
e_+ &= e + p_{00}, && (2.12) \\
\Omega_+ &= \Omega + p_{01}, && (2.13) \\
\omega_+ &= -B^\top \Omega_+, && (2.14) \\
h_+ &= h + \sum_{2 \le |i| \le m+1} p_{0i} y^i && (2.15) \\
N_+ &= N + [R] = e_+ + \langle \Omega_+(y_0), y\rangle + h_+, \\
P_+ &= \int_0^1 \{R_t, F\} \circ \phi_F^t \mathrm{d}t + (P - R) \circ \phi_F^1 + Q.
\end{aligned}$$

Then

$$H_+ = N_+ + P_+$$

is the new Hamiltonian in the next KAM cycle with the desired normal form N_+.

It should be pointed that it is due to the Jacobi identity (1.1) that the structure matrix $I(y+y_0)$ is kept unchanged at each KAM step. Let $I(z)$, $z=(y,x)$, be a structure matrix on $G\times T^n$ and $\phi_F^t(z)$ be the flow generated by a vector field $I(z)\nabla F(z)$. Then by (1.1) it is easy to see that

$$\frac{\partial\phi_F^t}{\partial z}(z)^\top I(z)\frac{\partial\phi_F^t}{\partial z}(z)=I(\phi_F^t(z)),$$

which implies the preservation of the Poisson structure on $G\times T^n$ under the transformation $z_1=\phi_F^1(z)$.

2.3 Estimate on the transformation

Lemma 2.2 *There is a constant c_2 such that the following holds.*

1) *On $\mathcal{O}_+$,*

$$|f_{ki}| \;\leq\; c_2|k|^\tau s^{m-2|i|}\mu e^{-|k|r},$$

for all $0<|k|\leq K_+$.

2) *On $D_*\times\mathcal{O}_+$,*

$$|F|,\ |F_x|,\ s^2|F_y|\leq c_2 s^m\mu\Gamma(r-r_+),$$

and, on $\tilde{D}\times\mathcal{O}_+$,

$$\partial_y^i F=0,\quad |i|>m+1,$$

and

$$|D^iF|\leq c_2\mu\Gamma(r-r_+),\quad |i|\leq m+1.$$

Proof: 1) By the standard Cauchy estimate, we have

$$s^{2|i|}|p_{ki}|\leq |P|_{D(s,r)}e^{-|k|r}\leq\gamma s^m\mu e^{-|k|r}. \tag{2.16}$$

The desired estimate now follows from (2.9),(2.10) and (2.16).

2) It follows immediately from 1) that, on $D_*\times\mathcal{O}_+$,

$$\begin{aligned}|F| &\leq \sum_{0<|k|\leq K_+,|i|\leq m+1}|f_{ki}||y^i|e^{|k|(r_++\frac34(r-r_+))}\\ &\leq cs^m\mu\Gamma(r-r_+).\end{aligned}$$

The rest of estimates on the derivatives of F follow from direct calculations and a similar argument as above. ■

Lemma 2.3 *Assume that*

H3) $c_2\mu\Gamma(r-r_+) < \frac{1}{8}(r-r_+)$;

H4) $c_2 s\mu\Gamma(r-r_+) < \frac{1}{4}s_+$;

H5) $c_2(s^{\frac{a_0}{2}} + \mu^{\frac{a_0}{2}})\Gamma^3(r-r_+) < 1$.

Then the following holds.

1) *Let* ϕ_F^t *be the flow generated by the equation* (2.11). *Then*

$$\phi_F^t : D_3 \longrightarrow D_4, \quad \text{for all} \quad 0 \le t \le 1.$$

2) $\Phi_+ : D_+ \to D(s,r)$.

3) *There is a constant* c_3 *such that*

$$\begin{aligned} |\phi_F^t - id|_{\tilde{D}\times\mathcal{O}_+} &\le c_3\mu\Gamma(r-r_+), \\ |D\phi_F^t - \mathrm{Id}|_{\tilde{D}\times\mathcal{O}_+} &\le c_3\mu\Gamma(r-r_+), \\ |D^i\phi_F^t|_{\tilde{D}\times\mathcal{O}_+} &\le c_3\mu\Gamma(r-r_+), \quad 2 \le i \le m+1, \end{aligned}$$

for all $0 \le |t| \le 1$.

4)

$$\begin{aligned} |\Phi_+ - id|_{\tilde{D}\times\mathcal{O}_+} &\le c_3\mu\Gamma(r-r_+), \\ |D\Phi_+ - \mathrm{Id}|_{\tilde{D}\times\mathcal{O}_+} &\le c_3\mu\Gamma(r-r_+), \\ |D^i\Phi_+|_{\tilde{D}\times\mathcal{O}_+} &\le c_3\mu\Gamma(r-r_+), \quad 2 \le i \le m+1. \end{aligned}$$

Proof: 1) Denote ϕ_{F1}^t, ϕ_{F2}^t as the components of ϕ_F^t in y, x planes respectively, and let X_F be the vector field on the right hand side of (2.11). Then

$$\phi_F^t = \mathrm{id} + \int_0^t X_F \circ \phi_F^\lambda \mathrm{d}\lambda.$$

For any $(y,x) \in D_3$, we let $t_* = \sup\{t \in [0,1] : \phi_F^t(y,x) \in D_4\}$. By H1), we have $D_4 \subset D_*$. It follows from H3), H4) and Lemma 2.2 that

$$\begin{aligned} |\phi_{F1}^t(y,x)| &= |y| + |\int_0^t B(\phi_{F1}^\lambda + y_0)F_x \circ \phi_F^\lambda \mathrm{d}\lambda| \le |y| + c|F_x|_{D_*} \\ &\le s_+ + c_2 s^m \mu\Gamma(r-r_+) < s_+ + 3s_+ = 4s_+, \\ |\phi_{F2}^t(y,x)| &= |x| + |\int_0^t (-B(\phi_{F1}^\lambda + y_0)F_y \circ \phi_F^\lambda + C(\phi_{F1}^\lambda + y_0)F_x \circ \phi_F^\lambda)\mathrm{d}\lambda| \\ &\le |x| + c(|F_x| + |F_y|)_{D_*} \le r_+ + \frac{2}{8}(r-r_+) + c_2 s^{m-2}\mu\Gamma(r-r_+) \\ &< r_+ + \frac{3}{8}(r-r_+), \end{aligned}$$

where B, C are the matrices defined in (1.4). This shows that $\phi_F^t(y,x) \in D_4$ for all $0 \le t \le t_*$. Hence $t_* = 1$ and 1) holds.

2) clearly follows from 1).

3) Using Lemma 2.2 and the argument above, we immediately have

$$|\phi_F^t - id|_{\tilde{D}} \le c_2 \mu \Gamma(r - r_+).$$

By applying H5), Lemma 2.2 and the Gronwall inequality to the following identity

$$\begin{aligned} D\phi_F^t &= \mathrm{Id} + \int_0^t (D(I\nabla F)) D\phi_F^\lambda \mathrm{d}\lambda \\ &= \mathrm{Id} + \int_0^t ((DI \cdot DF) \circ \phi_F^\lambda \cdot D\phi_F^\lambda + (ID^2F) \circ \phi_F^\lambda) \cdot D\phi_F^\lambda) \mathrm{d}\lambda, \end{aligned}$$

we have

$$\begin{aligned} |D\phi_F^t - \mathrm{Id}|_{\tilde{D}} &\le \int_0^t (|DI||DF| + |I||D^2F|)_{\tilde{D}} |D\phi_F^\lambda - \mathrm{Id}|_{\tilde{D}} \mathrm{d}\lambda \\ &\quad + (|DI||DF| + |I||D^2F|)_{\tilde{D}} \\ &\le c\mu\Gamma(r - r_+). \end{aligned}$$

The estimates on the higher order derivatives of ϕ_F^t follow from the induction and a similar argument.

