Persistence of Invariant Tori in Integrable Hamiltonian Systems Under Almost Periodic Perturbations

Peng Huang 1 · Xiong Li 1

Received: 13 November 2017 / Accepted: 14 May 2018 / Published online: 15 June 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract In this paper, we are concerned with the existence of invariant tori in nearly integrable Hamiltonian systems

\[ H = h(y) + f(x, y, t), \]

where \( y \in D \subseteq \mathbb{R}^n \) with \( D \) being an open bounded domain, \( x \in \mathbb{T}^n \), \( f(x, y, t) \) is a real analytic almost periodic function in \( t \) with the frequency \( \omega = (\ldots, \omega_\lambda, \ldots)_{\lambda \in \mathbb{Z}} \in \mathbb{R}^Z \). As an application, we will prove the existence of almost periodic solutions and the boundedness of all solutions for the second-order differential equations with superquadratic potentials depending almost periodically on time.

Keywords Invariant tori · Hamiltonian systems · Almost periodic solutions · Boundedness · Superquadratic potentials

Mathematics Subject Classification 70H08 · 70H12 · 70K43 · 34C11

Communicated by Changpin Li.

Xiong Li: Partially supported by the NSFC (11571041) and the Fundamental Research Funds for the Central Universities.

✉ Xiong Li
xli@bnu.edu.cn
Peng Huang
hp@mail.bnu.edu.cn

1 School of Mathematics Sciences, Beijing Normal University, Beijing 100875, People’s Republic of China
1 Introduction

In this paper, we study the existence of invariant tori in the nearly integrable Hamiltonian system

\[ H = h(y) + f(x, y, t), \quad (1.1) \]

where \( y \in D \subseteq \mathbb{R}^n \) with \( D \) being an open bounded domain, \( x \in \mathbb{T}^n \), \( f(x, y, t) \) is a real analytic almost periodic function in \( t \) with the frequency \( \omega = (\ldots, \omega_\lambda, \ldots)_{\lambda \in \mathbb{Z}} \) and admits a spatial series expansion

\[ f(x, y, t) = \sum_{k \in \mathbb{Z}_S^Z} f_k(x, y) e^{i(k, \omega)t}. \quad (1.2) \]

Here \( \omega = (\ldots, \omega_\lambda, \ldots)_{\lambda \in \mathbb{Z}} \) is a bilateral infinite sequence of rationally independent frequency, that is to say, any finite segments of \( \omega \) are rationally independent, \( S \) is a family of finite subsets \( A \) of \( \mathbb{Z} \), which represents a spatial structure on \( \mathbb{Z} \) with \( \mathbb{Z} = \bigcup_{A \in S} A \) depended on the Fourier exponents \( \{ \Omega_\lambda : \lambda \in \mathbb{Z} \} \) of a kind of real analytic almost periodic functions similar to the almost periodic perturbation \( f \), and its basis is the frequency \( \{ \omega_\lambda : \lambda \in \mathbb{Z} \} \) which is contained in \( \{ \Omega_\lambda : \lambda \in \mathbb{Z} \} \), \( \mathbb{Z}_S^Z \) is the space of bilateral infinite integer sequences \( k = (\ldots, k_\lambda, \ldots)_{\lambda \in \mathbb{Z}} \) whose support \( \text{supp} k = \{ \lambda : k_\lambda \neq 0 \} \) is a finite set of \( \mathbb{Z} \) contained in a subset \( A \) that belongs to \( S \), \( \langle k, \omega \rangle = \sum_{\lambda \in \mathbb{Z}} k_\lambda \omega_\lambda \).

Kolmogorov–Arnold–Moser (or KAM) theory is a powerful method about the persistence of quasi-periodic solutions and almost periodic solutions under small perturbations. KAM theory is not only a collection of specific theorems, but rather a methodology, a collection of ideas of how to approach certain problems in perturbation theory connected with small divisors.

The classical KAM theory was developed for the stability of motions in Hamiltonian systems that are small perturbations of integrable Hamiltonian systems. Integrable systems in their phase space contain lots of invariant tori, and the classical KAM theory establishes persistence of such tori, which carry quasi-periodic motions. The classical KAM theory concludes that most of invariant tori of integrable Hamiltonian systems can survive under small perturbation and with Kolmogorov’s nondegeneracy condition (Arnold 1963; Eliasson 1988; Kolmogorov 1954; Pöschel 2001).

