

Birkhoff–Kolmogorov–Arnold–Moser Tori in Convex Hamiltonian Systems

Chong-Qing Cheng

Department of Mathematics, Nanjing University, Nanjing 210008, People's Republic of China

Received: 1 September 1993 / Accepted: 15 August 1995

Abstract: For the Hamiltonian systems of KAM type, it is proved that some lower dimensional invariant tori always exist in the resonance gaps although those maximum tori can not survive small perturbations in the generic case.

1. Introduction and Results

In this paper we study perturbed integrable Hamiltonian systems with n degrees of freedom and investigate what happens to those n -dimensional invariant tori of the unperturbed completely integrable systems in the zones of instability [A1].

To be more precise, the Hamiltonian system under consideration

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q} \quad (1.1)$$

is determined by a Hamiltonian of KAM type

$$H(p, q) = N(p) + P(p, q) \quad (1.2)$$

which is assumed to be real analytical in $\mathbb{D} \times \mathbb{T}^n$, where $\mathbb{D} \subseteq \mathbb{R}^n$ is an open set, P is a small Hamiltonian perturbation and N is the main part. As usually, $q = (q_1, q_2, \dots, q_n) \in \mathbb{T}^n$ denotes a vector of angular variables and $p = (p_1, p_2, \dots, p_n) \in \mathbb{R}^n$ is a vector of action variables. Clearly, when the perturbation P vanishes, the system (1.1) is integrable and $\mathbb{D} \times \mathbb{T}^n$ is stratified by a family of n -dimensional invariant tori $p = \text{const.}$ carrying a quasi-periodic flow $q = \omega t + q_0$ with torus frequency vector ω given by

$$\omega(p) = (\omega_1, \omega_2, \dots, \omega_n) = \frac{\partial N}{\partial p}. \quad (1.3)$$

If the frequency vector $\omega(p)$ is not too well approximated by rationals, in other words, if it satisfies some Diophantine condition, the famous KAM theory tells us that the corresponding invariant torus $p = p_0$ survives small perturbations with only

slight deformations and still carries a quasi-periodic flow with the same frequencies $\omega(p_o)$ ([K, A2, Mo1]). Suppose $N(p)$ satisfies the Kolmogorov nondegeneracy condition, i.e. the Hessian matrix of N is invertible. It is then clear that there are some frequency vectors in $\{\omega(p)|p \in \mathbb{D}\}$ which do not satisfy Diophantine conditions. In this paper we suppose that there is a $p_o \in \mathbb{D}$ such that $\omega = \omega(p_o)$ satisfies a rational resonance condition, which means

$$\langle I_n, \omega \rangle = \sum_{i=1}^n I_{ni} \omega_i = 0 \tag{1.4}$$

holds for some $I_n \in \mathbb{Z}^n \setminus \{0\}$. Here we assume $\lambda I_n \notin \mathbb{Z}^n \setminus \{0\}$ for any real $|\lambda| < 1$. By noticing that the map $\frac{\partial N}{\partial p}$ from the action variable space to the frequency space is invertible, we know there is a $(n - 1)$ -dimensional manifold containing p_o , where (1.4) is always satisfied. This is the zone of instability. In this case, the torus $p = p_o$ is foliated into a family of $(n - 1)$ -dimensional submanifolds which are determined by

$$\langle q, I_n \rangle = \text{const.}, \tag{1.5}$$

and invariant under quasi-periodic flow $q \rightarrow q + \omega t$ if the perturbation P vanishes. So it is quite natural to ask what happens to such a n -dimensional invariant torus of the unperturbed system when a perturbation appears. It seems that such a n -torus breaks down under perturbations in general. In the case of a monotone twist area-preserving map of the annulus, which is equivalent to a non-degenerate Hamiltonian system with two degrees of freedom, these phenomena were already discovered by Poincaré; a modern rigorous account can be found in [Rob]. Here we would like to investigate whether there is some $(n - 1)$ -dimensional submanifold in the unperturbed system surviving small perturbations and still carrying a quasi-periodic flow, therefore it can be viewed as “trace” or “ghost” of that torus [BK]. These “ghosts” are expected to have a great influence on the dynamical behaviour of Arnold diffusion [Ch, L].

To formulate our results, we need to say some words about the flow on those $(n - 1)$ -dimensional submanifolds. From the theory of abelian groups we know that there is a unimodular matrix (i.e. $\det I = 1$)

$$I = [I_1, I_2, \dots, I_n]$$

in which I_n is specified as above and all $I_j \in \mathbb{Z}^n$ for $1 \leq j \leq n$. We define an injection from \mathbb{T}^{n-1} to \mathbb{T}^n by

$$\mathcal{H}_\xi : (\phi_1, \phi_2, \dots, \phi_{n-1}) \rightarrow q = (\phi_1, \phi_2, \dots, \phi_{n-1}, \xi)I^{-1}, \tag{1.6}$$

where fixed $\xi \in \mathbb{T}$ is treated as a parameter. Such a map induces a diffeomorphism from \mathbb{T}^{n-1} to one of those $(n - 1)$ -dimensional submanifolds satisfying (1.5), depending on ξ . These submanifolds are invariant under the flow $\Phi_t(q) = \omega t + q$, and we have a pull back flow of Φ_t on \mathbb{T}^{n-1} denoted by Φ_t^*

$$\mathcal{H}_\xi \Phi_t^*(0) = \Phi_t(q)$$

which is also quasi-periodic; its frequency vector $\omega^* = (\omega_1^*, \omega_2^*, \dots, \omega_{n-1}^*)$ is determined by

$$(\omega^*, 0) = \omega I. \tag{1.7}$$

To identify which submanifold \mathcal{H}_ξ maps \mathbb{T}^{n-1} onto, we need to specify the parameter ξ . When ξ ranges over the whole \mathbb{T} all the corresponding images of \mathbb{T}^{n-1} under \mathcal{H}_ξ covers \mathbb{T}^n exactly. In this paper, when there is no danger of confusion we call these manifolds \mathbb{T}^{n-1} tori, and ω^* the frequency vector of the quasi-periodic flow on \mathbb{T}^{n-1} .

Theorem. *Let the Hamiltonian be of KAM type $H(p, q) = N(p) + P(p, q)$ and real analytic in a complex neighbourhood*

$$\Sigma_{s, \tau} : |\operatorname{Im} q| \leq \tau, \quad |p - p_o| < s$$

of a n -torus $p = p_o$, where $|x| = \max_{1 \leq j \leq n} |x_j|$, and assume the following conditions:

1. *Convexity: the main part N is a convex function, hence the eigenvalues of its Hessian matrix $\left(\frac{\partial^2 N}{\partial p_i \partial p_j}\right)$ satisfy:*

$$\min\{\lambda_1, \lambda_2, \dots, \lambda_n\} \geq \lambda > 0 \quad \text{in } \operatorname{Re}(\Sigma_{s, \tau}).$$

2. *Resonance: $\langle \omega(p_o), I_n \rangle = 0$.*

3. *Relative Diophantine condition: the frequency ω^* of the unperturbed flow on $(n - 1)$ -dimensional submanifolds satisfies Diophantine inequalities*

$$\left| \sum_{j=1}^{n-1} k_j \omega_j^* \right| \geq D |k|_1^{-\mu}, \quad \forall k \in \mathbb{Z}^{n-1} \setminus \{0\} \tag{1.8}$$

for some $D > 0$ and $\mu \geq n$, here $|k|_1 = \sum_{j=1}^{n-1} |k_j|$.

Then there is a positive $d = d(I_n, N, n, D, \mu, s, \tau)$ such that for $|P| \leq d$ in $\Sigma_{s, \tau}$, the flow induced by (1.1) admits at least one $(n - 1)$ -dimensional submanifold of the form

$$\begin{aligned} p &= p_o + \Gamma(\phi_1, \phi_2, \dots, \phi_{n-1}), \\ q &= \mathcal{H}_{\xi_o}(\phi_1, \phi_2, \dots, \phi_{n-1}) + \Theta(\phi_1, \phi_2, \dots, \phi_{n-1}) \end{aligned} \tag{1.9}$$

with $\Gamma = (\Gamma_1, \Gamma_2, \dots, \Gamma_n)$, $\Theta = (\Theta_1, \Theta_2, \dots, \Theta_n)$ being real analytical functions of period 2π in the complex domain $|\operatorname{Im} \phi| < \frac{\varepsilon}{2}$. The parametrization is chosen so that the induced flow on \mathbb{T}^{n-1} is still given by

$$\phi = \phi_o + \omega^* t.$$

Moreover, for each $\varepsilon > 0$, a positive $d' = d'(\varepsilon, I_n, N, n, D, \mu, s, \tau) < d$ can be found such that if $|P| \leq d'$ in $\Sigma_{s, \tau}$ further, then the functions Γ and Θ satisfy

$$|\Gamma| + |\Theta| < \varepsilon. \tag{1.10}$$

We impose the convexity of the Hamiltonian to prove our theorem, and we do not assume any other conditions on P except smallness and analyticity. It can be seen from the following proof that the surviving \mathbb{T}^{n-1} is of hyperbolic type generically. With regard to Birkhoff’s fixed point theorem, it seems that another \mathbb{T}^{n-1} of elliptic type could survive the perturbation as well, but we failed to prove its existence without additional conditions on P .

Given a fixed I_n , there is an $(n - 1)$ -dimensional plane in the frequency space where the resonance condition (1.4) is satisfied, corresponding to which, there is a $(n - 1)$ -dimensional submanifold in action variable space. Clearly, in such a submanifold almost all points satisfy some relative Diophantine condition. So when perturbation is small enough there is a set with positive $(n - 1)$ -dimensional Lebesgue measure where our result holds. On the other hand, restricted to the case of area-preserving twist mapping such invariant torus corresponds to hyperbolic fixed point predicted by Birkhoff fixed point theorem. Therefore we would like to call such survived submanifolds \mathbb{T}^{n-1} Birkhoff–Kolmogorov–Arnold–Moser tori, or BKAM tori in brief. The theorem shows that in every zone of instability there are some tori surviving small perturbations.

The analyticity of the Hamiltonian in our theorem is clearly not necessary. We require it just for simplicity. The same result can be expected in the C^m case if we exploit Nash–Moser’s technique [Mo1, H1, H2, Rüs1].

The structural stability of lower dimensional invariant manifolds has been a very interesting and important subject, which is related to finding quasi-periodic solutions of nonlinear partial differential equations [W]. For Hamiltonian systems, remarkable progress has been made when lower dimensional tori are of hyperbolic, elliptic or mixed type [G, Mo2, E, P, Rüs2]. These results are not applicable here because the lower dimensional submanifolds considered here are degenerate, i.e. the linearized variational equation for normal coordinates has zero eigenvalue. For example, we assume that a torus \mathbb{T}^n is stratified into a family of trivially embedded tori \mathbb{T}^{n-1} ,

$$(\phi_1, \phi_2, \dots, \phi_{n-1}) \rightarrow (\phi_1 + q_1, \phi_2 + q_2, \dots, \phi_{n-1} + q_{n-1}, q_n)$$

with fixed q as a parameter. In this case we have $\omega_n = 0$ and the normal form of unperturbed equations reads

$$\begin{aligned} \dot{q}_j &= \omega_j, & \dot{p}_j &= 0, & 1 \leq j < n, \\ \dot{z}_n &= 0, & \dot{\bar{z}}_n &= 0, \end{aligned}$$

where z_n is so-called normal coordinate. Consequently we can not predict which specific $(n - 1)$ -torus is not destroyed by the perturbation if we do not assume other conditions on the perturbation except smallness.

It should be mentioned that for a $2n$ -dimensional convex Hamiltonian system of KAM type, the existence of at least n Birkhoff periodic orbits has been proved in [BK] via some variational approach invented by Conley and Zehnder [CZ]. Variational methods have been proved a powerful method for the study of Hamiltonian systems, the existence of periodic orbits, Aubry–Mather set [Au, Ma1] as well as action minimizing invariant measures for positive definite Lagrangians with multi-degree of freedom [Ma2]. However it seems always necessary to deal with some small divisor problems when we manage to prove the existence of manifolds with dimension ≥ 2 .

In Sect. 2 we outline our proof, using a Newtonian iteration scheme to overcome the small divisor problem, using the variational principle to localize surviving BKAM torus, etc. In Sect. 3, we study in detail one step of iteration of symplectic transformation, and in Sect. 4 the convergence of the iteration is proved.

2. Outline of the Proof

Our problem is proved by basically using the KAM method employing a rapidly converging iteration scheme. The specific difficulty of our problem is that we can not prescribe nested domains as is done in the proof of the classical KAM theory, shrinking to the domain $|\operatorname{Im} q| \leq \frac{\tau}{2}$, $p = p_0$, in every intermediate domain one step of symplectic transformation is carried out. We need to bound the domain for any step of the transformation according to a special property of the main part of the Hamiltonian which is due to the degeneracy as we pointed out in the last section.

To simplify the notation, we introduce the linear symplectic transformation $(p, q) \rightarrow (x, y)$ by

$$x = qI, \quad p = yI^t.$$

Here I is the unimodular matrix, I^t denotes its transposition. Substituting (x, y) for (q, p) in the Hamiltonian $H(x, y)$, clearly, H is analytical in the domain

$$\Sigma_{s_*, \tau_*} \quad |\operatorname{Im} x| \leq \tau_*, \quad |y - y_0| \leq s_*,$$

where $p_0 = I^t y_0$, $\tau_* \leq \tau$ and $s_* \leq s$, (τ_*, s_*) depends on (I_n, τ, s) . In this case, we have

$$\frac{\partial N}{\partial y}(y_0) = (\omega_1, \omega_2, \dots, \omega_{n-1}, 0) = (\omega, 0).$$

N is also a convex Hamiltonian in y . In the following, we use $(\omega, 0)$ to denote $(\omega^*, 0)$ for the abbreviation of notation.