4) follows from 3). ■

2.4 Estimate on the new Hamiltonian

Lemma 2.4 *There is a constant c_4 such that*

$$\begin{aligned} |\Omega^+ - \Omega|_{\mathcal{O}_+} &\le c_4 \gamma s^{m-2} \mu, \\ |\omega^+ - \omega|_{\mathcal{O}_+} &\le c_4 \gamma s^{m-2} \mu, \\ |e_+ - e|_{\mathcal{O}_+} &\le c_4 \gamma s^m \mu, \\ |h_+ - h|_{D(s_+) \times \mathcal{O}_+} &\le c_4 \gamma s^m \mu. \end{aligned}$$

Proof: The proof immediately follows from (2.4) and (2.12)-(2.15). ■

Lemma 2.5 *Assume that*

H6) $c_4 \gamma_0 s^{m-2} \mu < \dfrac{\gamma - \gamma_+}{K_+^{\tau+1}}$.

Then

$$|\langle k, \omega_+(y_0) \rangle| > \frac{\gamma_+}{|k|^\tau},$$

for all $y_0 \in \mathcal{O}_+$ and $0 < |k| \le K_+$.

Proof: By H6) and Lemma 2.4, one has

$$\begin{aligned} |\langle k, \omega_+(y_0)\rangle| &\geq |\langle k, \omega(y_0)| - c_4\gamma_0 s^{m-2}\mu K_+ \\ &\geq \frac{\gamma}{|k|^\tau} - c_4\gamma_0 s^{m-2}\mu K_+ > \frac{\gamma_+}{|k|^\tau}, \end{aligned}$$

as desired. ■

Lemma 2.6 *There is a constant c_5 such that*

$$|P_+|_{D_+} \leq c_5\Delta_+.$$

Thus, if

H7)

$$c_5\Delta_+ \leq \gamma_+ s_+^m \mu_+,$$

then

$$|P_+|_{D_+\times\mathcal{O}_+} \leq \gamma_+ s_+^m \mu_+.$$

Proof: Using (2.8), Lemma 2.2 2), and the Cauchy estimate, we have that

$$\begin{aligned} |Q|_{D_+\times\mathcal{O}_+} &\leq (|I||h_y||F_x|)_{D_*\times\mathcal{O}_+} + cs^{2+m}\mu\Gamma(r-r_+) \\ &\leq cs^{2+m}\mu\Gamma(r-r_+). \end{aligned}$$

Let

$$W = \int_0^1 \{R_t, F\}\circ\phi_F^t \mathrm{d}t.$$

Then Cauchy's estimate yields

$$\begin{aligned} |W|_{D_+\times\mathcal{O}_+} &\leq (|I|(|R_x||F_y| + |R_y||F_x| + |R_x||F_x| + |\{\{h, F\}, F\}|))_{D*\times\mathcal{O}_+} \\ &\leq cs^{2m}\mu^2\Gamma^2(r-r_+). \end{aligned}$$

Recall that

$$P_+ = W + (P-R)\circ\phi_F^1 + Q.$$

The above estimates together with Lemma 2.1 imply that

$$|P_+|_{D_+\times\mathcal{O}_+} \leq c(\gamma(s_+^{m+1}\mu + \frac{s_+^{m+1}}{s}\mu) + s^{2m}\mu^2 + s^{m+2}\mu)\Gamma^3(r-r_+) = c\Delta_+.$$

■

3 Iteration Lemma

In this section, we will prove an iteration lemma which guarantees the inductive construction of the canonical transformations in all KAM steps.

Let $r_0, s_0, \gamma_0, \mu_0, \mathcal{O}_0, H_0, N_0, e_0, \Omega_0, P_0$ be given as in the beginning of Section 2 and let $\tilde{D}_0 = D(r_0, \beta_0)$, $D_0 = D(r_0, s_0)$, $K_0 = 0$, $\Phi_0 = id$. For any $\nu = 0, 1\cdots$, we label all index-free qualities in Section 2 by ν and all '$+'$'-indexed qualities in Section 2 by $\nu + 1$. This defines the following sequences:

$$r_\nu,\ s_\nu,\ \mu_\nu,\ K_\nu,\ \mathcal{O}_\nu,\ D_\nu,\ \tilde{D}_\nu,\ H_\nu,$$
$$N_\nu,\ e_\nu,\ \Omega_\nu,\ \omega_\nu,\ h_\nu,\ P_\nu,\ \Phi_\nu,$$

for $\nu = 0, 1, \cdots$. In particular,

$$\begin{aligned} H_\nu &= H_\nu(y,x) = N_\nu + P_\nu, \\ N_\nu &= e_\nu + \langle \Omega_\nu, y\rangle + h_\nu(y), \end{aligned}$$

where $(y,x) \in \tilde{D}_\nu$, $y_0 \in \mathcal{O}_\nu$, $e_\nu = e_\nu(y_0)$, $\omega_\nu = \omega_\nu(y_0) = -B^\top(y_0)\Omega_\nu(y_0)$, $\Omega_\nu = \Omega_\nu(y_0)$ is analytic on $\mathcal{O}_\nu$, h_ν is a polynomial in y of order less or equal to $m+1$, and, $h_\nu = h_\nu(y)$ and $P_\nu = P_\nu(y,x)$ are analytic in $y_0 \in \mathcal{O}_\nu$ and $(y,x) \in \tilde{D}_\nu$. Moreover, for $\nu = 1, 2, \cdots$,

$$\begin{aligned} s_\nu &= s_{\nu-1}^{1+b+\sigma}, \\ \mu_\nu &= c_0 s_{\nu-1}^{\sigma}\mu_{\nu-1}, \quad c_0 = \max\{1, c_1, \cdots, c_5\}, \\ \gamma_\nu &= \gamma_0(1 - \sum_{i=1}^{\nu}\frac{1}{2^{i+1}}), \\ K_\nu &= ([\log\frac{1}{s_{\nu-1}}] + 1)^3, \\ \Delta_\nu &= (\gamma_\nu(s_\nu^{m+1}\mu_\nu + \frac{s_\nu^{m+1}}{s_{\nu-1}}\mu_\nu) + s_{\nu-1}^{2m}\mu_\nu^2 + s_{\nu-1}^{m+2}\mu_\nu)\Gamma^3(r_{\nu-1} - r_\nu), \\ \mathcal{O}_\nu &= \{y_0 \in \mathcal{O}_{\nu-1} : |\langle k, \omega_{\nu-1}(y_0)\rangle| > \frac{\gamma_{\nu-1}}{|k|^\tau},\ 0 < |k| \le K_\nu\}, \\ D_\nu &= D(r_\nu, s_\nu), \\ \tilde{D}_\nu &= D(r_\nu + \frac{7}{8}(r_{\nu-1} - r_\nu), \beta_0). \end{aligned}$$

Lemma 3.1 (**Iteration Lemma**) *Let $\mu_* = \mu_0^{1-a_0}$. If $\mu_0 = \mu_0(\varepsilon_0)$ is sufficiently small, then the following holds for all $\nu = 0, 1, \cdots$.*

1)

$$|e_\nu - e_0|_{\mathcal{O}_\nu} \le 2\gamma_0\mu_*, \tag{3.1}$$
$$|e_{\nu+1} - e_\nu|_{\mathcal{O}_{\nu+1}} \le \frac{\gamma_0\mu_*}{2^{\nu+1}}, \tag{3.2}$$

$$|\Omega_\nu - \Omega_0|_{\mathcal{O}_\nu} \le 2\gamma_0\mu_*, \tag{3.3}$$
$$|\Omega_{\nu+1} - \Omega_\nu|_{\mathcal{O}_{\nu+1}} \le \frac{\gamma_0\mu_*}{2^{\nu+1}}, \tag{3.4}$$
$$|\omega_\nu - \omega_0|_{\mathcal{O}_\nu} \le 2\gamma_0\mu_*, \tag{3.5}$$
$$|\omega_{\nu+1} - \omega_\nu|_{\mathcal{O}_{\nu+1}} \le \frac{\gamma_0\mu_*}{2^{\nu+1}}, \tag{3.6}$$
$$|h_\nu - h_0|_{D(s_\nu)\times\mathcal{O}_\nu} \le 2\gamma_0\mu_*, \tag{3.7}$$
$$|h_{\nu+1} - h_\nu|_{D(s_{\nu+1})\times\mathcal{O}_{\nu+1}} \le \frac{\gamma_0\mu_*}{2^{\nu+1}}, \tag{3.8}$$
$$\frac{1}{s_\nu^m}|P_\nu|_{D_\nu\times\mathcal{O}_\nu} \le \gamma_\nu\mu_\nu. \tag{3.9}$$