Later important generalizations of the classical KAM theorem were made to the Rüssmann’s nondegeneracy condition (Cheng and Sun 1994; Rüssmann 1990, 2001; Sevryuk 1996; Xu et al. 1997). However, in the case of Rüssmann’s nondegeneracy condition, one can only get the existence of a family of invariant tori while there is no information on the persistence of frequency of any torus.

Chow et al. (2002) and Sevryuk (2006) considered perturbations of moderately degenerate integrable Hamiltonian systems and proved that the first \( d \) frequencies (\( d < n \), \( n \) denotes the freedom of Hamiltonian systems) of unperturbed invariant \( n \)-tori can persist. Xu and You (2010) proved that if some frequency satisfies certain nonresonant condition and topological degree condition, the perturbed system still has an invariant
torus with this frequency under Rüssmann’s nondegeneracy condition. Zhang and Cheng (2010) concerned with the persistence of invariant tori for nearly integrable Hamiltonian systems under time quasi-periodic perturbations; they proved that if the frequency of unperturbed system satisfies the Rüssmann’s nondegeneracy condition and has nonzero Brouwer’s topological degree at some Diophantine frequency, then invariant torus with frequency (Diophantine frequency and frequency of time quasi-periodic perturbation) satisfying the Diophantine condition persists under time quasi-periodic perturbations.

However, there are only few results available to obtain the existence of almost periodic solutions via KAM theory, because it is difficult to treat small divisor problem of infinite frequencies.

In this paper, we focus on the almost periodic case, that is, the perturbation in (1.1) is almost periodic in $t$ and admits the spatial series expansion, and want to prove most of invariant tori of the integrable Hamiltonian system can survive under small almost periodic perturbations and with Kolmogorov’s nondegeneracy condition.

After we get the invariant tori theorem, as an application, we shall prove the existence of almost periodic solutions and the boundedness of all solutions for the differential equation

$$\ddot{x} + x^{2l+1} = \sum_{j=0}^{2l} p_j(t)x^j, \quad x \in \mathbb{R},$$

(1.3)

where $p_0, p_1, \ldots, p_{2l}$ are real analytic almost periodic functions with the frequency $\omega = (\ldots, \omega_0, \ldots, \lambda, \ldots) \in \mathbb{Z}$.

External forced problem is an important feature of the classical perturbation for Hamiltonian systems. It is well known that the longtime behavior of a time-dependent nonlinear differential equation

$$\ddot{x} + f(t, x) = 0,$$

(1.4)

$f$ being periodic in $t$, can be very intricate. For example, there are equations having unbounded solutions but with infinitely many zeros and with nearby unbounded solution having randomly prescribed numbers of zeros and also periodic solution (see Moser 1973).

In contrast to such unboundedness phenomena, one may look for conditions on the nonlinearity, in addition to the superlinear condition that

$$\frac{1}{x}f(t, x) \to \infty \quad \text{as} \quad |x| \to \infty,$$

which allow to conclude that all solutions of (1.4) are bounded.

The problem was studied extensively for the differential equation (1.3) with $p_0, p_1, \ldots, p_{2l}$ being real analytic quasi-periodic functions in $t$ with the frequency $\omega = (\omega_1, \omega_2, \ldots, \omega_m)$. The first result was due to Morris (1976), who proved that every solution of equation
\[ \ddot{x} + x^3 = f(t), \quad (1.5) \]

with \( p(t + 1) = p(t) \) being continuous, is bounded. Morris also proved that there are infinitely many quasi-periodic solutions and the boundedness of all solutions of \((1.5)\) via Moser’s twist theorem (Moser 1962). This result, prompted by Littlewood in Littlewood (1968), was extended to \((1.3)\) with sufficiently smooth periodic \( p_j(t) \) by Dieckerhoff and Zehnder (1987). Later, their result was extended to more general cases by several authors. We refer to Levi (1986), Laedrich and Levi (1991), Norris (1992), You (1990), Yuan (1998) and references therein.