Formally, for any step of iteration, a symplectic coordinate transformation $\mathcal{M} : (x, y) \rightarrow (x_+, y_+)$ is introduced taking the form:

$$y = y_+ + W_x, \quad x_+ = x + W_{y_+}$$

with a generating function $W(x, y_+)$. As the first step of that iteration W can be determined by looking for the solution of the homological equation as was done in the proof of the classical KAM theory [A2]:

$$\langle (\omega, 0), W_x \rangle = -P(x, y_+) + N_+(x_n, y_+),$$

where $N_+(x_n, y_+)$ is chosen in this way: we expand $P(x, y_+)$ into Fourier series in $(x_1, x_2, \dots, x_{n-1}) = \hat{x}$,

$$P(x, y_+) = \sum_{k \in \mathbb{Z}^{n-1}} P_k(x_n, y_+) e^{i\langle k, \hat{x} \rangle},$$

and let $N_+(x_n, y_+) = -P_0(x_n, y_+)$. If we also expand W into Fourier series in \hat{x} in the same way then

$$\langle (\omega, 0), W_x \rangle = \sum_{k \in \mathbb{Z}^{n-1} \setminus \{0\}} \langle k, \omega \rangle W_k(x_n, y_+) e^{i\langle k, \hat{x} \rangle},$$

so we have

$$W(x, y_+) = \sum_{k \in \mathbb{Z}^{n-1} \setminus \{0\}} \frac{iP_k(x_n, y_+)}{\langle k, \omega \rangle} e^{i\langle k, \hat{x} \rangle}.$$

If $|P(x, y_+)| \leq d$ in $\Sigma_{s, \tau}$ Cauchy's estimate leads to

$$|W(x, y_+)| \leq \frac{c_1 d}{D(\tau - \tau_1)^\mu}$$

in $\{|\operatorname{Im} \hat{x}| \leq \tau_1 < \tau\}$, c_1 is a positive number only depending on n and μ . Throughout this paper we use c_i to denote some positive number depending only on n and μ if not specified otherwise.

After the change of variables, the Hamiltonian takes the form

$$H(x_+, y_+) = (N + N_+)(x_{n+}, y_+) + P_+(x_+, y_+).$$

The new perturbation term P_+ is much smaller than the original one if d is small enough, and the main part $N + N_+$ depends not only on action variables but also on one angular variable, so the n -tori near y_o are no longer invariant tori of (1.1) when $H = N + N_+$, which implies the iteration scheme on the whole \mathbb{T}^n would fail. Nevertheless we find that there are some $(n - 1)$ dimensional tori embedded in some n -torus near p_o which are invariant under the flow governed by (1.1) when $H = N + N_+$:

Lemma 1. *Suppose $N(y)$ is nonsingular on \mathbb{D} , i.e. the eigenvalues $\lambda_i(y)$ of its Hessian matrix are lower bounded in the sense of modulus: $|\lambda_i(y)| \geq \lambda > 0$, and $\frac{\partial N}{\partial y}(y_o) = (\omega_1, \omega_2, \dots, \omega_k, 0, \dots, 0) =: \Omega$. $N_+ = N_+(x_{k+1}, \dots, x_n, y)$, defined on $\mathbb{T}^{n-k} \times \mathbb{D}$, satisfies some smallness condition: $|N_{+y}| < \lambda n^{-1} \operatorname{dist}\{y_o, \partial\mathbb{D}\}$ and the spectrum of its Hessian matrix lie in the ball with radius smaller than λ , centered at the origin. Then Eq. (1.1) with Hamiltonian $N + N_+$ admits at least $(n - k + 1)$ invariant tori \mathbb{T}^k ,*

$$x_j = \omega_j t + x_{jl}, \quad (1 \leq j \leq k)$$

$$x_j = x_{jl}, \quad (k < j \leq n)$$

$$y = y_l,$$

with $1 \leq l \leq (n - k + 1)$, $|y_l - y_o| \leq \frac{\eta}{\lambda} \max_{\mathbb{D}} |N_{+y}|$.

Proof. The function $F = N(y) + N_+(x_{k+1}, \dots, x_n, y) - \langle \Omega, y \rangle$ is invariant under the flow $x(t) = x + \Omega t$, $y(t) = y$, for F is independent of (x_1, x_2, \dots, x_k) . If we find some critical points of F with respect to (x_{k+1}, \dots, x_n) and y , which are also invariant under such flow, we have

$$\frac{\partial}{\partial y}(N + N_+) = \Omega, \quad \frac{\partial}{\partial x}(N + N_+) = 0,$$

which implies that starting from these points, the flow governed by (1.1) keeps y and (x_{k+1}, \dots, x_n) fixed and $x_j = \omega_j t + x_j(0)$ ($j \leq k$). Thus we only need to find critical points of F .

For any fixed $(x_{k+1}, \dots, x_n) \in \mathbb{T}^{n-k}$ we consider the map $N_y + \varepsilon N_{+y} : \mathcal{B}_r(y_o) \rightarrow \mathbb{R}^n$, where $0 \leq \varepsilon \leq 1$ and $\mathcal{B}_r(y_o)$ is a ball of radius $r > \frac{\eta}{\lambda} \max_{\mathbb{D}} |N_{+y}|$, centered at y_o . By the condition assumed there is such a ball contained in \mathbb{D} . When $\varepsilon = 0$ the degree of the map

$$\operatorname{deg}(N_y + \varepsilon N_{+y}, \mathcal{B}_r(y_o), \Omega) = \pm 1,$$

and for $0 \leq \varepsilon \leq 1$,

$$\operatorname{dist}\{(N_y + \varepsilon N_{+y})(\partial\mathcal{B}_r), \Omega\} \geq \min_{(\partial\mathcal{B}_r)} |N_y(\partial\mathcal{B}_r) - N_y(y_o)| - \varepsilon n \max_{\mathbb{D}} |N_{+y}| > 0$$

which implies that the equation

$$\frac{\partial}{\partial y}(N + N_+)(x_{k+1}, \dots, x_n, y) = \Omega$$

has solution for any $(x_{k+1}, \dots, x_n) \in \mathbb{T}^{n-k}$. From the assumed conditions we see that $(N + N_+)_{y^2}$ is nonsingular, so we obtain a differentiable function $y = y(x_{k+1}, \dots, x_n)$ by the implicit function theorem. The function $F(x_{k+1}, \dots, x_n, y(x_{k+1}, \dots, x_n))$ is defined on the torus \mathbb{T}^{n-k} , it then has at least $(n - k + 1)$ critical points by Lusternik–Schnirelman theory.

In this paper we are interested in the critical point of $F(x_n, y(x_n))$, where F reaches its maximum. Such a critical point corresponds to a hyperbolic $(n - 1)$ -dimensional torus in general since we have assumed the convexity of $N(y)$. The existence of such a \mathbb{T}^{n-1} tori, invariant under Hamiltonian flow given by $N(y) + N_+(x_n, y)$ suggests the possibility that the iteration scheme for the symplectic transformation may converge on these lower dimensional BKAM tori instead of on the whole torus \mathbb{T}^n .

In the following we still use (p, q) to denote the action and angular variables.

To proceed to the next step of iteration we have to examine whether or not the critical point has “stronger persistency” for perturbations from an average part of P_+ with respect to $\hat{q} = (q_1, q_2, \dots, q_{n-1})$. Here “stronger persistency” is such a condition which guarantees that the corresponding critical point of $N + N_+$ is sufficiently close to that of N . Since we assume that $N(p)$ is convex and that the perturbation P can be set small enough, there is no problem to estimate the displacement in the direction of p , it is always as small as the derivative of first order in p of the perturbation. The problem is that we can not expect the displacement to be also small enough in the direction of q_n . We deal with this problem in two ways.

2a. The Critical Point has Weaker Persistency. By definition, the following relation holds:

$$\left| \left(\frac{\partial^2 N}{\partial q_n^2} - \sum_{i,j=1}^n a_{ij} \frac{\partial^2 N}{\partial p_i \partial q_n} \frac{\partial^2 N}{\partial p_j \partial q_n} \right)_{p=p^*, q_n=q_n^*} \right| \leq c_2 d^{\frac{2}{13}}, \tag{2.1}$$

where (a_{ij}) is the inverse of the Hessian matrix of the main part $N(p, q_n)$ with respect to p , (p^*, q_n^*) is the critical point of $N(p, q_n) - \langle \omega, \hat{p} \rangle$, $(\hat{p} = p_1, p_2, \dots, p_{n-1})$. To see what (2.1) means, we consider the main part N . By assumption, the main part of the Hamiltonian $N(p, q_n)$ is nonsingular with respect to p in the sense that the Hessian matrix of N in p is nonsingular for all $q_n \in L \subset \mathbb{T}$, and $\rho(q_n)$ is the critical point of $N(p, q_n) - \langle \omega, \hat{p} \rangle$. As q_n ranges over L it actually defines a curve $\{p = \rho(q_n), q_n \in L\}$ ($\rho(q_n)$ is clearly real analytical). Denoting by q_n^* the critical point of $N(\rho(q_n), q_n) - \langle \omega, \hat{p}(q_n) \rangle$, then $(p^* = \rho(q_n^*), q_n^*)$ is the critical point of $N(p, q_n) - \langle \omega, \hat{p} \rangle$. The condition (2.1) actually is the following:

$$\left| \frac{d^2}{dq_n^2} (N(\rho(q_n), q_n) - \langle \omega, \hat{p}(q_n) \rangle)_{q_n=q_n^*} \right| \leq c_2 d^{\frac{2}{13}}. \tag{2.1'}$$

To distinguish whether the critical point has stronger persistency or not there is no precise criterion, it depends on how we run the KAM machine. There is a range for the constant c_2 , where it is neither too large nor too small. When c_2 takes its value in this region we can deal with such a problem in both ways.

In the case of weaker persistency, the critical point of $(N + N_+)(p, q_n) - \langle \omega, \hat{p} \rangle$ might be far away from that of $N(p, q_n) - \langle \omega, \hat{p} \rangle$. So we can not expand the Hamiltonian in Taylor series of (p_n, q_n) as in previous work to study lower dimensional tori ([G, E, P]), because we do not know how large the domain for q_n is so that the critical point of $(N + N_+)(p, q_n) - \langle \omega, \hat{p} \rangle$ is still within. However, since the nonsingularity of the Hessian matrix of N is assumed with respect to p , the critical point of $N(p, q_n) - \langle \omega, \hat{p} \rangle$ in p for fixed q_n , $\rho(q_n)$ is not far from the critical point of $(N + N_+)(p, q_n) - \langle \omega, \hat{p} \rangle$ in p for the same q_n , denoted by $\rho_+(q_n)$. So we only need to study the Hamiltonian in the neighbourhood of the line $\{(\rho(q_n), q_n), q_n \in L\}$. The choice of L is very delicate. On one hand, it has to be big enough so that the domain for q_n is much bigger than that for p and the critical point of ϕ_+ in the next step of iteration is still in this region; on the other hand, it can not be too big, because we need the estimate $|\phi'(q_n)| \leq O(d^{\frac{4}{13}})$ in the domain $\{|Im q| \leq d^{\frac{2}{13}}, Re q_n \in L\}$.

Expanding the perturbation in Taylor series of $p_n - \rho_n(q_n)$ and retaining the sum up to order 5,

$$TP = \sum_{j=0}^5 P_j(\hat{p}, q)(p_n - \rho_n(q_n))^j, \tag{2.2}$$

then the remaining part $RP = P - TP$ is much smaller than P in some smaller domain of $p_n - \rho_n(q_n)$. We also develop TP into the Fourier series of \hat{q} and truncate it up to order K ,

$$\mathcal{F}_K P_j = \sum_{\substack{k \in \mathbb{Z}^{n-1} \setminus \{0\} \\ |k| \leq K}} P_{jk} e^{i\langle k, \hat{q} \rangle};$$

the remaining part $\mathcal{R}_K(TP) = TP - \mathcal{F}_K(TP)$ can also be made much smaller than the truncated part. We choose the generating function W in the form

$$W(p, q) = \sum_{j=0}^5 W_j(\hat{p}, q)(p_n - \rho_n(q_n))^j, \tag{2.3}$$

where W_j is the solution of the following equation:

$$\langle \omega, W_{j\hat{q}} \rangle - j \frac{\partial^2 N}{\partial p_n \partial q_n}(\rho(q_n), q_n) W_j = -\mathcal{F}_K P_j(\hat{p}, q). \tag{2.4}$$