2) $\Phi_{\nu+1} : \tilde{D}_{\nu+1}\times\mathcal{O}_{\nu+1} \longrightarrow \tilde{D}_\nu$, *is canonical and real analytic with respect to* $(y,x) \in \tilde{D}_{\nu+1}$, *and analytic with respect to* $y_0 \in \mathcal{O}_{\nu+1}$. *Moreover,*

$$H_{\nu+1} = H_\nu \circ \Phi_{\nu+1},$$

and, on $\tilde{D}_{\nu+1}\times\mathcal{O}_{\nu+1}$,

$$|\Phi_{\nu+1} - id|,\ |D\Phi_{\nu+1} - \mathrm{Id}|,\ |D^i\Phi_{\nu+1}| \le \frac{\mu_*}{2^{\nu+1}},\quad 2 \le i \le m+1. \tag{3.10}$$

3)

$$\mathcal{O}_{\nu+1} = \{y_0 \in \mathcal{O}_\nu : |\langle k, \omega_\nu(y_0)\rangle| > \frac{\gamma_\nu}{|k|^\tau},\ K_\nu < |k| \le K_{\nu+1}\}.$$

Proof: The lemma will be proved by performing the KAM steps inductively. We first verify the conditions H1)-H7) in Section 2 for all $\nu = 0, 1, \cdots$.

Note that

$$\mu_\nu = c_0^\nu \mu_0 s_0^{\frac{\sigma}{b+\sigma}((1+b+\sigma)^\nu - 1)}, \tag{3.11}$$
$$s_\nu = s_0^{(1+b+\sigma)^\nu}, \tag{3.12}$$
$$s_0 = \varepsilon^{\frac{1}{2m}},\ \mu_0 = \frac{1}{4}\varepsilon^{\frac{1}{m+1}},\ \gamma_0 = 4\varepsilon^{\frac{1}{4m}}. \tag{3.13}$$

By (3.13), we see that if ε_0 is small, then

$$s_{\nu+1} \le s_0^{b+\sigma} s_\nu \le \frac{s_\nu}{16},$$

i.e., H1) holds.

To verify H2), we denote

$$E_\nu = \frac{r_\nu - r_{\nu+1}}{8} = \frac{1}{16} r_0(1-\delta)\delta^{\nu+2}.$$

Since $\delta(1+b+\sigma) > 1$, we have

$$\frac{E_\nu}{2}\log\frac{1}{s_\nu} = -\frac{1}{32}r_0(1-\delta)\delta^2(\delta(1+b+\sigma))^\nu \log s_0 \ge -\frac{1}{32}r_0(1-\delta)\delta^2\log s_0 \ge 1,$$

as ε_0 small.

It follows from the above and (3.12),(3.13) that

$$\begin{aligned}
&\log(n+1)! + 3n\log([\log\frac{1}{s_\nu}]+1) - \frac{E_\nu}{2}([\log\frac{1}{s_\nu}]+1)^3 \\
&\quad \le \log(n+1)! + 3n\log(\log\frac{1}{s_\nu}+2) - (\log\frac{1}{s_\nu})^2 \\
&\quad \le -(m+1)(1+b+\sigma)\log\frac{1}{s_\nu},
\end{aligned}$$

as s_ν small, which, by (3.12),(3.13), is ensured by making ε_0 small. Thus,

$$\int_{K_{\nu+1}}^{\infty} \lambda^n e^{-\lambda\frac{E_\nu}{2}} \mathrm{d}\lambda \le (n+1)! K_{\nu+1}^n e^{-K_{\nu+1}\frac{E_\nu}{2}} \le s_{\nu+1}^{m+1},$$

i.e., H2) holds. Similarly, we have

$$\begin{aligned}
&(\nu+1)\log(2c_0) + \log 2 + (m-2)\log s_\nu + \log\mu_\nu \\
&\quad + 3(\tau+1)\log(\log\frac{1}{s_\nu}+2) < 0,
\end{aligned}$$

as ε_0 small. It follows that

$$c_0\gamma_0 s_\nu^{m-2}\mu_\nu K_{\nu+1}^{\tau+1} < \frac{\gamma_0}{2^{\nu+2}} < \gamma_\nu - \gamma_{\nu+1},$$

i.e., H6) holds.

Let

$$\begin{aligned}
l_0 &= \min\{b, \frac{a_0}{2}\}, \\
\eta &= 8+n+2m[\tau]+2m,
\end{aligned}$$

where $[\tau]$ is the integral part of τ. We note that

$$\Gamma_\nu = \Gamma(r_\nu - r_{\nu+1}) \le \int_1^\infty \lambda^{8+n+2m[\tau]+2m} e^{-\lambda E_\nu}\mathrm{d}\lambda \le \frac{\eta!}{E_\nu^\eta}. \tag{3.14}$$

Since by (3.11),

$$\frac{\mu_\nu^{l_0}}{E_\nu^{4\eta}} = \left(\frac{16}{r_0(1-\delta)\delta^2}\right)^{4\eta} \mu_0^{l_0} c_0^{l_0\nu} \frac{s_0^{\frac{l_0\sigma}{b+\sigma}((1+b+\sigma)^\nu - 1)}}{\delta^{(4\eta)\nu}} \le c_*\mu_0^{l_0}\left(\frac{c_0^{l_0} s_0^{l_0\sigma}}{\delta^{4\eta}}\right)^\nu \le c_*\mu_0^{l_0}, \tag{3.15}$$

where $c_* = \left(\frac{16}{r_0(1-\delta)\delta^2}\right)^{4\eta}$, we have by (3.14) that

$$\frac{c_0\mu_\nu\Gamma_\nu}{E_\nu} \le c_0\eta!\frac{\mu_\nu}{E_\nu^{\eta+1}} \le c_0\eta!\frac{\mu_\nu^{l_0}}{E_\nu^{3\eta}} \le c_0 c_*\eta!\mu_0^{l_0} \le 1,$$

as ε_0 small. This verifies H3).

Similarly,

$$
\begin{aligned}
\frac{c_0 s_\nu \mu_\nu \Gamma_\nu}{s_{\nu+1}} &= \frac{c_0 \mu_\nu \Gamma_\nu}{s_\nu^{b+\sigma}} \le \frac{c_0 \mu_\nu}{s_\nu^{b+\sigma}} \frac{\eta!}{E_\nu^\eta} \le c_0 \eta! \left(\frac{16}{r_0(1-\delta)\delta^2} \right)^\eta \frac{\mu_0}{s_0^{\frac{\sigma}{b+\sigma}}} \frac{s_0^{(1+b+\sigma)^\nu (\frac{\sigma}{b+\sigma} - (b+\sigma))}}{\delta^{\eta\nu}} \\
&\le c_0 \eta! \left(\frac{16}{r_0(1-\delta)\delta^2} \right)^\eta s_0^{(3-\frac{\sigma}{b+\sigma})} \left(\frac{c_0 s_0^{2(b+\sigma)a_0}}{\delta^\eta} \right)^\nu \\
&\le c_0 \eta! \left(\frac{16}{r_0(1-\delta)\delta^2} \right)^\eta s_0^{(3-\frac{\sigma}{b+\sigma})} \le \frac{1}{4}, \qquad (3.16) \\
c_0 \mu_\nu^{l_0} \Gamma_\nu^4 &\le c_0 (\eta!)^4 \frac{\mu_\nu^{l_0}}{E_\nu^{4\eta}} \le c_0 c_* (\eta!)^4 \mu_0^{l_0} \left(\frac{c_0^{l_0} s_0^{l_0 \sigma}}{\delta^{4\eta}} \right)^\nu \le c_0 c_* (\eta!)^4 \mu_0^{l_0} \le \frac{1}{4}, \qquad (3.17) \\
c_0 s_\nu^{l_0} \Gamma_\nu^4 &\le c_0 (\eta!)^4 \frac{s_\nu^{l_0}}{E_\nu^{4\eta}} \le c_0 c_* (\eta!)^4 \frac{s_0^{(1+b+\sigma)^\nu}}{\delta^{4\eta\nu}} \le c_0 c_* (\eta!)^4 s_0 \left(\frac{s_0^{b+\sigma}}{\delta^{4\eta}} \right)^\nu \\
&\le c_0 c_* (\eta!)^4 s_0 \le \frac{1}{16}, \qquad (3.18)
\end{aligned}
$$

as ε_0 small. Now, (3.16) is just H4), and, (3.17) together with (3.18) verifies H5). Moreover, by making ε_0 small, we have by (3.17) that

$$
c_0 \mu_\nu^{a_0} \Gamma_\nu^3 \le \frac{1}{2^\nu}. \qquad (3.19)
$$

Next, for each $\nu \ge 1$, by (2.2) and (3.13), we have

$$
\begin{aligned}
\frac{c_0 \Delta_{\nu+1}}{s_{\nu+1}^m \mu_{\nu+1} \gamma_{\nu+1}} &\le 4c_0 (s_\nu^{1+b} + s_\nu^b + s_\nu^{2m-m(1+b+\sigma)-\sigma} \frac{\mu_{\nu+1}}{\gamma_0} + s_\nu^{m+2-m(1+b+\sigma)-\sigma} \frac{1}{\gamma_0}) \Gamma_\nu^4 \\
&\le 4c_0 (2 s_\nu^b + 2 \frac{s_\nu^{2-m(b+\sigma)-\sigma}}{\gamma_0}) \Gamma_\nu^4 \le 8c_0 (s_\nu^b + \frac{s_\nu^{\frac{3}{2}}}{\gamma_0}) \Gamma_\nu^4 \\
&\le 8c_0 (s_\nu^b + s_\nu) \Gamma_\nu^4 \le 16 c_0 s_\nu^{l_0} \Gamma_\nu^4.
\end{aligned}
$$

Hence, by (3.18), H7) is satisfied.

We now proceed the induction. First, we see immediately from Lemmas 2.1-2.4, 2.6 and (3.19) that parts 1) 2) of the present lemma is true for $\nu = 1$. Now assume that for some ν the 1) 2) holds for all $i = 1, 2, \cdots, \nu$. Then, by Lemmas 2.2-2.4, 2.6 and (3.19), we see that the KAM step in Section 2 is valid for $i = \nu + 1$. In particular, for $i = \nu + 1$, all formulas (3.1)-(3.10) hold. This proves parts 1) 2) of the Lemma.

Part 3) clearly holds for $\nu = 0$. We now assume that $\nu > 0$. Then by Lemma 2.5,

$$
\mathcal{O}_\nu = \{ y_0 \in \mathcal{O}_\nu : |\langle k, \omega_\nu(y_0) \rangle| > \frac{\gamma_\nu}{|k|^\tau}, \ 0 < |k| \le K_\nu \}.
$$

Hence

$$
\mathcal{O}_{\nu+1} = \{ y_0 \in \mathcal{O}_\nu : |\langle k, \omega_\nu(y_0) \rangle| > \frac{\gamma_\nu}{|k|^\tau}, \ 0 < |k| \le K_{\nu+1} \}
$$

$$
\begin{aligned}
&= \{y_0 \in \mathcal{O}_\nu : |\langle k, \omega_\nu(y_0)\rangle| > \frac{\gamma_\nu}{|k|^\tau},\ 0 < |k| \le K_\nu\} \\
&\qquad \cap \{y_0 \in \mathcal{O}_\nu : |\langle k, \omega_\nu(y_0)\rangle| > \frac{\gamma_\nu}{|k|^\tau},\ K_\nu < |k| \le K_{\nu+1}\} \\
&= \mathcal{O}_\nu \cap \{y_0 \in \mathcal{O}_\nu : |\langle k, \omega_\nu(y_0)\rangle| > \frac{\gamma_\nu}{|k|^\tau},\ K_\nu < |k| \le K_{\nu+1}\} \\
&= \{y_0 \in \mathcal{O}_\nu : |\langle k, \omega_\nu(y_0)\rangle| > \frac{\gamma_\nu}{|k|^\tau},\ K_\nu < |k| \le K_{\nu+1}\}.
\end{aligned}
$$

The lemma is now complete. ■

4 Proof of main results

4.1 Proof of Parts 1) 2) of Theorem 1

We first show the convergence.

Let $\mu_* = \mu_*(\varepsilon_0)$ be sufficiently small. Then Lemma 3.1 yields the following sequences:

$$
\begin{aligned}
&D_{\nu+1} \times \mathcal{O}_{\nu+1} \subset D_\nu \times \mathcal{O}_\nu, \\
&\Psi^\nu = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_\nu : D_{\nu+1} \times \mathcal{O}^*_{\nu+1} \to D_0, \\
&H \circ \Psi^\nu = H_\nu = N_\nu + P_\nu, \\
&N_\nu = e_\nu + \langle \Omega_\nu, y\rangle + h_\nu(y),
\end{aligned}
$$

$\nu = 0, 1, \cdots$, which satisfy all properties described in the lemma.

Let

$$
\mathcal{O}_* = \bigcap_{\nu=0}^{\infty} \mathcal{O}_\nu, \quad G_* = D(\frac{\beta_0}{2}, \frac{r_0}{2}) \times \mathcal{O}_*, \quad G^* = D(\frac{\beta_0}{2}) \times \mathcal{O}_*.
$$

Then $\mathcal{O}_*$ is a Cantor set. By Lemma 3.1 1), it is clear that e_ν, Ω_ν converge uniformly on $\mathcal{O}_*$, say, to e_∞, Ω_∞, respectively, and, h_ν converges uniformly on G^*, say, to h_∞. Hence N_ν converges uniformly on G_* to

$$
N_\infty = e_\infty + \langle \Omega_\infty, y\rangle + h_\infty(y).
$$

We now show the uniform convergence of Ψ^ν on G_*. Note that

$$
\begin{aligned}
\Psi^\nu - \Psi^{\nu-1} &= \Phi_0 \circ \cdots \circ \Phi_\nu - \Phi_0 \circ \cdots \circ \Phi_{\nu-1} \\
&= \int_0^1 D(\Phi_0 \circ \cdots \circ \Phi_{\nu-1})(id + \theta(\Phi_\nu - id))\mathrm{d}\theta(\Phi_\nu - id).
\end{aligned}
$$

By Lemma 3.1 2), we have

$$
|\Phi_\nu - id|_{G_*} \le \frac{\mu_*}{2^\nu},
$$

and,

$$\begin{aligned}
&|D(\Phi_1\circ\cdots\circ\Phi_{\nu-1})(id+\theta(\Phi_\nu-id))|\\
\leq\;& |D\Phi_1(\Phi_2\circ\cdots\circ\Phi_{\nu-1})(id+\theta(\Phi_\nu-id))|\cdots|D\Phi_{\nu-1}(id+\theta(\Phi_\nu-id))|\\
\leq\;& (1+\frac{\mu_*}{2})\cdots(1+\frac{\mu_*}{2^{\nu-1}})\\
\leq\;& e^{\frac{\mu_*}{2}+\cdots+\frac{\mu_*}{2^{\nu-1}}}\leq e^{\mu_*}.
\end{aligned}$$

It follows that

$$|\Psi^\nu-\Psi^{\nu-1}|_{G_*}\leq e^{\mu_*}\frac{\mu_*}{2^\nu},$$

which implies the uniform convergence of Ψ^ν. Let Ψ^∞ be the limit of Ψ^ν. Then,

$$\begin{aligned}
|\Psi^\infty-id|_{G_*} &\leq |\Phi_0-id|_{G_*}+\sum_{\nu=1}^{\infty}|\Psi^\nu-\Psi^{\nu-1}|_{G_*}\\
&\leq 2\mu_*.
\end{aligned}$$

Thus, Ψ^∞ is uniformly close to the identify and is real analytic on $D(\frac{\beta_0}{2},\frac{r_0}{2})$. Similarly, one can show that, for $i=1,2,\cdots,m+1$, $D^i\Psi^\nu$ converge uniformly to $D^i\Psi^\infty$ respectively, on G_*. By a standard argument using the Whitney extension theorem, one can further show that Ψ^∞ is Whitney smooth with respect to $y_0\in\mathcal{O}_*$ (see [5],[18],[25] for details).