By restricting the domain of q_n properly small, e.g. $|Im q_n| \leq d^{\frac{2}{13}}$, $Im(N_{p_n q_n}(\rho(q_n), q_n))$ can be set so small such that

$$|\langle \omega, k \rangle - j Im(N_{p_n q_n}(\rho(q_n), q_n))| \geq \frac{1}{2} |\langle \omega, k \rangle|, \quad \forall j \leq 5, \forall |k| \leq K$$

if we choose K carefully [MP]. In this way we then find a well defined generating function if ω satisfies the Diophantine condition (1.8). If a symplectic transformation $\mathcal{M} : (p, q) \rightarrow (p_+, q_+)$:

$$p = p_+ + W_q, \quad q_+ = q + W_{p_+} \tag{2.5}$$

is introduced with the generating function $W(p_+, q)$ being determined as above, it takes the Hamiltonian $H(p, q)$ into the form

$$H(p, q) = (N + N_+)(p_+, q_{n_+}) + P_+(p_+, q_+), \tag{2.6}$$

where

$$N_+ = \frac{1}{(2\pi)^{n-1}} \int_0^{2\pi} P(p, q) d\hat{q},$$

$$P_+ = (M_1 + M_2 + M_3 + M_4)(p_+, q_+),$$

$$M_1 = N(p_+ + W_q, q_n) - N(p_+, q_n + W_{p_{n+}}) - \langle \omega, W_{\hat{q}} \rangle$$

$$+ \frac{\partial^2 N}{\partial p_n \partial q_n}(\rho(q_n), q_n)(p_{n+} - \rho_n(q_n))W_{p_{n+}},$$

$$M_2 = (P - N_+)(p, q) - (P - N_+)(p_+, q),$$

$$M_3 = R(P - N_+)(p_+, q) + \mathcal{R}_K(T(P - N_+))(p_+, q),$$

$$M_4 = N_+(p, q) - N_+(p_+, q_+).$$

It is easier to see that M_2 and M_3 as well as M_4 can be set very small. To see M_1 is also small we need to write it in some more transparent way. Let

$$N(p, q_n) = \phi(q_n) + \langle \omega, \hat{p} \rangle + \tilde{N}(p, q_n), \tag{2.7}$$

where $\phi(q_n) = N(\rho(q_n), q_n) - \langle \omega, \hat{\rho}(q_n) \rangle = \min_p \{N(p, q_n) - \langle \omega, \hat{p} \rangle\}$, we have

$$\tilde{N}_{p_i q_n}(p, q_n) = N_{p_i q_n}(p, q_n),$$

$$\tilde{N}_{p_i p_j}(p, q_n) = N_{p_i p_j}(p, q_n),$$

$$\tilde{N}(\rho(q_n), q_n) = \tilde{N}_p(\rho(q_n), q_n) = \tilde{N}_{q_n}(\rho(q_n), q_n) = 0, \tag{2.8}$$

since on the real line $\{(\rho(q_n), q_n), q_n \in L\}$ they are identical with zero. Then

$$M_1 = M_{11} + M_{12} + M_{13}, \tag{2.9}$$

$$M_{11} = \phi(q_{n+} - W_{p_{n+}}) - \phi(q_{n+}) = \phi'(q_{n+})W_{p_{n+}} + \frac{1}{2}\phi''(q_{n+} + \eta W_{p_{n+}})W_{p_{n+}}^2,$$

$$M_{12} = \tilde{N}(p_+ + W_q, q_n) - \tilde{N}(p_+, q_n) = \frac{1}{2} \sum_{i,j=1}^n N_{p_i p_j}(\rho(q_n) + \xi_1 \Delta p_+, q_n) \Delta p_{i+} W_{q_j}$$

$$+ \frac{1}{2} \sum_{i,j=1}^n N_{p_i p_j}(p_+ + \xi_2 W_q, q_n) W_{q_i} W_{q_j},$$

$$M_{13} = \tilde{N}(p_+, q_n) - \tilde{N}(p_+, q_n + W_{p_{n+}}) + N_{p_n q_n}(\rho(q_n), q_n) \Delta p_{n+} W_{p_{n+}}$$

$$= - \sum_{i=1}^{n-1} N_{p_i q_n}(\rho(q_n), q_n) \Delta p_i + W_{p_{n+}} - \frac{1}{2} N_{q_n^2}(p_+, q_n + \eta W_{p_{n+}}) W_{p_{n+}}^2$$

$$- \frac{1}{2} \sum_{i,j=1}^n N_{p_i p_j q_n}(\rho(q_n) + \xi_3 \Delta p_+, q_n) \Delta p_{i+} \Delta p_j + W_{p_{n+}},$$

where some notations are used as in the following:

$$\xi_j(W_q) = (\xi_{j1}W_{q_1}, \dots, \xi_{jn}W_{q_n}), \quad \Delta p_+ = p_+ - \rho(q_n),$$

$|\xi_{ij}| \leq 1$, ($i = 1, 2, 3$; $j \leq n$) and $|\eta| \leq 1$. If we let $|\text{Im } q_n| \leq O(d^{\frac{2}{13}})$, $|p_n - \rho_n(q_n)| \leq O(d^{\frac{3}{13}})$ and $|p_j - \rho_j(q_n)| \leq O(d^{\frac{4}{13}})$, ($j < n$), and choose the domain of q_n carefully, which will be demonstrated in detail in the next section, we are able to make P_+ and the moduli of its gradient much smaller than the original one on one hand; on the other hand, the critical point of $N + N_+(p, q_n) - \langle \omega, \hat{p} \rangle$ remains in the prescribed domain, which is one of the key points making the iteration scheme work.

If the perturbation is small enough, such a transformation can then be repeated. In every step, a symplectic change of coordinates \mathcal{M}_l is set up in this way so that

$$H_l \circ \mathcal{M}_l = N_{l+1} + P_{l+1}$$

with another main part N_{l+1} and a much smaller error term P_{l+1} , for instance

$$|P_{l+1}| \leq |P_l|^\kappa$$

for some $\kappa > 1$. Such a procedure might be repeated infinitely many times, if those critical points always have weaker persistency. It results in a sequence of transformations $\mathcal{M}_0, \mathcal{M}_1 \dots$ whose infinite product converges on a set containing at least one trivially embedded $(n - 1)$ -torus $(p, q_n) = (\tilde{p}, \tilde{q}_n)$ and transforms it to a set containing at least one $(n - 1)$ -torus with its tangent map taking constant vector fields on (\tilde{p}, \tilde{q}_n) to the vector field governed by (1.1). Such a procedure also might be repeated only for a finite number of times, when the critical point gets stronger persistency we may switch to the other procedure described in the following, which guarantees that the iteration scheme works successfully.

2b. The Critical Point has Stronger Persistency. Here (2.1) does not hold. In this case the critical point of $N + N_+ - \langle \omega, \hat{p} \rangle$ remains close to that of $N - \langle \omega, \hat{p} \rangle$. We shall show below that once the critical point gets stronger persistency the critical point in the following steps of iteration shall possess stronger persistency as well. Therefore, we can restrict (p_n, q_n) in a smaller domain $|p_n - p_n^*| \leq O(d^{\frac{3}{13}})|q_n - q_n^*| \leq O(d^{\frac{3}{13}})$ and expand P in a Taylor series in (p_n, q_n) ,

$$P = \sum P_{ij}(\hat{p}, \hat{q})(p_n - p_n^*)^i (q_n - q_n^*)^j,$$

and use the truncation TP of P to approximate P . TP will be chosen such that these sums only extend over $|i + j| \leq 5$. By restricting the domains of (p_n, q_n) properly small, $P - TP$ is much smaller than P in the relevant domain. To find a symplectic transformation (2.5) we introduce the generating function W by

$$W(p_+, q) = \sum_{|i+j| \leq 5} W_{ij}(\hat{p}_+, \hat{q})(p_{n+} - p_n^*)^i (q_n - q_n^*)^j, \tag{2.10}$$

where W_{ij} are determined by the following equations:

$$\partial \mathbf{W}_l + \mathbf{A}_l \mathbf{W}_l = -\tilde{\mathbf{P}}_l, \quad (0 \leq l \leq 5), \tag{2.11}$$

where $\mathbf{W}_l = (W_{l0}, W_{(l-1)1}, \dots, W_{0l})^t$, $\tilde{\mathbf{P}}_l = (\tilde{P}_{l0}, \tilde{P}_{(l-1)1}, \dots, \tilde{P}_{0l})^t$, $\partial W_{ij} = \langle \omega, W_{ij\hat{q}} \rangle$, $\tilde{P}_{ij} = P_{ij} - [P_{ij}]$,

$$[P_{ij}] = \frac{1}{(2\pi)^{n-1}} \int_0^{2\pi} P_{ij}(\hat{p}, \hat{q}) d\hat{q},$$

and $\mathbf{A}_0 = 0$

$$\mathbf{A}_1 = \begin{pmatrix} -N_{p_n q_n} & N_{p_n^2} \\ -(\phi'' - N_{p_n q_n} \rho'_n) & -N_{p_n^2} \rho'_n \end{pmatrix} := \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},$$

and other matrices are defined through the entries of \mathbf{A}_1 in the following way:

$$\mathbf{A}_l = (a_{ij})_{(l+1) \times (l+1)}, \quad 2 \leq l \leq 5,$$

$$a_{jj} = (l + 1 - j)b_{11} + (j - 1)b_{22},$$

$$a_{j(j+1)} = j b_{12},$$

$$a_{(j+1)j} = (l + 1 - j)b_{21},$$

$$a_{ij} = 0, \quad (|i - j| \geq 2).$$

All entries in these matrices are evaluated at the critical point.

These equations in (2.11) can be solved in this way. First we expand W_{ij} and \tilde{P}_{ij} into the Fourier series

$$W_{ij} = \sum_{k \in \mathbb{Z}^{n-1} \setminus \{0\}} W_{ijk}(\hat{p}) e^{i\langle k, \hat{q} \rangle},$$

$$\tilde{P}_{ij} = \sum_{k \in \mathbb{Z}^{n-1} \setminus \{0\}} P_{ijk}(\hat{p}) e^{i\langle k, \hat{q} \rangle}, \quad i, j \leq 5,$$

then

$$(i\langle k, \omega \rangle \mathbf{I} + \mathbf{A}_j) \mathbf{W}_{jk} = -\tilde{\mathbf{P}}_{jk},$$

where \mathbf{I} is the identity matrix. Since the spectrum of \mathbf{A}_j is as follows:

$$Sp(\mathbf{A}_j) = \{(j - l)\sigma_1 + l\sigma_2, \quad l = 0, 1, \dots, j\},$$

$\sigma_{1,2} = \frac{1}{2} \{ -(N_{p_n q_n} + N_{p_n^2} \rho'_n) \pm \sqrt{(N_{p_n q_n} + N_{p_n^2} \rho'_n)^2 - 4\phi'' N_{p_n^2}} \}$, we find

$$\mathbf{W}_{0k} = \frac{i\tilde{\mathbf{P}}_{0k}}{\langle k, \omega \rangle},$$

$$\mathbf{W}_{jk} = -\frac{(i\langle k, \omega \rangle \mathbf{I} + \mathbf{A}_j)^* \tilde{\mathbf{P}}_{1k}}{\prod_{\lambda_l \in Sp(\mathbf{A}_j)} (i\langle k, \omega \rangle + \lambda_l)}, \quad (j \leq 5), \tag{2.12}$$

where \mathbf{A}_j^* denotes the adjoint matrix of \mathbf{A}_j . Because the critical point of $N(\rho(q_n), q_n) - \langle \omega, \hat{\rho}(q_n) \rangle$ is found by looking for its maximum, we have $\phi''(q_n) < 0$, which guarantees that all $\lambda_l \in Sp(\mathbf{A}_j)$ are real numbers of $N_{p_n^2} > 0$ due to the definite positivity of the Hamiltonian. So we need not worry about the problem of small divisor if ω satisfies the Diophantine condition (1.8). Now the generating function is well defined, and the symplectic map of (2.5) transforms the Hamiltonian $H(p, q)$ into

$$H(p, q) = (N + N_+)(p_+, q_{n+}) + P_+(p_+, q_+), \tag{2.13}$$

where

$$\begin{aligned}
 N_+ &= \frac{1}{(2\pi)^{n-1}} \int_0^{2\pi} TP(p_+, q) d\hat{q}, \\
 P_+ &= (M_1 + M_2 + M_3 + M_4)(p_+, q_+), \\
 M_1 &= N(p_+ + W_q, q_n) - N(p_+, q_n + W_{p_{n+}}) - \langle \omega, W_{\hat{q}} \rangle \\
 &\quad - (p_{n+} - p_n^*, q_n - q_n^*) \begin{pmatrix} N_{p_n^2} & N_{p_n q_n} \\ -N_{p_n^2} \rho'_n & \phi'' - N_{p_n q_n} \rho'_n \end{pmatrix} \begin{pmatrix} W_{q_n} \\ -W_{p_{n+}} \end{pmatrix}, \\
 M_2 &= (P - TP)(p, q), \\
 M_3 &= (TP - N_+)(p, q) - (TP - N_+)(p_+, q), \\
 M_4 &= N_+(p, q_n) - N_+(p_+, q_{n+}),
 \end{aligned}$$

where the entries of the matrix are evaluated at the critical point. To make M_1 easier to study, we also use the notation of (2.7) and identities in (2.8), then

$$M_1 = M_{11} + M_{12} + M_{13} + M_{14} + M_{15}, \tag{2.14}$$

where

$$\begin{aligned}
 M_{11} &= \phi(q_n) - \phi(q_n + W_{p_{n+}}) + \phi''(q_n^*)(q_n - q_n^*)W_{p_{n+}}, \\
 M_{12} &= \tilde{N}(p_+ + W_q, q_n) - \tilde{N}(p_+, q_n) - N_{p_n^2}(\rho(q_n), q_n)\Delta p_{n+}W_{q_n} \\
 &= \sum_{i,j=1}^n N_{p_i p_j}(\rho(q_n), q_n)\Delta p_{i+}W_{q_j} - N_{p_n^2}(\rho(q_n), q_n)\Delta p_{n+}W_{q_n} \\
 &\quad + \frac{1}{2} \sum_{i,j,k=1}^n N_{p_i p_j p_k}(\rho(q_n) + \xi_1 \Delta p_+, q_n)\Delta p_i + \Delta p_j + W_{q_k} \\
 &\quad + \frac{1}{2} \sum_{i,j=1}^n N_{p_i p_j}(p_+ + \xi_2 W_q, q_n)W_{q_i}W_{q_j}, \\
 M_{13} &= \tilde{N}(p_+, q_n) - \tilde{N}(p_+, q_n + W_{p_{n+}}) + N_{p_n q_n}(\rho(q_n), q_n)\Delta p_{n+}W_{p_{n+}} \\
 &= -\sum_{i=1}^{n-1} N_{p_i q_n}(\rho(q_n), q_n)\Delta p_{i+}W_{p_{n+}} - \frac{1}{2}N_{q_n^2}(p_+, q_n + \eta W_{p_n})W_{p_{n+}}^2 \\
 &\quad - \frac{1}{2} \sum_{i,j=1}^n N_{p_i p_j q_n}(\rho(q_n) + \xi_3 \Delta p_+, q_n)\Delta p_{i+}\Delta p_j + W_{p_{n+}}, \\
 M_{14} &= N_{p_n^2}(\rho(q_n), q_n)\Delta p_{n+}W_{q_n} - N_{p_n^2}(p^*, q_n^*)(p_{n+} - p_{n+}^*)W_{q_n} \\
 &\quad + N_{p_n^2}(p^*, q_n^*)\rho'_n(q_n^*)(q_n - q_n^*)W_{q_n}, \\
 M_{15} &= -N_{p_n q_n}(\rho(q_n), q_n)\Delta p_{n+}W_{p_n} + N_{p_n q_n}(p^*, q_n^*)(p_{n+} - p_{n+}^*)W_{p_n} \\
 &\quad - N_{p_n q_n}(p^*, q_n^*)\rho'_n(q_n^*)(q_n - q_n^*)W_{p_n}, \\
 \Delta p_+ &= p_+ - \rho(q_n).
 \end{aligned}$$

By shrinking the domain of the new variables properly, we are able to make P_+ and the moduli of its gradient much smaller than the original ones in a smaller domain. Since “stronger persistency” has been assumed, in such a smaller domain the new main part of the Hamiltonian possesses a critical point of the same type in such a domain. This procedure yields a sequence of transformations $\mathcal{M}_0, \mathcal{M}_1 \dots$ whose infinite product converges on a trivially embedded $(n - 1)$ -torus $(p, q_n) = (\tilde{p}, \tilde{q}_n)$ and transforms, it to a $(n - 1)$ -torus with its tangent map taking constant vector field on (\tilde{p}, \tilde{q}_n) to the vector field determined by (1.1).