Hence, on G_*,

$$P_\nu=H\circ\Psi^\nu-N_\nu,$$

converges uniformly to

$$P_\infty=H\circ\Psi^\infty-N_\infty.$$

Since

$$|P_\nu|_{D_\nu}\leq\gamma_\nu s_\nu^m\mu_\nu,$$

the Cauchy estimate implies that

$$|\partial_y^j P_\nu|_{D(r_{\nu+1},\frac{1}{2}s_\nu)}\leq\frac{4^{2m+1}}{r_\nu-r_{\nu+1}}\gamma\mu_\nu,$$

for all $1\leq|j|\leq m$ and $\nu=1,2,\cdots$. Let $\nu\to\infty$. We have that, on $D(0,\frac{r_0}{2})\times\mathcal{O}_*$,

$$\partial_y^j P_\infty=0,$$

for all $1\leq|j|\leq m$. Thus, for each $y_0\in\mathcal{O}_*$, the generalized Hamiltonian

$$H_\infty=N_\infty+P_\infty$$

or its associated vector field $I(y+y_0)\nabla H_\infty$ admits an analytic, quasi-periodic, invariant n-torus $\mathcal{T}_{y_0}=\{0\}\times T^n$ with the Diophantine frequency $\omega_\infty(y_0)=-B^\top(y_0)\Omega_\infty(y_0)$. Moreover, these invariant n-tori form a Whitney smooth family.

It remains to show the measure estimate. We note that **R)** implies that there is an open set $G^0 \subset G$ with $|G \setminus G^0| = 0$ such that

$$\mathrm{rank}\{\frac{\partial^\alpha \omega_0}{\partial p^\alpha} : |\alpha| \le n-1\} = n \tag{4.1}$$

for all $y \in G^0$. Hence without loss of generality, we assume that (4.1) holds on $G = \mathcal{O}_0$. By (3.5), the Cauchy estimate, and the Whitney extension theorem, ω_ν, $\nu = 0, 1, \cdots$, admit uniform smooth extensions on $\mathcal{O}_0$ such that

$$|\partial_p^\alpha(\omega_\nu(p) - \omega_0(p))| \le c\mu_*,$$

for all $|\alpha| \le m$, $p \in \mathcal{O}_0$, $\nu = 0, 1, \cdots$, where c is a constant independent of ν. It follows that if ε_0 is sufficient small, then

$$\mathrm{rank}\{\frac{\partial^\alpha \omega_\nu}{\partial p^\alpha} : |\alpha| \le n-1\} = n \tag{4.2}$$

on $\mathcal{O}_0$, for all $\nu = 0, 1, \cdots$.

In the case that $n = 1$, $\mathcal{O}_0$ is a closed interval $[d_1, d_2] \subset R^1$. Since the condition **R)** implies that $\omega(y) \ne 0$ on $\mathcal{O}_0$, there exists a $\sigma' > 0$ such that

$$|\omega(y)| \ge \sigma' \ \text{ for all } y \in \mathcal{O}_0.$$

Let ε be small such that

$$\gamma < \min\{\frac{\sigma'}{2}, \frac{d_2 - d_1}{2}\}.$$

Then one can simply take

$$\mathcal{O}_* = [d_1 + \gamma, d_2 - \gamma],$$

from which $|\mathcal{O}_0 \setminus \mathcal{O}_*| = O(\gamma)$ follows.

In the case that $n \ge 2$, the following lemma is vital to the measure estimate.

Lemma 4.1 *Let $\Lambda \subset R^d$, $d > 1$, be a bounded closed region and let $g : \Lambda \to R^d$ be a smooth map such that*

$$\mathrm{rank}\{\frac{\partial^\alpha g}{\partial p^\alpha} : |\alpha| \le d-1\} = d.$$

Then for a fixed $\tau > d(d-1) - 1$

$$\left|\{p \in \Lambda : |\langle g(p), k\rangle| \le \frac{\gamma}{|k|^\tau}\}\right| \le c(\Lambda, p, \tau)\left(\frac{\gamma}{|k|^{\tau+1}}\right)^{\frac{1}{d-1}}, \quad k \in Z^d \setminus \{0\},\ \gamma > 0.$$

Proof: See Theorem B in [32]. We note that the constant c above does not depend on g but rather on a lower bound of the derivatives of g up to order $d-1$. ∎

We divide the estimate of $|\mathcal{O}_0 \setminus \mathcal{O}_*|$ into the following three cases.

Case 1: $l = n$. Let

$$R_k^{\nu+1} = \{p \in \mathcal{O}_\nu : |\langle k, \omega_\nu(p)\rangle| \leq \frac{\gamma_\nu}{|k|^\tau}\},$$
$$\hat{R}_k^{\nu+1} = \{p \in \mathcal{O}_0 : |\langle k, \omega_\nu(p)\rangle| \leq \frac{\gamma_\nu}{|k|^\tau}\},$$

for all $k \in Z^n \setminus \{0\}$ and $\nu = 0, 1, \cdots$. Then by Lemma 3.1 3),

$$\mathcal{O}_{\nu+1} = \mathcal{O}_\nu \setminus \bigcup_{K_\nu < |k| \leq K_{\nu+1}} R_k^{\nu+1},$$

and,

$$\mathcal{O}_0 \setminus \mathcal{O}_* = \bigcup_{\nu=0}^{\infty} \bigcup_{K_\nu < |k| \leq K_{\nu+1}} R_k^{\nu+1}.$$

Since by Lemma 4.1 and (4.2),

$$|R_k^{\nu+1}| \leq |\hat{R}_k^{\nu+1}| \leq c \left(\frac{\gamma}{|k|^{\tau+1}} \right)^{\frac{1}{n-1}},$$

for all $k \in Z^n \setminus \{0\}$ and $\nu = 0, 1, \cdots$, where c is a constant independent of ν, we have

$$\begin{aligned} |\mathcal{O}_0 \setminus \mathcal{O}_*| &\leq \sum_{\nu=0}^{\infty} \sum_{K_\nu < |k| \leq K_{\nu+1}} |R_k^{\nu+1}| \leq c\gamma^{\frac{1}{n-1}} \sum_{\nu=0}^{\infty} \sum_{K_\nu < |k| \leq K_{\nu+1}} \frac{1}{|k|^{\frac{\tau+1}{n-1}}} \\ &= O(\gamma^{\frac{1}{n-1}}) = O(\gamma^{\frac{1}{l_*-1}}), \end{aligned}$$

as desired.