3. The Inductive Step

We begin the iteration by demonstrating in detail a single step which furnishes the building blocks. Such iteration is constructed in two ways which depends on whether the critical point has stronger persistency or not as we have seen in the last section. In the case of weaker persistency we need to construct the nested domains for canonical variables in some implicit way so that the critical point for the next step of iteration is still in the chosed domain.

3a. Weaker Persistency. We assume the Hamiltonian $N(p, q_n) + P(p, q)$ is real analytical in the domain $\mathcal{A}_t \times \Sigma_s$ where

$$\begin{aligned} \mathcal{A}_t &= \{ |\operatorname{Im} q_i| \leq t, \operatorname{Re} q_i \in \mathbb{T}, (i \leq n - 1) \}, \\ \Sigma_s &= \{ |p_i - \rho_i(q_n)| \leq s^4, (i \leq n - 1), \\ &\quad |p_n - \rho_n(q_n)| \leq s^3, |\operatorname{Im} q_n| \leq s^2, \operatorname{Re} q_n \in L \}, \end{aligned}$$

and L is either an interval contractable to one point or the whole \mathbb{T} . If $L \subset \mathbb{T}$ is homotopic to a point, then $L \supseteq [q_n^* - s^2, q_n^* + s^2]$.

Let s_+, t_+ be positive numbers such that

$$\begin{aligned} 2s_+ &\leq s, \quad s \leq t \leq 1, \quad s_+^4 \geq s^5, \\ s_+ &\leq t_+, \quad s_+ \leq \frac{1}{4}(t - t_+), \end{aligned} \tag{3.1}$$

$$s^2 |\log s|^\mu \leq \frac{D}{2} \left(\frac{3(t - t_+)}{64(n - 1)} \right)^\mu, \quad s^2 \leq \left(\frac{e}{2n - 2} \right)^{n-1} \left(\frac{t - t_+}{32} \right)^n. \tag{3.2}$$

We assume further in $\mathcal{A}_t \times \Sigma_s$:

(3.a1) The smallness of the perturbation P ,

$$|P(p, q)| \leq \delta,$$

where

$$\delta \leq \min \left\{ \min \left\{ \frac{D(t - t_+)^{\mu}}{4c_1 8^{\mu+2}}, \frac{D^6(t - t_+)^{6\mu}}{8c_6} \right\} (s^4 - 3s_+^4)(s^3 - 3s_+^3), \frac{\lambda}{12n} s_+^2 s^{10} \right\}. \tag{3.3}$$

(3.a2) The positivity of N : the Hessian matrix of N in p is definitely positive in $\text{Re } \Sigma_s$ and

$$\min_{\Sigma_s} \left\| \frac{\partial^2 N}{\partial p^2} \xi \right\| \geq \lambda \|\xi\|, \quad (\forall \xi \in \mathbb{C}^n)$$

Without losing generality we assume $0 < \lambda \leq 1$.

(3.a3) The critical point (p^*, q_n^*) is determined by looking for the maximum of the function $\phi(q_n) = N(\rho(q_n), q_n) - \langle \omega, \hat{p}(q_n) \rangle$ ($q_n \in L$), where $\rho(q_n)$ is determined by the condition

$$\frac{\partial}{\partial p} (N(p, q_n) - \langle \omega, \hat{p} \rangle)_{p=\rho(q_n)} = 0.$$

On the whole line $\{(\rho(q_n), q_n), q_n \in L\}$,

$$|\phi''(q_n)| \leq 2s^2. \tag{3.4}$$

If L is a contractible interval, denoted by $[l, r]$, it is assumed $l \leq q_n^* - s^2$, $r \geq q_n^* + s^2$, at each end $q_n = l$, $q_n = r$, the following hypothesis holds:

$$|\phi'(q_n)| \geq 2s^4; \tag{3.5}$$

$\rho(q_n)$ is real analytical in $\{| \text{Im } q_n | \leq s^2, \text{Re } q_n \in L\}$.

(3.a4) In Σ_s

$$\max_{j \geq 3} \left| \frac{\partial^j N}{\partial p^j} \right| \leq \beta, \quad \max_{|(l_1, l_2)| \leq 3, l_2 \geq 1} \left| \frac{\partial^{l_1+l_2} N}{\partial p^{l_1} \partial q_n^{l_2}} \right| \leq \frac{1}{7n^2 2^\mu},$$

$$\max_{j \geq 3} |\phi^{(j)}(q_n)| \leq \frac{1}{64}, \quad \max_{j=1,2} |\rho^{(j)}(q_n)| \leq \frac{1}{2}. \tag{3.6}$$

Consider the following inequality:

$$|\phi''(q_n)|_{q_n=q_n^s} \leq s_+^2. \tag{3.7}$$

When this relation holds the critical point is referred to as a point with weaker persistency. We study this case first.

We introduce several intermediate domains which are defined somewhat implicitly:

$$\mathcal{A}_{t_+} \subset \mathcal{A}_1 \subset \mathcal{A}_2 \subset \mathcal{A}_t, \quad \Sigma_{s_+} \subset \Sigma_0 \subset \Sigma_1 \subset \Sigma_2 \subset \Sigma_s,$$

where

$$\mathcal{A}_{t_+} = \{ |\text{Im } q_i| \leq t_+, \text{Re } q_i \in \mathbb{T}, i \leq n-1 \},$$

$$\mathcal{A}_1 = \{ |\text{Im } q_i| \leq t_+ + s_+^2, \text{Re } q_i \in \mathbb{T}, i \leq n-1 \},$$

$$\mathcal{A}_2 = \left\{ |\text{Im } q_i| \leq t - \frac{1}{2}(t - t_+), \text{Re } q_i \in \mathbb{T}, i \leq n-1 \right\},$$

$$\begin{aligned} \Sigma_{s_+} &= \{|p_i - \rho_{+i}(q_n)| \leq s_+^4, (i \leq n - 1), \\ &\quad |p_n - \rho_{+n}(q_n)| \leq s_+^3, |\operatorname{Im} q_n| \leq s_+^2, \operatorname{Re} q_n \in L_+\}, \\ \Sigma_0 &= \{|p_i - \rho_i(q_n)| \leq 2s_+^4, (i \leq n - 1), \\ &\quad |p_n - \rho_n(q_n)| \leq 2s_+^3, |\operatorname{Im} q_n| \leq s_+^2, \operatorname{Re} q_n \in L_+\}, \\ \Sigma_1 &= \{|p_i - \rho_i(q_n)| \leq 3s_+^4, (i \leq n - 1), \\ &\quad |p_n - \rho_n(q_n)| \leq 3s_+^3, |\operatorname{Im} q_n| \leq 2s_+^2, \operatorname{Re} q_n \in L_1\}, \\ \Sigma_2 &= \left\{ |p_i - \rho_i(q_n)| \leq s^4 - \frac{1}{2}(s^4 - 3s_+^4), (i \leq n - 1), \right. \\ &\quad |p_n - \rho_n(q_n)| \leq s^3 - \frac{1}{2}(s^3 - 3s_+^3), \\ &\quad \left. |\operatorname{Im} q_n| \leq s^2 - \frac{1}{2}(s^2 - 2s_+^2), \operatorname{Re} q_n \in L_2 \right\}. \end{aligned}$$

$\{(\rho_+(q_n), q_n), q_n \in L_+\}$ is the critical point line of $N + N_+ - \langle \omega, \hat{p} \rangle$ determined in the same way as $(\rho(q_n), q_n)$ for $N - \langle \omega, \hat{p} \rangle$ (N_+ is defined in (2.6)). L_+ is either \mathbb{T}^1 or some interval $[l_+, r_+]$ when (3.a3) holds for $N + N_+$. When $L = [l, r] \neq \mathbb{T}^1, L_2 = [l + \frac{1}{2}(l_1 - l), r - \frac{1}{2}(r - r_1)]$, where $L_1 = [l_1, r_1]$ and its determination is implicit, in this case $[l_+, r_+] = [l_1 + s_+^2, r_1 - s_+^2]$.

L_1 is defined in the following way. Choosing $q_n^* - l_1$ and $r_1 - q_n^*$ as big as possible so that

$$|\phi'(q_n)| < 4s_+^4 \tag{3.8}$$

holds in (l_1, r_1) and at each end of this interval it becomes an identity. Since (3.7) is assumed we see $L_1 \supseteq [q_n^* - 3s_+^2, q_n^* + 3s_+^2]$. Indeed, for $|q_n - q_n^*| \leq 3s_+^2$,

$$|\phi'(q_n)| \leq 3s_+^2 |\phi''(q_n^*)| + \frac{9}{2}s_+^4 \sup |\phi^{(3)}(q_n)| < 4s_+^4.$$

When $L \neq \mathbb{T}$, by virtue of such a choice of L_1 it is ensured by (3.4) and (3.5),

$$\operatorname{dist}(L_1, \partial L) \geq \frac{1}{2}s^2. \tag{3.9}$$

If $L = \mathbb{T}$, let $L_2 = \mathbb{T}$, L_1 may also be determined by (3.8). If $L_1 \neq \mathbb{T}$, we choose L_+ the same as above. If (3.8) is satisfied on the whole \mathbb{T} , let $L_1 = L_+ = \mathbb{T}$.

To find the upper bound of $|\phi'(q_n)|$ in the domain $\{q_n | |\operatorname{Im} q_n| \leq 2s_+^2, \operatorname{Re} q_n \in L_1\}$ we need to formulate a technical lemma first.

Lemma 2. *Assume a function $f(x) \in C^2[a, b]$ is defined, satisfies $|f(x)| \leq A$, $|f''(x)| \leq B$ for all $x \in [a, b]$, and at some $x_0 \in [a, b]$ $|f'(x_0)| \leq \xi$ then for all $x \in [a, b]$ we have*

$$|f'(x)| \leq 2\sqrt{AB} + \xi.$$

Proof. Suppose $|f'(x)|$ reaches its maximum in $[a, b]$ at x_1 , $|f'(x_1)| = M \geq \xi \geq 0$. Without losing generality, we assume $f'(x_1) = M$. By our assumption, there exists $x_2 \in [a, b]$, where $f'(x_2) \leq \xi$ and $\xi \leq f'(x) \leq M$ for all $x \in [x_1, x_2]$. Since both at x_1 and $x_2 |f(x)| \leq A$,

$$2A \geq |f(x_1) - f(x_2)| = \left| \int_{x_2}^{x_1} f'(x) dx \right| \geq \frac{1}{2B} |M - \xi|^2,$$

which leads to our conclusion.

Applying this lemma to $\phi'(q_n)$ we find for $q_n \in L_1$,

$$|\phi''(q_n)| \leq 2\sqrt{\sup \phi^{(3)}(q_n) \sup \phi'(q_n)} + |\phi''(q_n^*)| \leq \frac{3}{2}s_+^2, \tag{3.10}$$

in view of (3.6) and (3.8), from which combined with (3.8) we get that for any $q_n \in \{q_n | |\text{Im } q_n| \leq 2s_+^2, \text{Re } q_n \in L_1\}$,

$$|\phi'(q_n)| \leq |\phi'(\text{Re } q_n)| + 2|\phi''(\text{Re } q_n)|s_+^2 + 4 \max |\phi^{(3)}|s_+^4 \leq 8s_+^4. \tag{3.11}$$

Now we introduce a symplectic transformation \mathcal{M} with the form of (2.5), where W is determined by searching for the solution of Eq. (2.4), which maps the domain $\mathcal{A}_1 \times \Sigma_1$ into $\mathcal{A}_2 \times \Sigma_2$. Let us check it. Let

$$K = \left\lceil \frac{1}{n-1} \left(\frac{D}{2s^2} \right)^{\frac{1}{\mu}} \right\rceil, \tag{3.12}$$

where $[a]$ denotes the largest integer among those smaller than a , and assume s is so small such that (3.2) is satisfied, then we have a bound for the truncated parts of perturbation [A2]

$$|\mathcal{R}_K P| \leq \left(\frac{2n-2}{e} \right)^{n-1} \left(\frac{32}{t-t_+} \right)^n \delta s^2, \quad |\mathcal{T}_K P| \leq 2\delta \tag{3.13}$$

if $|\text{Im } q_i| \leq t - \frac{1}{8}(t - t_+)$ ($1 \leq i < n$). On the other hand, in view of (3.6), when $|\text{Im } q_n| \leq s^2$,

$$\left| \text{Im} \left(j \frac{\partial^2 N}{\partial p_n \partial q_n}(\rho(q_n), q_n) \right) \right| \leq \frac{1}{2^\mu} s^2, \quad \forall j \leq 5.$$

So for $|k| \leq K, 0 \leq j \leq 5$ we have from (3.2),

$$\left| \langle k, \omega \rangle - \text{Im} \left(j \frac{\partial^2 N}{\partial p_n \partial q_n}(\rho(q_n), q_n) \right) \right| \geq \frac{1}{2} |\langle k, \omega \rangle|, \tag{3.14}$$

which leads to the estimate in the domain $\mathcal{A}_2 \times \Sigma_2$ on the generating function defined as follows:

$$W_j(p, q) = - \sum_{\substack{k \in \mathbb{Z}^{n-1} \setminus \{0\} \\ |k| \leq K}} \frac{P_{jk}(\hat{p}, q_n) e^{i\langle k, \hat{q} \rangle}}{i\langle k, \omega \rangle - j \frac{\partial^2 N}{\partial p_n \partial q_n}(\rho(q_n), q_n)}$$

by Cauchy’s technique:

$$\begin{aligned}
 |W(p_+, q)| &\leq \theta, \\
 |W_{\hat{q}}| &\leq 4\theta(t - t_+)^{-1}, \\
 |W_{\hat{p}}| &\leq 2\theta(s^4 - 3s_+^4)^{-1}, \\
 |W_{q_n}| &\leq 2\theta(s^2 - 2s_+^2)^{-1}, \\
 |W_{p_n}| &\leq 2\theta(s^3 - 3s_+^3)^{-1}, \\
 \max\{|W_{p_i p_j}|, |W_{p_i q_j}|, |W_{q_i q_j}|\} &\leq 8\theta(s^4 - 3s_+^4)^{-2}, \quad \forall i, j \leq n,
 \end{aligned} \tag{3.15}$$

where

$$\theta = \frac{4c_1 8^\mu \delta}{D(t - t_+)^{\mu}}.$$

Since $W(p_+, q)$ is real analytical, $\mathbb{R}^n \subset \mathbb{C}^n$ is mapped into \mathbb{R}^n by \mathcal{M} defined in (2.5). It follows from (3.1), (3.3) and (3.15) that

$$\begin{aligned}
 |\hat{q} - \hat{q}_+| &\leq |W_{\hat{p}}| \leq \frac{1}{2}(t - t_+) - s_+^2, \\
 |q_n - q_{n+}| &\leq |W_{p_{n+}}| \leq \frac{1}{4}s^2, \\
 |\hat{p} - \hat{p}(q_n)| &\leq |(\hat{p} - \hat{p}(q_n)) - (\hat{p}_+ - \hat{p}(q_{n+}))| + |\hat{p}_+ - \hat{p}(q_{n+})| \\
 &\leq |W_{\hat{q}}| + |\rho'| |W_{p_{n+}}| + 3s_+^4 \leq \frac{1}{2}(s^4 + 3s_+^4), \\
 |p_n - \rho_n(q_n)| &\leq |(p_n - \rho_n(q_n)) - (p_{n+} - \rho_n(q_{n+}))| + |p_{n+} - \rho_n(q_{n+})| \\
 &\leq \frac{1}{2}(s^3 + 3s_+^3),
 \end{aligned} \tag{3.16}$$

which implies \mathcal{M} maps $\mathcal{A}_1 \times \Sigma_1$ into $\mathcal{A}_2 \times \Sigma_2$ (notice (3.9) and the definition of L_2). Furthermore it is obvious in $\mathcal{A}_1 \times \Sigma_1$,

$$|\mathcal{M} - \text{id}| \leq 4\theta(s^4 - 3s_+^4)^{-1}. \tag{3.17}$$

Finally we figure out the bound for the Jacobian matrix in $\mathcal{A}_1 \times \Sigma_1$,

$$\left[\frac{\partial(p, q)}{\partial(p_+, q_+)} - \text{Id} \right].$$

The principle difficulty is that \mathcal{M} is defined implicitly and therefore it is necessary to invert \mathcal{M} in an appropriate domain. For this, let us recall some lemmas; their proof can be found in [A2] and [G].

Lemma 3. *Suppose $x = y + \phi(y)$ in the k -dimensional strip $\Sigma_a = \{|\text{Im } y| \leq a\}$, where ϕ is real analytic. Then given a positive number ξ satisfying $a > 2\xi$ and $|\phi|/\xi \leq \frac{1}{2}$, there exist a unique real analytic function f on $\Sigma_{a-2\xi}$ such that $y = x + f(x)$, where $|f| \leq 2|\phi|, |f_x| \leq 4|\phi|/\xi$ in $\Sigma_{a-2\xi}$.*

Lemma 4. *Suppose $x = y + \phi(y)$ in the region U containing a ball of radius $5a$, where ϕ is real analytic and satisfies $|\phi| \leq \frac{a}{2}$. Then there exist a unique real analytic function f on $U - 5a$ such that $y = x + f(x)$ and $|f| \leq |\phi|, |f_x| \leq |\phi|/a$ for $x \in U - 5a$.*

We use $U - a$ to denote the set of those points contained in U whose a -neighbourhoods are also contained in U .

Due to Lemma 2 and Lemma 3, the equations

$$q_+ = q + W_{p_+}(p_+, q),$$

can be inverted to

$$q = q_+ + Q(p_+, q_+)$$

in $\{|\operatorname{Im} q| \leq t_+ + s_+^2\} \times \{|\operatorname{Im} q_n| \leq 2s_+^2, \operatorname{Re} q_n \in L_1\} \times \{|\hat{p} - \hat{\rho}(q_n)| \leq s^4 - \frac{1}{2}(s^4 - 3s_+^4)\} \times \{|p_n - \rho_n(q_n)| \leq s^3 - \frac{1}{2}(s^3 - 3s_+^3)\}$, provided δ satisfies (3.3). In that domain, by (3.16), we have

$$|Q| \leq 4\theta(s^4 - 3s_+^4)^{-1},$$

$$\left| \frac{\partial Q}{\partial q_+} \right| \leq 40\theta(s^4 - 3s_+^4)^{-1}s^{-2}.$$

The remaining entries in the Jacobian $|\frac{\partial(p, q)}{\partial(p_+, q_+)} - \operatorname{Id}|$ are easily estimated. In the sequel, it is assumed $(p, q) \in \mathcal{A}_2 \times \Sigma_2, (p_+, q_+) \in \mathcal{A}_1 \times \Sigma_1$. Plainly,

$$\left| \frac{\partial q}{\partial p_+} \right| \leq 8\theta(s^4 - 3s_+^4)^{-2}.$$

Expression for the partial derivatives of p is obtained formally from the relations

$$p = p_+ + W_q(p_+, q).$$

Each such partial derivatives can be bounded by using the previous estimates; e.g.

$$\frac{\partial p}{\partial q_+} = W_{q^2} \frac{\partial q}{\partial q_+}.$$

Therefore

$$\left| \frac{\partial(p, q)}{\partial(p_+, q_+)} - \operatorname{Id} \right| \leq 80\theta(s^4 - 3s_+^4)^{-2}. \tag{3.18}$$

The next step is to figure out the bound for the new error term P_+ in the domain $\mathcal{A}_{t_+} \times \Sigma_o$. Firstly, it is derived from (2.6) and (2.9) that in $\mathcal{A}_1 \times \Sigma_1$,

$$\begin{aligned} |M_{11}| &\leq |\phi'(q_n)| |W_{p_{n+}}| + \max |\phi^{(3)}| |W_{p_{n+}}|^2 \\ &\leq 16\theta s_+^4 (s^3 - 3s_+^3)^{-1} + 4\theta^2 (s^3 - 3s_+^3)^{-2}, \end{aligned} \tag{see (3.11)}$$

$$\begin{aligned} |M_{12}| &\leq n^2 \beta (|\hat{p}_+ - \hat{\rho}(q_n)| |W_{\hat{q}}| + |p_{n+} - \rho_n(q_n)| |W_{q_n}| + |W_q|^2) \\ &\leq n^2 \beta (6\theta s_+^3 (s^2 - 2s_+^2)^{-1} + 12\theta s_+^4 (t - t_+)^{-1} + 4\theta^2 (s^2 - 2s_+^2)^{-2}), \end{aligned}$$

$$\begin{aligned}
 |M_{13}| &\leq \beta(6n\theta s_+^4(s^3 - 3s_+^3)^{-1} + 4\theta^2(s^3 - 3s_+^3)^{-2} + 9n^2\theta s_+^6(s^3 - 3s_+^3)^{-1}), \\
 |M_2| &\leq 8n\delta\theta(s^4 - 3s_+^4)^{-1}s^{-4}, \\
 |M_3| &\leq c_3 \left(\delta \left(\frac{s_+}{s} \right)^{15} + \theta s^2 \right), \\
 |M_4| &\leq 16n\delta\theta(s^4 - 3s_+^4)^{-1}s^{-4}. \quad (\text{see (3.13)})
 \end{aligned}$$

In conclusion we have

$$|P_+| \leq c_4 \left\{ \theta s_+^3(s^2 - 2s_+^2)^{-1} + \delta \left(\frac{s_+}{s} \right)^{15} \right\} := \delta_+,$$

if we take (3.1), (3.2) and (3.3) into account. Here c_4 depends on β . Shrinking $\mathcal{A}_1 \times \Sigma_1$ to $\mathcal{A}_{t_+} \times \Sigma_o$, then

$$\begin{aligned}
 |P_+(p_+, q_+)| &\leq \delta_+, \\
 \max \left\{ \left| \frac{\partial P_+}{\partial p_+} \right|, \left| \frac{\partial P_+}{\partial q_+} \right| \right\} &\leq \frac{\delta_+}{s_+^4}, \\
 \max_{|l_1+l_2|=2} \left\{ \left| \frac{\partial^2 P_+}{\partial p_+^{l_1} \partial q_+^{l_2}} \right| \right\} &\leq \frac{2\delta_+}{s_+^8}, \\
 \max_{|l_1+l_2|=3} \left\{ \left| \frac{\partial^3 P_+}{\partial p_+^{l_1} \partial q_+^{l_2}} \right| \right\} &\leq \frac{6\delta_+}{s_+^{12}}. \quad (3.19)
 \end{aligned}$$

To complete one step of the iteration, we need to localize a critical point (p_+^*, q_{n+}^*) of the function $F = (N + N_+)(p, q_n) - \langle \omega, \hat{p} \rangle$ by the same minimax principle as above and find a domain $\mathcal{A}_{t_+} \times \Sigma_{s_+}$ centered at (p_+^*, q_{n+}^*) satisfying conditions similar to those shown from (3.a2) to (3.a4).

In Σ_2 the estimate

$$|\rho_+(q_n) - \rho(q_n)| \leq \frac{4n\delta}{\lambda(s^4 - 3s_+^4)}, \quad (3.20)$$

can be easily drawn from the fact that $|N_+| \leq 2\delta$ because of $|P| \leq \delta$. The argument is the following. Choosing $r > 0$ such that

$$\frac{4n\delta}{\lambda(s^4 - 3s_+^4)} < r \leq s^4 - \frac{1}{2}(s^4 - 3s_+^4),$$

which is possible due to (3.3), and defining \mathcal{D} is a n -multidisc in \mathbb{C}^n , depending on q_n ,

$$\mathcal{D} = \{ |p_i - \rho(q_n)| \leq r, 1 \leq i \leq n \}.$$

Since in $\Sigma_2 |N_{+p}| \leq \frac{4\delta}{s^4 - 3s_+^4}$, we then have for $0 \leq \varepsilon \leq 1$,

$$\text{dist}\{(N_p + \varepsilon N_{+p})(\partial\mathcal{D}), (\omega, 0)\} \geq \|N_p(\partial\mathcal{D}) - N_p(\rho(q_n))\| - \frac{4n\delta}{s^4 - 3s_+^4} > 0$$

because of (3.a2), while the degree

$$\text{deg}\{N_p + \varepsilon N_{+p}, \mathcal{D}, (\omega, 0)\} = 1$$

when $\varepsilon = 0$. Clearly $\rho_+(q_n)$ is also real analytical in $\{|\text{Im } q_n| \leq s^2, \text{Re } q_n \in L\}$.