Case 2: $l < n$. Let $\bar{\mathcal{O}}_0 = [1,2]^{n-l}$ and define

$$\tilde{\mathcal{O}}_0 = \mathcal{O}_0 \times \bar{\mathcal{O}}_0,$$
$$\tilde{\mathcal{O}}_* = \mathcal{O}_* \times \bar{\mathcal{O}}_0,$$
$$\tilde{p} = (p, \bar{p})^\top, \quad \bar{p} \in \bar{\mathcal{O}}_0,$$
$$\tilde{\omega}_\nu(\tilde{p}) = \omega_\nu(p), \quad \nu = 0, 1, \cdots, \ \tilde{p} \in \tilde{\mathcal{O}}_0.$$

Then it is clear that

$$\text{rank}\{\frac{\partial^\alpha \tilde{\omega}_\nu}{\partial \tilde{p}^\alpha} : \forall |\alpha| \leq n-1\} = n$$

on $\tilde{\mathcal{O}}_0$ for all $\nu = 0, 1, \cdots$, as μ_0 sufficiently small. Similar to Case 1, we have that

$$|\tilde{\mathcal{O}}_0 \setminus \tilde{\mathcal{O}}_*| = O(\gamma^{\frac{1}{n-1}}).$$

By Fubini's theorem,

$$|\mathcal{O}_0 \setminus \mathcal{O}_*| = O(\gamma^{\frac{1}{n-1}}) = O(\gamma^{\frac{1}{l_*-1}}),$$

as desired.

Case 3: $l > n$. For any $p \in \mathcal{O}_0$, **R)** implies that there exist indexes

$$\alpha^i \in \{\alpha \in Z_+^l : |\alpha| \le n-1\}, \quad i = 0, 1, \cdots, n-1,$$

such that

$$\text{rank}\{\frac{\partial^{\alpha^i}\omega}{\partial p^{\alpha^i}}(p) : i = 0, 1, \cdots, n-1\} = n.$$

Since $\text{rank}\{\frac{\partial \omega}{\partial p}(p)\} \le n$, there are $p_{i_1}, p_{i_2}, \cdots, p_{i_{l-n}}$ such that

$$\frac{\partial \omega}{\partial p_{i_j}}(p) \notin \{\frac{\partial^{\alpha^i}\omega}{\partial p^{\alpha^i}}(p) : i = 0, 1, \cdots, n-1\}, \quad j = 1, 2, \cdots, l-n.$$

Define

$$\begin{aligned}
&\Omega(p) = (p_{i_1}, p_{i_2}, \cdots, p_{i_{l-n}})^\top, \quad p \in \mathcal{O}_0, \\
&\tilde{\omega}_\nu(p) = (\omega_\nu(p), \Omega(p))^\top, \quad \nu = 0, 1, \cdots, \ p \in \mathcal{O}_0, \\
&\tilde{R}_k^{\nu+1} = \{p \in \mathcal{O}_\nu : |\langle k, \tilde{\omega}_\nu(p)\rangle| \le \frac{\gamma_\nu}{|k|^\tau}\}, \quad k \in Z^l \setminus \{0\}, \ \nu = 0, 1, \cdots, \\
&\tilde{\mathcal{O}}_{\nu+1} = \tilde{\mathcal{O}}_\nu \setminus \bigcup_{K_\nu < |k| \le K_{\nu+1}} \tilde{R}_k^{\nu+1}, \quad \nu = 0, 1, \cdots, \\
&\tilde{\mathcal{O}}_* = \bigcap_{\nu \ge 0} \tilde{\mathcal{O}}_\nu.
\end{aligned}$$

Then

$$\text{rank}\{\frac{\partial^{\alpha^i}\tilde{\omega}_\nu}{\partial p^{\alpha^i}}(p) : i = 0, 1, \cdots, n-1; \frac{\partial \tilde{\omega}_\nu}{\partial p_{i_j}}(p) : j = 1, \cdots, l-n\} = l$$

on $\mathcal{O}_0$ for all $\nu = 0, 1, \cdots$. It follows that

$$\text{rank}\{\frac{\partial^\alpha \tilde{\omega}_\nu}{\partial p^\alpha} : |\alpha| \le l-1\} = l$$

on $\mathcal{O}_0$ for all $\nu = 0, 1, \cdots$. Similar to Case 1), we have that

$$|\mathcal{O}_0 \setminus \tilde{\mathcal{O}}_*| = O(\gamma^{\frac{1}{l-1}}).$$

Since $\tilde{\mathcal{O}}_* \subset \mathcal{O}_*$,

$$|\mathcal{O}_0 \setminus \mathcal{O}_*| \le |\mathcal{O}_0 \setminus \tilde{\mathcal{O}}_*| = O(\gamma^{\frac{1}{l-1}}) = O(\gamma^{\frac{1}{l_*-1}}),$$

as desired.

Above all,

$$|\mathcal{O}_0 \setminus \mathcal{O}_*| = O(\gamma^{\frac{1}{l_*-1}}) = O(\varepsilon^{\frac{1}{4(2n+3)(l_*-1)}}) \to 0, \ \text{as } \varepsilon \to 0.$$

In particular, each $\mathcal{O}_\nu$, $\nu = 0, 1, \cdots$, is nonempty and hence each KAM step can be continued. Let $G_\varepsilon = \mathcal{O}_*$ and recall that $G = \mathcal{O}_0$. The proof of parts 1) 2) of Theorem 1 is now completed.

4.2 Outline of proof of part 3) of Theorem 1

Let $y_0 \in G$ be such that $\omega = \omega(y_0)$ is Diophantine associated to a fixed $\tau > n-1$.

Part 3) of Theorem 1 can be proved by introducing a translation

$$\phi: \ x \to x, \quad y \to y + y^*$$

at each KAM step after the canonical transformation ϕ_F^1. To determine y^* at a KAM step, one need to split the term h in N as

$$h(y) = \frac{1}{2}\langle y, A(y_0)y\rangle + \hat{h}(y),$$

where $\hat{h} = O(|y|^3)$, and, by induction hypothesis, $A(y_0)$ is non-singular on $\mathcal{O}$. We then use the implicit function theorem to choose y^* as the unique solution of the equation

$$A(y_0)y + \partial_y \hat{h}(y) = -p_{01}. \tag{4.3}$$

With the composite transformation

$$\Phi_+ = \phi_F^1 \circ \phi$$

the new Hamiltonian in Section 2 becomes

$$\begin{aligned} H \circ \Phi_+ &= N_+ + P_+, \\ N_+ &= e_+ + \langle \Omega_+, y\rangle + \frac{1}{2}\langle y, A_+ y\rangle + \hat{h}_+(y), \end{aligned}$$

where

$$\begin{aligned} \Omega_+ &= \Omega, \\ e_+ &= e + \langle \Omega, y^*\rangle + \frac{1}{2}\langle y^*, Ay^*\rangle + \hat{h}(y^*) + [R](y^*), \\ A_+ &= A + \partial_y^2 \hat{h}(y^*) + \partial_y^2 [R](y^*), \\ \hat{h}_+(y) &= \hat{h}(y+y^*) - \hat{h}(y^*) - \langle \partial_y \hat{h}(y^*), y\rangle - \frac{1}{2}\langle y, \partial_y^2 \hat{h}(y^*)y\rangle \\ &\quad + [R](y+y^*) - [R](y^*) - \langle \partial_y [R](y^*), y\rangle - \frac{1}{2}\langle y, \partial_y^2 [R](y^*)y\rangle, \end{aligned} \tag{4.4}$$

and, P_+ is defined accordingly.

Note by (4.3) that

$$|y^*|_{\mathcal{O}} \le c\gamma s^{m-2}\mu.$$

This gives an estimate of ϕ at each KAM step. Let $\tau > n-1$ be fixed at the begining of Section 2. The remaining of the proof can be carried out similarly by incorporating all translations ϕ and their estimates into Lemma 3.1 and showing the convergence of the composed transformations. We remark that since the toral frequency ω is kept unchanged in all KAM steps, no measure estimate is needed.

4.3 Proof of Theorem 2

By replacing $y_0 \in G$ with $(y_0, \xi_0) \in G \times \Xi =: \mathcal{O}_0$, we see that all arguments in the previous sections for the Hamiltonian (1.2) go through for the parameterized Hamiltonian (1.6). With the hypothesis **R1)**, the desired measure estimate for Theorem 2 can be carried out similarly to that for Theorem 1 above.