By the choice of δ and s , (3.20) implies that $\Sigma_{s_+} \subset \Sigma_o$. To make sure the maximum point of the function $(N + N_+)(\rho_+(q_n), q_n) - \langle \omega, \hat{\rho}_+(q_n) \rangle$ remains in L_+ we only need to notice that for $l_1 \leq q \leq l_1 + 2s_+^2$ and $r_1 - 2s_+^2 \leq q \leq r_1$,

$$\begin{aligned} \inf_{\substack{l_1 \leq q \leq l_1 + 2s_+^2 \\ r_1 - 2s_+^2 \leq q \leq r_1}} |\phi'(q_n)| &\geq 4s_+^4 - \max_{q_n \in L_1} |\phi''| 2s_+^2 \geq s_+^4, \\ |\phi'(q_n)|_{q_n=l_1+s_+^2, r_1-s_+^2} &\geq \frac{5}{2}s_+^4, \end{aligned} \tag{3.21}$$

from which it follows

$$\begin{aligned} \phi(l_1 + s_+^2) &\leq \max\{\phi(l_1), \phi(l_1 + 2s_+^2)\} - s_+^6, \\ \phi(r_1 - s_+^2) &\leq \max\{\phi(r_1), \phi(r_1 - 2s_+^2)\} - s_+^6, \end{aligned}$$

it leads to

$$\begin{aligned} \phi(q_n^*) &\geq \max\{\phi(l_1), \phi(l_1 + 2s_+^2), \phi(r_1), \phi(r_1 - 2s_+^2)\} \\ &\geq \max\{\phi(l_1 + s_+^2), \phi(r_1 - s_+^2)\} + s_+^6. \end{aligned}$$

On the other hand, from the definition of ϕ_+ we have

$$\begin{aligned} \phi(q_n) - \delta &\leq \phi_+(q_n) \leq \phi(q_n) + \delta, \\ \max \phi_+(q_n) &\geq \phi_+(q_n^*) \geq \phi(q_n^*) - \delta, \end{aligned}$$

from which we finally obtain

$$\max \phi_+(q_n) \geq \max\{\phi_+(l_1 + s_+^2), \phi_+(r_1 - s_+^2)\} + s_+^6 - 2\delta.$$

It shows only in the interior of L_+ can ϕ_+ reach its maximum. In (3.a3) we still need to verify

$$\begin{aligned} |\phi_+''(q_n)|_{l_+ \leq q_n \leq r_+} &\leq 2s_+^2, \\ |\phi_+'(q_n)|_{q_n=l_+, r_+} &\geq 2s_+^4, \quad |\phi_+'(q_n)|_{q_n \in [l_+, l_+ + s_+^2], (r_+^2 - s_+^2, r_+]} > 0. \end{aligned}$$

In fact, it is derived from (3.20) that

$$\max_{\Sigma_1} |\rho_+'(q_n) - \rho'(q_n)| \leq \frac{16n\delta}{\lambda(s^4 - 3s_+^4)s^2}.$$

Since $\phi''(q_n)$ has another expression

$$\phi''(q_n) = \frac{\partial^2 N}{\partial q_n^2} + \sum_{i=1}^n \frac{\partial^2 N}{\partial p_i \partial q_n} \rho_i', \tag{3.22}$$

consequently we obtain

$$\begin{aligned} \max_{\Sigma_1} |\phi''(q_n) - \phi''_+(q_n)| &\leq s_+^2 s^4, \\ \max_{\Sigma_1} |\phi'(q_n) - \phi'_+(q_n)| &\leq s_+^2 s^6, \end{aligned}$$

which completes the check of all of (3.a3) if we take (3.3), (3.10) and (3.21) into account.

As for (3.a1) for P_+ in Σ_{s_+} we shall not deal with it here; it is closely connected to the proof of convergence, so it will be explained in the next section. We also leave (3.a2) and (3.a4) to be verified in the next section. So far, one step of iteration in the case of weaker persistency is completed.

3b. Stronger Persistency. In this case

$$\left| \left(\frac{\partial^2 N}{\partial q_n^2} - \sum_{i,j=1}^n a_{ij} \frac{\partial^2 N}{\partial p_i \partial q_n} \frac{\partial^2 N}{\partial p_j \partial q_n} \right)_{p=p^*, q_n=q_n^*} \right| \geq s_+^2. \tag{3.23}$$

Under such conditions, the critical point of $(N + N_+)(p, q_n) - \langle \omega, \hat{p} \rangle$ determined by the same minimax principle remains close to that of $N(p, q_n) - \langle \omega, \hat{p} \rangle$. Indeed we have

$$|q_n^* - q_{n+}^*| \leq s^5, \tag{3.24}$$

which guarantees that

$$|P_n^* - P_{n+}^*| < s^5. \tag{3.25}$$

Equation (3.24) follows from the argument:

$$\begin{aligned} \phi(q_n) - \delta &\leq \phi_+(q_n) \leq \phi(q_n) + \delta, \\ \phi(q_n^*) - \phi(q_n) &\geq \frac{1}{2} |\phi''(q_n^*)| |q_n - q_n^*|^2 - \frac{1}{6} \max \phi^{(3)}(q_n) |q_n - q_n^*|^3, \end{aligned}$$

so for $s^5 \leq |q_n - q_n^*| \leq 2s_+^3$,

$$\phi_+(q_n^*) - \phi_+(q_n) \geq \frac{1}{2} \left(s_+^2 - \frac{1}{192} |q_n - q_n^*| \right) |q_n - q_n^*|^2 - 2\delta > 0,$$

it then verifies (3.24) if we take (3.3) into account, which also shows that ϕ_+ can not reach its maximum in the interval $[q_n^* - 2s_+^3, q_n^* - s^5]$ and $[q_n^* + s^5, q_n^* + 2s_+^3]$.

To go on with the demonstration, we also assume (3.1)(3.3)(3.a2) and (3.a4) hold. Actually they have been used in deriving (3.24).

The condition of stronger persistency enables us to find nested domains in an explicit way and a related symplectic transformation of coordinates (2.5) with the generating function (2.10).

Instead of considering the Hamiltonian $N + P$ in $\mathcal{A}_t \times \Sigma_s$, we only suppose it is real analytic in the domain $\mathcal{A}_t \times \mathcal{D}_s$,

$$\mathcal{D}_s = \{ |\hat{p} - \hat{\rho}(q_n)| \leq s^4 \} \times \{ |p_n - p_n^*| \leq s^3 \} \times \{ |q_n - q_n^*| \leq s^3 \},$$

$\mathcal{D}_{s_+} \subseteq \Sigma_o$. We deal with it in this way because there is another possibility that in the last step of the iteration the corresponding critical point has stronger persistency

already. In order to obtain the corresponding estimate, the Hamiltonian $N + P$ is only defined in such a domain $\mathcal{A}_t \times \mathcal{D}_s$ rather than in the domain $\mathcal{A}_t \times \Sigma_s$. To see $\mathcal{D}_{s_+} \subseteq \Sigma_o$, we notice that (3.20) makes $|\rho_+(q_n) - \rho(q_n)| \leq \frac{1}{2}s_+^3$ through (3.3), consequently,

$$\{p_n \mid |p_n - p_{n+}^*| \leq s_+^3\} \subseteq \left\{p_n \mid |p - \rho_{n+}(q_n)| \leq \frac{3}{2}s_+^3\right\} \subseteq \{p_n \mid |p - \rho_n(q_n)| \leq 2s_+^3\}$$

for any fixed $q_n \in \{|q_n - q_{n+}^*| \leq s_+^3\}$ since $|\rho'| \leq \frac{1}{2}$ (see (3.6)). By noticing the estimate (3.18) holds in $\mathcal{A}_{t_+} \times \Sigma_o$ we find the switch of iteration from the weaker persistency case to the stronger one does not break down the convergence.

Several intermediate domains also need to be introduced:

$$\mathcal{A}_{t_+} \subset \mathcal{A}_1 \subset \mathcal{A}_2 \subset \mathcal{A}_t, \quad \mathcal{D}_{s_+} \subset \mathcal{D}_0 \subset \mathcal{D}_1 \subset \mathcal{D}_2 \subset \mathcal{D}_s,$$

where

$$\mathcal{A}_{t_+} = \{|\operatorname{Im} q_i| \leq t_+, \operatorname{Re} q_i \in \mathbb{T}, i \leq n - 1\},$$

$$\mathcal{A}_1 = \{|\operatorname{Im} q_i| \leq t_+ + s_+^2, \operatorname{Re} q_i \in \mathbb{T}, i \leq n - 1\},$$

$$\mathcal{A}_2 = \left\{|\operatorname{Im} q_i| \leq t - \frac{1}{2}(t - t_+), \operatorname{Re} q_i \in \mathbb{T}, i \leq n - 1\right\},$$

$$\mathcal{D}_{s_+} = \{|\hat{p} - \hat{p}_+(q_n)| \leq s_+^4, |p_n - p_{n+}^*| \leq s_+^3, |q_n - q_{n+}^*| \leq s_+^3\}$$

$$\mathcal{D}_0 = \{|\hat{p} - \hat{p}(q_n)| \leq 2s_+^4, |p_n - p_n^*| \leq 2s_+^3, |q_n - q_n^*| \leq 2s_+^3\}$$

$$\mathcal{D}_1 = \{|\hat{p} - \hat{p}(q_n)| \leq 3s_+^4, |p_n - p_n^*| \leq 3s_+^3, |q_n - q_n^*| \leq 3s_+^3\}$$

$$\mathcal{D}_2 = \left\{|\hat{p} - \hat{p}(q_n)| \leq s^4 - \frac{1}{2}(s^4 - 3s_+^4), |p_n - p_n^*| \leq s^3 - \frac{1}{2}(s^3 - 3s_+^3),\right.$$

$$\left. |q_n - q_n^*| \leq s^3 - \frac{1}{2}(s^3 - 3s_+^3)\right\}.$$

Supposing $|P| \leq \delta$ in $\mathcal{A}_t \times \mathcal{D}_s$ and expanding P in a Taylor series in (p_n, q_n) , we have the Cauchy estimate in $\{|\operatorname{Im} \hat{q}| \leq t\} \times \{|\hat{p} - \hat{p}(q_n)| \leq s^4\}$,

$$\max \{|P_{ij}(\hat{p}, \hat{q})|\} \leq \frac{(i+j)! \delta}{s^{3(i+j)}}. \tag{3.26}$$

Together with (2.12), (3.26) shows that in $\mathcal{A}_2 \times \mathcal{D}_2$,

$$|W| = \frac{c_5 \delta}{D^6 (t - t_+)^{6\mu}} := \gamma,$$

$$|W_{\hat{q}}| \leq 4\gamma (t - t_+)^{-1},$$

$$|W_{\hat{p}}| \leq 2\gamma (s^4 - 3s_+^4)^{-1},$$

$$|W_{q_n}| \leq 2\gamma (s^3 - 3s_+^3)^{-1},$$

$$|W_{p_n}| \leq 2\gamma (s^3 - 3s_+^3)^{-1},$$

$$\max \{|W_{p_i p_j}|, |W_{p_i q_j}|, |W_{q_i q_j}|\} \leq 8\gamma (s^4 - 3s_+^4)^{-2}, \quad \forall i, j \leq n. \tag{3.27}$$

If (3.3) is assumed we then have

$$\begin{aligned}
 |\hat{q} - \hat{q}_+| &\leq |W_{\hat{p}}| \leq \frac{1}{2}(t - t_+) - s_+^2, \\
 |q_n - q_{n+}| &\leq |W_{p_{n+}}| \leq \frac{1}{2}(s^3 - 3s_+^3), \\
 |\hat{p} - \hat{p}(q_n)| &\leq |(\hat{p} - \hat{p}(q_n)) - (\hat{p}_+ - \hat{p}(q_{n+}))| + |\hat{p}_+ - \hat{p}(q_{n+})| \\
 &\leq |W_{\hat{q}}| + |\hat{p}||W_{p_{n+}}| + 3s_+^4 \leq \frac{1}{2}(s^4 + 3s_+^4), \\
 |p_n - p_n^*| &\leq |W_{q_n}| \leq \frac{1}{2}(s^3 - 3s_+^3),
 \end{aligned} \tag{3.28}$$

which implies \mathcal{M} maps $\mathcal{A}_1 \times \mathcal{D}_1$ into $\mathcal{A}_2 \times \mathcal{D}_2$. By the argument similar to that in the case of weaker persistency, which we omit here because it could be done by following the footsteps exactly as we do when we deal with the case of weaker persistency, $\mathcal{M} : (p_+, q_+) \rightarrow (p, q)$ maps $\mathcal{A}_1 \times \mathcal{D}_1$ into $\mathcal{A}_2 \times \mathcal{D}_2$, and it is obvious in $\mathcal{A}_1 \times \mathcal{D}_1$,

$$|\mathcal{M} - \text{id}| \leq 4\gamma(s^4 - 3s_+^4)^{-1}. \tag{3.29}$$

Thanks to Lemma 2 and Lemma 3 we also have

$$\left| \frac{\partial(p, q)}{\partial(p_+, q_+)} - I \right| \leq 80\gamma(s^4 - 3s_+^4)^{-2}. \tag{3.30}$$

Now the aim is to show how large the new error term P_+ is in the domain $\mathcal{A}_1 \times \mathcal{D}_1$. It is induced from (2.13) and (2.14) that

$$\begin{aligned}
 |M_{11}| &\leq \max |\phi''| |W_{p_{n+}}|^2 + \max |\phi^{(3)}| |q_n - q_n^*|^2 |W_{p_{n+}}| \\
 &\leq 4\gamma^2(s^3 - 3s_+^3)^{-2} + 18\gamma s_+^6 (s^3 - 3s_+^3)^{-1}, \\
 |M_{12}| &\leq 6n^2 \beta \gamma s_+^4 (s^3 - 3s_+^3)^{-1} + 9n^3 \beta s_+^6 (s^3 - 3s_+^3)^{-1} + 2n^2 \beta \gamma^2 (s^3 - 3s_+^3)^{-2}, \\
 |M_{13}| &\leq 6(n - 1) \beta \gamma s_+^4 (s^3 - 3s_+^3)^{-1} + 2\beta \gamma^2 (s^3 - 3s_+^3)^{-2} + 9n^2 \beta \gamma s_+^6 (s^3 - 3s_+^3)^{-1}, \\
 |M_{14}| &\leq 18(n + 2) \beta \gamma s_+^6 (s^3 - 3s_+^3)^{-1}, \\
 |M_{15}| &\leq 18(n + 2) \beta \gamma s_+^6 (s^3 - 3s_+^3)^{-1}, \\
 |M_2| &\leq c_6 \delta \left(\frac{s_+}{s} \right)^{15}, \\
 |M_3| &\leq 4\delta \gamma (s^3 + 3s_+^3)^{-1} (s^3 - 3s_+^3)^{-1}, \\
 |M_4| &\leq 8\delta \gamma (s^3 + 3s_+^3)^{-1} (s^3 - 3s_+^3)^{-1}.
 \end{aligned}$$

Putting these terms together we finally obtain:

$$|P_+| \leq c_7 \left\{ \gamma s_+^4 (s^3 - 3s_+^3)^{-1} + \delta \left(\frac{s_+}{s} \right)^{15} \right\} := \delta_+,$$

c_7 depends on β . Shrinking $\mathcal{A}_1 \times \mathcal{D}_1$ to $\mathcal{A}_{t_+} \times \mathcal{D}_0$ then

$$\begin{aligned}
 & |P_+(p_+, q_+)| \leq \delta_+, \\
 & \max \left\{ \left| \frac{\partial P_+}{\partial p_+} \right|, \left| \frac{\partial P_+}{\partial q_+} \right| \right\} \leq \frac{\delta_+}{s_+^4}, \\
 & \max_{|l_1+l_2|=2} \left\{ \left| \frac{\partial^2 P_+}{\partial p_+^{l_1} \partial q_+^{l_2}} \right| \right\} \leq \frac{2\delta_+}{s_+^5}, \\
 & \max_{|l_1+l_2|=3} \left\{ \left| \frac{\partial^3 P_+}{\partial p_+^{l_1} \partial q_+^{l_2}} \right| \right\} \leq \frac{6\delta_+}{s_+^{12}}. \tag{3.31}
 \end{aligned}$$

The same as in the case of weaker persistency we need to proceed to the next step of iteration in the neighbourhood of the critical point line $p = \rho_+(q_n)$ of the function $(N + N_+)(p, q_n) - \langle \omega, \hat{p} \rangle$. Equations (3.20), (3.24) and (3.25) show that $\mathcal{D}_0 \subseteq \mathcal{D}_1$, because of the choice of relevant parameters given by (3.1), (3.3) and (3.6).