5 Applications

5.1 Perturbation of the three dimensional steady Euler fluid path flows

As shown by Mezić and Wiggins in [19], for a three dimensional, inviscid, incompressible, steady fluid flow, the fluid particle paths, under suitable coordinate, can be described by a three dimensional volume preserving flow generated by a divergence free system of ODEs of the following form:

$$\begin{cases} \dot{z}_1 = \dfrac{\partial H(z_1, z_2)}{\partial z_2} \\ \dot{z}_2 = -\dfrac{\partial H(z_1, z_2)}{\partial z_1} \\ \dot{z}_3 = h(z_1, z_2). \end{cases} \tag{5.1}$$

The right hand side of (5.1) describes the velocity field of the Euler flow (under the present coordinate). We assume that the steady Euler flow admits a family of elliptic vortex lines, i.e.,

H) there is a region $\mathcal{D}$ of the (z_1, z_2)-plane in which the level sets $H(z_1, z_2) = c$ are closed curves.

The assumption is generally satisfied for steady Euler flow. Following from a fundamental result of Arnold ([2]) on three dimensional volume preserving flows, if the steady Euler velocity field is not everywhere collinear with its vorticity field in a domain, then (5.1) admits either invariant tori with trajectories all closed or all dense, or invariant annuli with trajectories all closed.

By assumption **H)**, one can further reduce $(z_1, z_2) \in \mathcal{D}$ into the action-angle variables $(\mathcal{I}, \theta)$ in the usual way, with respect to which the system (5.1) has the form

$$\begin{cases} \dot{\mathcal{I}} = 0 \\ \dot{\theta} = \omega_1(\mathcal{I}) \\ \dot{z}_3 = h(\mathcal{I}, \theta). \end{cases} \tag{5.2}$$

By [19], if $\omega_1(\mathcal{I}) \neq 0$ in $\mathcal{D}$, 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 (5.2) to the system

$$\begin{cases} \dot{\mathcal{I}} = 0 \\ \dot{\theta} = \omega_1(\mathcal{I}) \\ \dot{\phi} = \omega_2(\mathcal{I}), \end{cases} \tag{5.3}$$

where $\phi \in S^1$ or R^1, and,

$$\omega_2(\mathcal{I}) = \frac{\omega_1(\mathcal{I})}{2\pi} \int_0^{2\pi} \frac{h(\mathcal{I}, \theta)}{\omega_1(\mathcal{I})} \mathrm{d}\theta.$$

In other word, under the assumption **H)**, the particle phase space R^3 of the steady Euler flow is foliated into either two dimensional tori or cylinders carrying either action-angle-angle or action-action-angle variables, respectively.

One important approach to understand the barriers to fluid transport and mixing is to study the persistence of the invariant 2-tori or 1-tori (on the cylinder) of (5.3) after suitable perturbations. This brings a KAM type of theory into play.

The persistence of invariant 2-tori of (5.3) under volume preserving perturbations was shown in [19] by using the KAM theory developed in [7] for volume preserving maps. We now consider a similar persistence problem under the generalized Hamiltonian framework. By considering certain non-volume-preserving perturbations to (5.3), it is our hope that the KAM type of results presented in this paper can be of help for a general understanding of the existence of barriers to fluid transport and mixing arising in viscid fluids.

To apply our results, we note that (5.3) can be easily put into the generalized Hamiltonian framework under a variety choice of Poisson structures ω^2 (or the associated structure matrices $I(\mathcal{I})$) of the phase manifold. Taking the action-angle-angle case for example, it is immediately seen that the right hand side of (5.3) can be written as

$$I(\mathcal{I})\nabla N(\mathcal{I}) = \begin{pmatrix} 0 & B^\top(\mathcal{I}) \\ -B(\mathcal{I}) & C \end{pmatrix} \begin{pmatrix} N'(\mathcal{I}) \\ 0 \\ 0 \end{pmatrix}, \tag{5.4}$$

where C is an arbitrary 2×2 skew-symmetric constant matrix, and, $B(\mathcal{I})$ and $N(\mathcal{I})$ are such that

$$\begin{pmatrix} \omega_1(\mathcal{I}) \\ \omega_2(\mathcal{I}) \end{pmatrix} + B(\mathcal{I})N'(\mathcal{I}) = 0.$$

Since C is a constant matrix, $I(\mathcal{I})$ clearly satisfies the Jacobi identity, and hence defines a structure matrix. To apply our results to such a system, the perturbation can be any sufficient smooth vector field of the form $\varepsilon I(\mathcal{I})\nabla P(\mathcal{I},\theta,\phi)$ (which certainly need not be divergence free). In applications, one does have the freedom to determine the Poisson matrix $I(\mathcal{I})$ according to either the form of a perturbation or the nature of a particular problem (e.g., certain fluid flows are restricted to preserve a given Poisson structure).

In the following, we give two examples of perturbed three dimensional volume preserving flows to illustrate the application of our results.

Example 5.1 (Action-angle-angle). Consider the following generalized Hamiltonian system:

$$\begin{pmatrix} \dot{r} \\ \dot{\theta} \\ \dot{\varphi} \end{pmatrix} = I\nabla(\frac{1}{2}r^2 + \varepsilon P(r,\theta,\varphi)), \tag{5.5}$$

where $r \in R^1_+$, $(\theta,\varphi) \in T^2$, $\varepsilon > 0$ is a small parameter, P is real analytic, and I is a constant structure matrix. Thus I must have the form

$$I = \begin{pmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & -\gamma \\ -\beta & \gamma & 0 \end{pmatrix},$$

where α,β,γ are arbitrary real numbers with $|\alpha|+|\beta|+|\gamma| \neq 0$. It is clear that

$$\omega = (\omega_1,\omega_2)^\top = -(\alpha r, \beta r)^\top.$$

Since $\text{rank}\frac{\partial\omega}{\partial r} \leq 1$, we see that the non-degenerate condition **R)** is not satisfied and hence Theorem 1 is not applicable. However, if $r \neq 0$, then

$$\text{rank}\frac{\partial\omega}{\partial(r,\alpha)} = \text{rank}\frac{\partial\omega}{\partial(r,\beta)} = \text{rank}\frac{\partial\omega}{\partial(\alpha,\beta)} \equiv 2.$$

We can then apply Theorem 2 to conclude the following.

Proposition 5.1 *Consider* (5.5) *and let* $0 < \delta < r_0$ *be arbitrarily given,* A, B *be any two compact intervals. Then the following holds.*

1) *Fix* $r_0, \gamma \neq 0$. *There is a family of Cantor sets* $G_\varepsilon \subset G = A \times B$ *with* $|G \setminus G_\varepsilon| \to 0$ *as* $\varepsilon \to 0$ *such that all unperturbed 2-tori* $T_{r_0,\alpha,\beta}$ *associated to* $(\alpha,\beta) \in G_\varepsilon$ *will persist as* ε *small;*

2) *Fix* β, γ *with* $|\beta| + |\gamma| \neq 0$. *There is a family of Cantor sets* $G_\varepsilon \subset G = [\delta, r_0] \times A$ *with* $|G \setminus G_\varepsilon| \to 0$ *as* $\varepsilon \to 0$ *such that all unperturbed* 2*-tori* $T_{r,\alpha}$ *associated to* $(r, \alpha) \in G_\varepsilon$ *will persist as* ε *small;*

3) *Fix* α, γ *with* $|\alpha| + |\gamma| \neq 0$. *There is a family of Cantor sets* $G_\varepsilon \subset G = [\delta, r_0] \times B$ *with* $|G \setminus G_\varepsilon| \to 0$ *as* $\varepsilon \to 0$ *such that all unperturbed* 2*-tori* $T_{r,\beta}$ *associated to* $(r, \beta) \in G_\varepsilon$ *will persist as* ε *small.*

Example 5.2 (Action-action-angle). Consider the following generalized Hamiltonian system:

$$\begin{pmatrix} \dot{r} \\ \dot{y} \\ \dot{\theta} \end{pmatrix} = I\nabla(\frac{1}{2}(r^2 + y^2) + \varepsilon P(r, y, \theta)), \tag{5.6}$$

where $r \in R^1_+$, $y \in R^1$, $\theta \in T^1$, $\varepsilon > 0$ is a small parameter, P is real analytic, and I is a constant structure matrix . Thus, I must have the form

$$I = \begin{pmatrix} 0 & 0 & \beta \\ 0 & 0 & -\gamma \\ -\beta & \gamma & 0 \end{pmatrix},$$

where β, γ are arbitrary real numbers with $|\beta| + |\gamma| \neq 0$. Clearly,

$$\omega = -\beta r + \gamma y, \quad \text{and} \quad \text{rank}\frac{\partial \omega}{\partial(r, y)} \equiv 1.$$

Hence, Theorem 1 is immediately applicable to yield the following.