In the demonstration of one cycle of iteration, we take (3.a2) and (3.a4) as granted. We also leave them unproved until the next section, it will simplify the notation a lot.

To complete one step of iteration we need to show that once the critical point gets stronger persistency, the forthcoming critical point shall possess stronger persistency as well. In fact, it is easy to see (3.20) also holds in \mathcal{D}_2 , consequently,

$$\max_{\mathcal{D}_1} |\rho'_+(q_n) - \rho'(q_n)| \leq \frac{8n\delta}{\lambda(s^4 - 3s_+^4)(s^3 - 3s_+^3)},$$

hence from (3.22)

$$\max_{\mathcal{D}_1} |\phi'_+(q_n) - \phi''(q_n)| \leq s_+^2 s^3.$$

Setting $s_{++} \leq \frac{1}{2}s_+$ we finally have

$$\begin{aligned}
 \phi''_+(q_{n++}^*) & \leq \phi''(q_n^*) + \max |\phi^{(3)}| |q_n^* - q_{n+}^*| + |\phi''_+(q_n^*) - \phi''(q_n^*)| \\
 & \leq -s_+^2 + 2s^5 < -s_{++}^2.
 \end{aligned}$$

4. Convergence

With the preliminary work done in the last two sections, we are now at the stage to complete the last part of our proof. As assumed in the first section, the Hamiltonian

$$H_o(p_o, q_o) = N_o(p_o) + P_o(p_o, q_o)$$

is real analytic in the region

$$\Sigma_o = \{ |\text{Im } q_o| \leq \tau_* \} \times \{ |p_o - p_o^*| \leq s_* \}.$$

$N_o(p_o)$ is convex in $\text{Re } \Sigma_o$ and

$$\min_{\Sigma_o} \left\| \frac{\partial^2 N_o}{\partial p^2} \xi \right\| \geq \lambda_o \|\xi\|, \quad (\forall \xi \in \mathbb{C}^n)$$

$$\sup_{|l|_1 \leq 3, \Sigma_o} \left| \frac{\partial^l N_o}{\partial p^l} \right| \leq \beta_o.$$

Let $\tau_* = \tau_o, s_* = s_o^3$ and assume

$$\max_{\Sigma_o} |P_o(p_o, q_o)| \leq \delta_o.$$

Using the results obtained in the last section, we construct inductively a sequence of symplectic transformations of coordinates, \mathcal{M}_m on domains Σ_m whose range is contained in Σ_{m-1} ; i.e.,

$$\mathcal{M}_m : \Sigma_m \rightarrow \Sigma_{m-1}.$$

In Σ_m the Hamiltonian H_o is transformed into

$$H_m(p_m, q_m) = N_m(p_m, q_{nm}) + P_m(p_m, q_m), \tag{4.1}$$

where $N(p_m, q_{nm})$ is convex in p_m for any fixed q_{nm} . The critical point line $\rho_m(q_{nm})$ is defined by looking for the minimum of the function $N_m(p_m, q_{nm}) - \langle \omega, \hat{p}_m \rangle$ and relevant critical point is $\rho_m(q_{nm})$ for any fixed q_{nm} in the real domain. The function $\phi_m(q_{nm}) = N_m(\rho_m(q_{nm}), q_{nm}) - \langle \omega, \hat{p}_m(q_{nm}) \rangle$ attains its maximum at the point q_{nm}^* . Let $p_m^* = \rho_m(q_{nm}^*)$. The domain Σ_m is described in two ways. In the case of weaker persistency,

$$\Sigma_m = \{ |\text{Im } \hat{q}_m| \leq \tau_m, |\hat{p}_m - \hat{\rho}_m(q_{nm})| \leq s_m^4, \\ |p_{nm} - \rho_{nm}(q_{nm})| \leq s_m^3, |\text{Im } q_{nm}| \leq s_m^2, \text{Re } q_{nm} \in L_m \},$$

and in the case of stronger persistency,

$$\Sigma_m = \{ |\text{Im } \hat{q}_m| \leq \tau_m, |\hat{p}_m - \hat{\rho}_m(q_{nm})| \leq s_m^4, \\ |p_{nm} - p_{nm}^*| \leq s_m^3, |q_{nm} - q_{nm}^*| \leq s_m^3 \},$$

where $L_m = [l_m, r_m] \supset [q_{nm}^* - s_m^2, q_{nm}^* + s_m^2]$,

(a) $\tau_m = \frac{\tau_o}{2}(1 + 2^{-m})$;

(b) $s_m = s_{m-1}^{\frac{14}{13}}, s_o \leq \tau_o$;

(c) $\delta_m = s_m^{13}$;

(d) The modification of the main part of the Hamiltonian after each step of iteration is small:

$$|(N_m - N_{m+1})| \leq 2\delta_m; \quad (\text{in } \Sigma_{m+1})$$

(e)

$$|P_m| \leq \delta_m,$$

$$\max_{\Sigma|l_j|=1} \left\{ \frac{\partial P_m}{\partial p_m^{l_1} \partial q_m^{l_2}} \right\} \leq \delta_m^{\frac{9}{13}},$$

$$\max_{\Sigma|l_j|=2} \left\{ \frac{\partial^2 P_m}{\partial p_m^{l_1} \partial q_m^{l_2}} \right\} \leq 2\delta_m^{\frac{5}{13}},$$

$$\max_{\Sigma|l_j|=3} \left\{ \frac{\partial^3 P_m}{\partial p_m^{l_1} \partial q_m^{l_2}} \right\} \leq 6\delta_m^{\frac{1}{13}}; \quad (\text{in } \Sigma_m)$$

(f)

$$|\mathcal{M}_m - \text{id}| \leq \delta_m^{\frac{7}{13}}; \quad (\text{in } \Sigma_m)$$

(g)

$$|\mathcal{M}'_m - \text{Id}| \leq \delta_m^{\frac{3}{13}}. \quad (\text{in } \Sigma_m)$$

\mathcal{M}'_m is the Jacobian of \mathcal{M}_m . At the beginning of the iteration we have to deal with the problem in the way of weaker persistency because $N_o = N_o(p_o)$ is independent of q_{no} . Such a process may be repeated for finitely or infinitely many times, depending on the fact whether the critical point at each step has weaker persistency or not. Once it has stronger persistency at some step of the iteration, then it always possesses stronger persistency at the following steps, which can be seen from the proof in the last section.

Such a construction is possible if all conditions introduced in the last section are satisfied for each m which we need to verify.

It follows from (e) in Σ_m

$$\left| \frac{\partial^2 \bar{P}_m}{\partial p_m^2} \right| \leq 2\delta_m^{\frac{5}{13}}, \quad \left| \frac{\partial^3 \bar{P}_m}{\partial p_m^3} \right| \leq 6\delta_m^{\frac{1}{13}},$$

where \bar{P}_m is the average part of P_m with respect to q_m

$$N_m(p_m, q_{nm}) = \sum_{v=1}^{m-1} \bar{P}_v(p_m, q_{nm}) + N_o(p_m),$$

$$\min_{\Sigma_m} \left\| \frac{\partial^2 N_m \xi}{\partial p^2} \right\| \geq \left(\lambda_o - 2 \sum_{v=1}^{m-1} \delta_v^{\frac{5}{13}} \right) \|\xi\|. \quad (\forall \xi \in \mathbb{C}^n)$$

Since N_o is independent of q_{no} we also have

$$\sup_{|l| \leq 3, l_2 \geq 1} \left| \frac{\partial^l N_m}{\partial p_m^{l_1} \partial q_{nm}^{l_2}} \right| \leq 6 \sum_{v=1}^{m-1} \delta_v^{\frac{1}{13}}.$$

Let $\lambda = \lambda_o - 2 \sum_{v=1}^{\infty} \delta_v^{\frac{5}{13}}$ and $\beta = \beta_o + 6 \sum_{v=1}^{\infty} \delta_v^{\frac{1}{13}}$ then there is a $d_1(N_o, n) > 0$ such that if $\delta_o \leq d_1$ then $\lambda \geq \frac{\lambda_o}{2}, \beta \leq 2\beta_o$ and $6 \sum_{v=1}^{m-1} \delta_v^{\frac{1}{13}} \leq (7n^2 2^\mu)^{-1}$. Let the λ and β appearing in the proof in the last section, take the value defined here then they satisfy the required conditions. To complete the verification of (3.a4) we

get in Σ_m by noticing (3.20)

$$\begin{aligned} |\rho_m(q_n) - \overline{\rho_{m+1}}(q_n)| &\leq c_8 s_m^9, \\ |\rho'_m(q_n) - \rho'_{m+1}(q_n)| &\leq c_8 s_m^6, \\ |\rho''_m(q_n) - \rho''_{m+1}(q_n)| &\leq c_8 s_m^3, \end{aligned}$$

and

$$\begin{aligned} |\phi'_m(q_n) - \phi'_{m+1}(q_n)| &\leq c_8 s_m^7, \\ |\phi''_m(q_n) - \phi''_{m+1}(q_n)| &\leq c_8 s_m^4, \\ |\phi_m^{(3)}(q_n) - \phi_{m+1}^{(3)}(q_n)| &\leq c_8 s_m, \end{aligned}$$

c_8 depends on λ . Since

$$|\rho_m^{(j)}| \leq \sum_{v=1}^m |\rho_v^{(j)} - \rho_{v-1}^{(j)}|, \quad |\phi_m^{(j)}| \leq \sum_{v=1}^m |\phi_v^{(j)} - \phi_{v-1}^{(j)}|,$$

because of $\rho_o^{(j)} = \phi_o^{(j)} = 0$, we can find $d_2(N_o, n) > 0$, if $\delta \leq d_2$ further then

$$\max_{\Sigma_{m,j \leq 3}} |\phi_m^{(j)}(q_n)| \leq \frac{1}{64}, \quad \max_{\Sigma_{m,j=1,2}} |\rho_m^{(j)}(q_n)| \leq \frac{1}{2}.$$

Up to now we have verified (3.a2) and (3.a4).

Let

$$d_3 = \min \left\{ \left(\frac{65D\tau_o^\mu}{c_1 8^{\mu+5} 2^{2\mu}} \right)^{\frac{13}{6}}, \left(\frac{65D^6\tau_o^{6\mu}}{2^{12\mu+10} c_5} \right)^{\frac{13}{6}}, \left(\frac{\lambda}{12n} \right)^{\frac{169}{11}}, 2^{-169} \right\},$$

since $\delta_m \leq \delta_o^{1+\frac{m}{13}}$, it can be seen that if $\delta_o \leq d_3$, $s_o = \delta_o^{\frac{1}{13}}$, then

$$\begin{aligned} 2s_{m+1} &\leq s_m, \quad s_m \leq \tau_m \leq 1, \quad s_{m+1}^4 = s_m^{\frac{56}{13}} \geq s_m^5, \\ s_{m+1} &\leq \tau_{m+1}, \quad s_{m+1} \leq \frac{1}{4}(\tau_m - \tau_{m+1}), \end{aligned}$$

and

$$\delta_m \leq \min \left\{ \frac{65D\tau_o^\mu}{c_1 8^{\mu+5} 2^{\mu(m+2)}} s_m^6, \frac{65D^6\tau_o^{6\mu}}{2^{6\mu(m+2)+10} c_5} s_m^6, \frac{\lambda}{12n} s_m^{10} s_{m+1}^2 \right\},$$

which verifies (3.1) and (3.3). Clearly $\exists d_4 > 0$, if $\delta_o \leq d_4$, $s_o |\log s_o|^\mu \leq 1$. Let

$$d_5 = \min \left\{ d_4, \frac{D}{2} \left(\frac{3\tau_o}{64(n-1)} \right)^{13\mu}, 2^{-169\mu}, \left(\left(\frac{e}{2n-2} \right)^{n-1} \left(\frac{\tau_o}{32} \right)^n \right)^{\frac{13}{2}} \right\},$$

(3.2) is also satisfied if $\delta_o \leq d_5$.

The most important thing we need to show is $|P_m| \leq \delta_m$ in Σ_m . When the case of weaker persistency is dealt with, we have from (3.18),

$$|P_{m+1}| \leq c_4 \left\{ \frac{2^{\mu(m+2)+4} 8^\mu c_1}{D \tau_o^\mu} s_m^{\frac{3}{13}} + s_m^{\frac{2}{13}} \right\} \delta_{m+1} ;$$

when the case of stronger persistency is dealt with, we have from (3.32),

$$|P_{m+1}| \leq c_7 \left\{ \frac{2^{6\mu(m+2)+3} c_5}{5D^6 \tau_o^{6\mu}} s_m^{\frac{4}{13}} + s_m^{\frac{2}{13}} \right\} \delta_{m+1} .$$

By noticing that s_m decreases to zero faster than exponentially, we see that there is a positive number $d_6(N_o, D, \mu, \tau_o, s_o) > 0$, such that if $\delta_o \leq d_6$ then

$$|P_{m+1}| \leq \delta_{m+1} .$$

The remaining two estimates in (e) follow directly from (3.19) and (3.31). From (3.17), (3.18) we have

$$|\mathcal{M}_m - \text{id}| \leq \frac{2^{\mu(m+2)} 8^{\mu+4} c_1}{13D \tau_o^\mu} s_m^2 \delta_m^{\frac{7}{13}} ,$$

$$|\mathcal{M}'_m - \text{Id}| \leq \frac{2^{\mu(m+2)} 8^{\mu+7} 10c_1}{169D \tau_o^\mu} s_m^2 \delta_m^{\frac{3}{13}} ,$$

similarly from (3.29) and (3.30) we have

$$|\mathcal{M}_m - \text{id}| \leq \frac{2^{6\mu(m+2)+6} c_5}{13D^6 \tau_o^{6\mu}} s_m^2 \delta_m^{\frac{7}{13}} ,$$

$$|\mathcal{M}'_m - \text{Id}| \leq \frac{2^{6\mu(m+2)+9} c_5}{D^6 \tau_o^{6\mu}} s_m^2 \delta_m^{\frac{3}{13}} ,$$

which results in (f) and (g) by the choice of δ_o . So far we have shown how to construct the sequence of transformations \mathcal{M}_m inductively and therefore assertions (a)~(g) are valid for all $m \geq 0$ if $\delta_o \leq \min\{d_j, j \leq 6\}$. As $m \rightarrow \infty$ the domain Σ_m shrink to a set containing the submanifold

$$\Sigma_\infty \supseteq \left\{ |\text{Im } \hat{q}_\infty| \leq \frac{\tau_o}{2} \right\} \times \{q_{n\infty} = q_\infty^*, p_\infty = \rho_\infty(q_\infty^*)\}$$

and the transformations

$$\mathcal{T}_m = \mathcal{M}_1 \circ \mathcal{M}_2 \circ \dots \circ \mathcal{M}_m : \Sigma_m \rightarrow \Sigma_o$$

with their derivatives

$$\mathcal{T}'_m = \mathcal{M}'_1 \circ \mathcal{M}'_2 \circ \dots \circ \mathcal{M}'_m : \Sigma_m \rightarrow \Sigma_o$$

converge uniformly to an embedding

$$\mathcal{T}_\infty : \Sigma_\infty \rightarrow \Sigma_o .$$

The uniform convergence of \mathcal{T}_m and \mathcal{T}'_m follows from the assertions (f) and (g) and the rapid convergence of the sequence $\{\delta_m\}$ to zero. Clearly the embedding \mathcal{T}_∞ is real analytic in Σ_∞ since it is the uniform limit of real analytic functions.

The next step is to show that

$$\mathcal{T}_\infty(\omega t + \hat{q}_\infty, q_{n\infty}^*, \rho_\infty(q_{n\infty}^*)) = \Phi'_o(\mathcal{T}_\infty(\hat{q}_\infty, q_{n\infty}^*, \rho_\infty(q_{n\infty}^*))), \quad (4.2)$$

where Φ'_o is the flow determined by the vector field V_o afforded by Hamiltonian $H_o, \rho_\infty = \lim_{m \rightarrow \infty} \rho_m$, ((3.20) implies this limit exists). If at any step of iteration, the related critical point always has weaker persistency, let $L_\infty = \bigcap_{m=1}^\infty L_m$, the facts that $L_m \supseteq L_{m+1}$ and L_m is closed for all $m \geq 0$ guarantee that $L_\infty \neq \{\emptyset\}$. In this case we let $q_{n\infty}^*$ be any point in L_∞ . If after some steps of iteration the related critical point has stronger persistency, due to (3.24), $q_{n\infty}^* = \lim_{m \rightarrow \infty} q_{nm}^*$. The proof is essentially the same as those for classical KAM theory (cf. [A2, G]). For completeness, we demonstrate the proof briefly by invoking some lemma developed in [G].

To abbreviate the notation, let us write the canonical equations with the Hamiltonian H_m in the form

$$\dot{w}_m = \Psi_m(w_m), \quad w_m = (x_m, y_m, u_m, v_m) \in \Sigma_m .$$

Since the transformation $\mathcal{T}_m : \Sigma_m \rightarrow \Sigma_o$ is symplectic, if $w_m(t)$ is a solution of the last equation, $w_0(t) = \mathcal{T}_m(w_m(t))$ satisfies $\dot{w}_0 = \Psi_0(w_0)$. By letting $\Psi_\infty(w_\infty) = (\omega, 0, 0, 0)$ and observing from (d) and (e) ((e')) we readily verify:

$$|\Psi_m - \Psi_\infty| \leq \text{distance}(\Sigma_{m+1}, \partial\Sigma_m) = \eta_{m+1} ;$$

$$\left| \frac{\partial \Psi_m}{\partial \omega_m} \right| \leq \text{const} .$$

It is derived from Lemma 5 that $\mathcal{T}_\infty(\omega t + \hat{q}_\infty, q_{n\infty}^*, \rho_\infty(q_{n\infty}^*))$ is a solution of $\dot{w}_0 = \Psi_0(w_0)$, which exactly is what (4.2) means.

Lemma 5 [G]. *Let $\Psi_o(w)$ be a smooth vector field on Σ_o determining a flow $\Phi'_o(w)$:*

$$\frac{d}{dt} \Phi'_o(w) = \Psi_o(\Phi'_o(w)), \quad \Phi'_o(w) = w .$$

Suppose there exist invertible transformations $\mathcal{M}_m : \Sigma_m \rightarrow \Sigma_o$ such that $|\prod_{m=1}^\infty \mathcal{M}'_m| < \infty$, where \mathcal{M}'_m denotes the Jacobian matrix of \mathcal{M}_m . The transformations

$$\mathcal{T}_m = \mathcal{M}_o \circ \mathcal{M}_2 \circ \dots \circ \mathcal{M}_m : \Sigma_m \rightarrow \Sigma_o$$

naturally induce flows

$$\Phi'_m = \mathcal{T}_m^{-1} \Phi'_o \mathcal{T}_m$$

with corresponding vector field Ψ_m on Σ_m ,

$$\Psi_m(w) = \left. \frac{d}{dt} \Phi'_m(w) \right|_{t=0} .$$

Assume

(1) Ψ_m converges to Ψ_∞ on Σ_∞ as $m \rightarrow \infty$ and $|\Psi_m - \Psi_\infty| \leq C\eta_m$ on Σ_∞ with some constant C independent of m and $\eta_m \rightarrow 0$ as $m \rightarrow \infty$;

(2) the segment $w = w_o + \sigma t, 0 \leq t \leq \frac{1}{B+C}$ belongs to Σ_∞ and on this segment $\Psi_\infty = \sigma$:

(3) $\left| \frac{\partial \Psi_m}{\partial w} \right| \leq B$ on Σ_m , where B is independent of m ; and

(4) $\mathcal{T}_\infty = \lim_{m \rightarrow \infty} \mathcal{T}_m$ exists and is continuous.

Then for $0 \leq t \leq \frac{1}{B+C}$, $\Phi'_o(\mathcal{T}_\infty(w_0)) = \mathcal{T}_\infty(w_o + \sigma t) \subset \Sigma_o$.

The last step is to check that the surviving invariant $(n - 1)$ -torus remains in a ε -neighbourhood of some $(n - 1)$ -torus (p_o, q_{no}) trivially embedded in the n -torus p_o . In fact, if we let

$$\begin{aligned} \hat{q}_o &= \hat{q}_\infty + \hat{Q}_\infty(\hat{q}_\infty, q_{n\infty}^*, \rho_\infty(q_{n\infty}^*)); \\ q_{no} &= q_{n\infty} + Q_{n\infty}(\hat{q}_\infty, q_{n\infty}^*, \rho_\infty(q_{n\infty}^*)); \\ p_o &= \rho_\infty(q_{n\infty}^*) + A_\infty(\hat{q}_\infty, q_{n\infty}^*, \rho_\infty(q_{n\infty}^*)); \end{aligned} \tag{4.3}$$

then from the estimate (f) we have

$$\max_{\Sigma_\infty} \{|A_\infty|, |Q_\infty|\} \leq \sum_{i=0}^\infty \delta_i^{\frac{7}{13}},$$

if $\delta_o \leq \min\{d_j, j \leq 6\}$. Given any $\varepsilon_1 > 0$, $d_7 = d_7(\varepsilon_1) > 0$ can be found such that if $\delta \leq d^*(\varepsilon_1, N, D, \mu, s, \tau) = \min\{d_j, j \leq 7\}$

$$\max_{\Sigma_\infty} \{|X_\infty|, |Y_\infty|, |U_\infty|, |V_\infty|\} \leq \varepsilon_1.$$

Recalling a linear symplectic transformation is introduced through the unimodular matrix I at the beginning of Sect. 2 and taking the inverse of such transformation we then obtain (1.9) plus the estimate (1.10) if we let $\varepsilon_1 = |\lambda_M(I)^{-1}| \varepsilon$, where $\lambda_M(I)$ is the largest eigenvalue of the matrix I in terms of the absolute value.

Acknowledgement. This work was done when the author visited the Mathematisches Institut der Universität zu Köln, supported by Alexander von Humboldt Foundation of Germany, and the Forschungsinstitut für Mathematik at ETH Zürich for their hospitality, he is very grateful. He would like to thank T. Küpper, J. Moser, J. Pöschel and H. Rüssmann for helpful discussions and comments. He also thanks the referee for his comments. The referee and J. Pöschel pointed out a crucial error in the original manuscript, which is important for the completion of the second version of this paper.

After the author finished the second version of this paper, he was informed by D.V. Treshchev the results in [T], in which the existence of some lower dimensional tori is proved under some more conditions on the perturbation P , such that one can consider the case of stronger persistency only and d , the scale of perturbation, must depend on the extra condition on P , i.e. the eigenvalue of the Hessian of the average of P over $\mathbb{T}^{(n-1)}$.

References

[A1] Arnol'd, V.I.: "Mathematical Methods of Classical Mechanics." Berlin, Heidelberg, New York: Springer, 1978
 [A2] Arnol'd, V.I.: Proof of A.N. Kolmogorov's theorem on the preservation of quasi periodic motions under small perturbations of Hamiltonian. Usp. Math. USSR **18**, 9–36 (1963)
 [Au] Aubry, S.: Theory of the devil's staircase. Lect. Notes Math. **925** (1978–79)
 [BK] Bernstein, D., Katok, A.: Birkhoff periodic orbits for small perturbations of completely integrable Hamiltonian systems with convex Hamiltonians. Invent. Math. **88**, 225–241 (1987)
 [Ch] Chierchia, L.: On the stability problem for nearly-integrable Hamiltonian systems. Seminar on Dynamical Systems (1994). Basel, Boston: Birkhäuser, pp. 35–46

- [CZ] Conley, C.C., Zehnder, E.: The Birkhoff–Lewis fixed point theorem and a conjecture of V.I. Arnol'd. *Invent. Math.* **73**, 33–49 (1983)
- [E] Eliasson, L.H.: Perturbations of stable invariant tori for Hamiltonian systems. *Ann. Sc. Norm. Super. Pisa. Cl. Sci. IV. Ser.* **15**, 115–147 (1988)
- [G] Graff, S.M.: On the continuation of hyperbolic invariant tori for Hamiltonian systems. *J. Diff. Eqns.* **15**, 1–69 (1974)
- [H1] Herman, M.R.: Sur les courbes invariantes par les difféomorphismes de l'anneau. *Astérisque* **103** (1983)
- [H2] Herman, M.R.: Sur les courbes invariantes par les difféomorphismes de l'anneau. *Astérisque* **144** (1986)
- [K] Kolmogorov, A.N.: On quasi-periodic motions under small perturbations of the Hamiltonian. *Doklady Akad. Nauk. USSR* **98**, 1–20 (1954)
- [L] Lochak, P.: Canonical perturbation theory via simultaneous approximation. *Uspekhi Mat. Nauk* **47**, 59–140 (1992)
- [Ma1] Mather, J.N.: Existence of quasi-periodic orbits for twist homeomorphisms of the annulus. *Topology* **21**, 457–467 (1982)
- [Ma2] Mather, J.N.: Action minimizing invariant measures for positive definite Lagrangian systems. *Math. Z.* **207**, 169–207 (1991)
- [Mo1] Moser, J.: On invariant curves of area preserving mappings of an annulus. *Nachr. Akad. Wiss. Gött. Math-Phys* **K1**, 1–20 (1962)
- [Mo2] Moser, J.: Convergent series expansions for quasi-periodic motions. *Math. Ann.* **169**, 136–176 (1967)
- [MP] Moser, J., Pöschel, J.: An extension of a result by Dinaburg and Sinai on quasi-periodic potentials. *Comment. Math. Helvetici* **59**, 39–85 (1984)
- [P] Pöschel, J.: On elliptic lower dimensional tori in Hamiltonian systems. *Math. Z.* **202**, 559–608 (1989)
- [Rob] Robinson, C.: Generic properties of conservative systems. *Am. J. Math.* **92**, 562–603, 897–906 (1970)
- [Rüs1] Rüssmann, H.: On the existence of invariant curves of twist mappings of an annulus. *Geometric Dynamics. Lect. Notes Math.* **1007**, 677–718 (1983)
- [Rüs2] Rüssmann, H.: On twist Hamiltonians. Talk on the Colloque International: Mécanique céleste et systèmes Hamiltonian, Marseille
- [SM] Siegel, C.L., Moser, J.: *Lectures on Celestial Mechanics*. Berlin, Heidelberg, New York: Springer, 1971
- [T] Treshchev, D.V.: The mechanism of destruction of resonance tori of Hamiltonian systems. *Math. USSR Sbornik* **68**, 181–203 (1991)
- [W] Wayne, C.E.: Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory. *Commun. Math. Phys.* **127**, 479–528 (1990)

Communicated by M. Herman