Proposition 5.2 *Consider* (5.6) *and let* $A \subset R^1_+, B \subset R^1$ *be compact intervals. Then there is a family of Cantor sets* $G_\varepsilon \subset G = A \times B$ *with* $|G \setminus G_\varepsilon| \to 0$ *as* $\varepsilon \to 0$ *such that all unperturbed* 1*-tori* $T_{r,y}$ *associated to* $(r, y) \in G_\varepsilon$ *will persist as* ε *small.*

5.2 Invalidity of closing lemma on a Poisson manifold

In [26], Pugh and Robinson proved the following C^2 closing lemma: on a symplectic manifold (M, ω^2) the set $\{H \in C^2$: the periodic orbits of the flow ϕ^t_H generated by the Hamiltonian vector field of H are dense in $M\}$ is dense in $C^2(M, R)$ under the C^2-topology. In [16], Herman gave a counterexample to the closing lemma on (T^{2n+2}, ω^2) for the C^k-topology $(k > 2n + 1)$.

Considering generalized Hamiltonian systems, a natural question is that whether or to what extend a C^k closing lemma can hold on a Poisson manifold. We do not have answer to this general question but can give a Herman type of counterexample on a Poisson manifold for sufficiently large k, based on Theorem 2.

Example 5.3 (Invalidity of closing lemma). Consider the Poisson manifold $([1,2]\times T^n, \omega_\xi^2)$, where $n \geq 2$ and ω_ξ^2 is defined by the structure matrix

$$I = \begin{pmatrix} 0 & -\xi^\top \\ \xi & J \end{pmatrix}$$

which depends on a parameter $\xi \in [1,2]^n = \Xi$ to be determined later.

Consider the generalized Hamiltonian

$$N = \frac{1}{2}y^2$$

for $y \in [1,2]$. It is clear that

$$I(y,\xi)\nabla N(y) = (0, y\xi)^\top,$$

hence $\omega = y\xi$.

Fix a $y_0 \in [1,2]$ and let ξ be such that

$$|\langle k, y_0\xi\rangle| > \frac{\gamma}{|k|^\tau}, \quad k \in Z^n \setminus \{0\}, \tag{5.7}$$

where $0 < \gamma \ll 1$ and $\tau > n-1$ is fixed. Since $\text{rank}\frac{\partial\omega}{\partial\xi} = n$, an application of the smooth version of Theorem 2 3) (see Remark 1.1) yields that there is an integer k sufficiently large and $\varepsilon > 0$ sufficiently small such that each Diophantine torus $T_{y,\xi}$ of the vector field $I(y,\xi)\nabla N(y)$ will persist and gives rise to a slightly deformed invariant torus of the perturbed vector field $I(y,\xi)\nabla(N(y) + P(y,x))$ for $|P|_{C^k([1,2]\times T^n)} \leq \varepsilon$. Now let ξ_0 be a fixed vector satisfying (5.7) (the Diophantine constant γ in (5.7) is determined by the smallness of ε). Then on the Poisson manifold $([1,2]\times T^n, \omega_{\xi_0}^2)$ any generalized Hamiltonian in the ε neighborhood of $N(y)$ under the C^k topology admits a quasi-periodic, invariant n-torus perturbed from the torus T_{y_0,ξ_0}. Hence C^k closing lemma is invalid on the Poisson manifold for sufficiently large k.

Note that in the above construction we have crucially used the fact of preservation of toral frequency, for otherwise the unperturbed torus cannot be chosen in *a priori*.

References

[1] V. I. Arnold, Proof of A. N. Kolmogorov's theorem on the preservation of quasi-periodic motions under small perturbations of the Hamiltonian, *Russ. Math. Surv.* **18** (1963), 9-36

[2] V. I. Arnold, Sur la topologie des ecoulements stationaries des fluides parfaits, *C. R. Acad. Sci. Paris* **261** (1965), 17-20

[3] H. Broer, G. Huitema and F. Takens, Unfoldings of quasi-periodic tori, *Mem. Amer. Math. Soc.* **83** (1990), 13-42

[4] H. Broer, G. Huitema and M. B. Sevryuk, Families of quasi-periodic motions in dynamical systems depending on parameters, Nonlinear Dynamical Systems and Chaos (*Proc. dyn. syst. conf.,* H. W. Broer et. al., eds.), Birkhäuser, Basel, 1996, 171-211

[5] H. Broer, G. Huitema and M. Sevryuk, Quasi-periodic motions in families of dynamical systems, *Lect. Notes Math.* **1645**, Springer-Verlag, 1996

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

[7] C. Q. Cheng and Y. S. Sun, Existence of invariant tori in three dimensional measure-preserving mappings, *Celestial Mech. Dyn. Astronom.* **47** (1990), 275-292

[8] C. Q. Cheng and Y. S. Sun, Existence of KAM tori in degenerate Hamiltonian systems, *J. Diff. Eqs.* **114** (1994), 288-335

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

[10] F. Z. Cong, Y. Li and M. Y. Huang, Invariant tori for nearly twist mappings with intersection property, *Northeast. Math. J.* **12** (1996), 280-298

[11] R. de la Llave, Recent progress in classical mechanics, Mathematical Physics, X (Leipzig, 1991), 3-19, Springer, Berlin, 1992

[12] A. Delshams and R. de la Llave, Existence of quasi-periodic orbits and absence of transport for volume-preserving transformations and flows, preprint (1991)

[13] L. H. Eliasson, Perturbations of stable invariant tori for Hamiltonian systems, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. Ser.IV* **15** (1988), 115-147

[14] M. R. Herman, Topological stability of the Hamiltonian and volume-preserving dynamical systems, Lecture at the International Conference on Dynamical Systems, Evanston, Illinois, 1991

[15] M. R. Herman, Différéntiabilité optimae 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

[16] 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

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

[18] Y. Li and Y. Yi, A quasi-periodic Poincaré's theorem, *Math. Ann.*,to appear

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

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

[21] J. Moser, Convergent series expansions for quasi- periodic motions, *Math. Ann.* **169** (1967), 136 - 176

[22] 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

[23] P. J. Olver, Applications of Lie groups to differential equations, Springer-Verlag, New York, 1986

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

[25] J. Pöschel, Integrability of Hamiltonian systems on Cantor sets, *Comm. Pure Appl. Math.* **35** (1982), 653-696

[26] C. C. Pugh and C. Robinson, The C^1 closing lemma including Hamiltonians, *Erg. Th. Dyn. Syst.* **3**(1983), 261-314

[27] H. Rüssmann, On twist Hamiltonians, Talk on the Colloque International: Mécanique céleste et systemes hamiltoniens, Marseille, 1990

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

[29] M. B. Sevryuk, Some problems of the KAM-theory: conditionally periodic motions in typical systems, *Russian Math. Surveys* **50** (1995), 341-353

[30] Z. H. Xia, Existence of invariant tori in volume-preserving diffeomorphisms, *Erg. Th. Dyn. Syst.* **12** (1992), 621-631

[31] Z. H. Xia, Existence of invariant tori for certain non-symplectic diffeomorphisms, Hamiltonian dynamical systems: History, Theory and Applications (IMA Vol. Math. Appl. 63, ed H. S. Dumas, K. R. Meyer and D. S. Schmidt), New York, Springer, 1995, 373-385

[32] 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

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