When \( p_0, p_1, \ldots, p_{2l} \) are quasi-periodic, by using the KAM iterations, Levi and Zehnder (1995) and Liu and You (1998) independently proved that there are infinitely many quasi-periodic solutions and the boundedness of all solutions for \((1.3)\) with \( p_0, p_1, \ldots, p_{2l} \) being sufficiently smooth and the frequency \( \omega = (\omega_1, \omega_2, \ldots, \omega_m) \) being Diophantine

\[ |\langle k, \omega \rangle| \geq \frac{\gamma}{|k|\tau}, \quad 0 \neq k \in \mathbb{Z}^m \]

for some \( \gamma > 0, \tau > m - 1 \). On the other hand, by establishing the invariant curve theorem of planar smooth quasi-periodic twist mappings in Huang et al. (2018), we obtained the existence of quasi-periodic solutions and the boundedness of all solutions for an asymmetric oscillation with a quasi-periodic external force (Huang et al. 2016).

One knows that the Diophantine condition is crucial when applying the KAM theory. A natural question is whether the boundedness for all solutions, called Lagrangian stability, still holds if \( \omega = (\omega_1, \omega_2, \ldots, \omega_m) \) is not Diophantine but Liouvillean? Wang and You (2016) proved the boundedness of all solutions of \((1.3)\) with \( m = 2 \) and \( \omega = (1, \alpha), \alpha \in \mathbb{R} \setminus \mathbb{Q} \), without assuming \( \alpha \) to be Diophantine.

Recently, in Huang et al. (2016) we established the invariant curve theorem of planar almost periodic twist mappings, as an application, we proved the existence of almost periodic solutions and the boundedness of all solutions of \((1.5)\) when \( f(t) \) is a real analytic almost periodic function with the frequency \( \omega = (\ldots, \omega_\lambda, \ldots)_{\lambda \in \mathbb{Z}} \). In Huang et al. (2017) we also established some variants of the invariant curve theorem on the basis of the invariant curve theorem obtained in Huang et al. (2016) and used them to study the existence of almost periodic solutions and the boundedness of all solutions for an asymmetric oscillation with an almost periodic external force.

Before ending the introduction, an outline of this paper is as follows. In Sect. 2, we first define real analytic almost periodic functions and their norms, then list some properties of them. The main invariant tori theorem (Theorem 3.4) is given in Sect. 3. The proof of the invariant tori theorem and the measure estimate are given in Sects. 4, 5, 6 and 7, respectively. In Sect. 8, we will prove the existence of almost periodic solutions and the boundedness of all solutions for \((1.3)\) with superquadratic potentials depending almost periodically on time.
2 Real Analytic Almost Periodic Functions and Their Norms

2.1 The Frequency of Real Analytic Almost Periodic Functions

Throughout the paper, we always assume that the real analytic almost periodic function has the Fourier exponents \( \{ \Omega_\lambda : \lambda \in \mathbb{Z} \} \), and its basis is the frequency \( \{ \omega_\lambda : \lambda \in \mathbb{Z} \} \) which is contained in \( \{ \Omega_\lambda : \lambda \in \mathbb{Z} \} \). Then for any \( \lambda \in \mathbb{Z} \), \( \Omega_\lambda \) can be uniquely expressed into

\[
\Omega_\lambda = r_{\lambda,1} \omega_{\lambda,1} + \cdots + r_{\lambda,j(\lambda)} \omega_{\lambda,j(\lambda)},
\]

where \( r_{\lambda,1}, \ldots, r_{\lambda,j(\lambda)} \) are rational numbers. Let

\[
S = \{ (\lambda_1, \ldots, \lambda_{j(\lambda)}) : \lambda \in \mathbb{Z} \}.
\]

Thus, \( S \) is a family of finite subsets \( A \) of \( \mathbb{Z} \), which reflects a spatial structure on \( \mathbb{Z} \subseteq \bigcup_{A \in S} A \).

For the bilateral infinite sequence of rationally independent frequency \( \omega = (\ldots, \omega_\lambda, \ldots) \),

\[
\langle k, \omega \rangle = \sum_{\lambda \in \mathbb{Z}} k_\lambda \omega_\lambda,
\]

where due to the spatial structure of \( S \) on \( \mathbb{Z} \), \( k \) runs over integer vectors whose support

\[
\text{supp } k = \{ \lambda : k_\lambda \neq 0 \}
\]

is a finite set of \( \mathbb{Z} \) contained in a subset \( A \) which belongs to \( S \). Moreover, we define

\[
\mathbb{Z}^S_{\subseteq} : = \left\{ k = (\ldots, k_\lambda, \ldots) \in \mathbb{Z}^\mathbb{Z} : \text{supp } k \subseteq A, \ A \in S \right\}.
\] (2.1)

In the following, we will give a norm for real analytic almost periodic functions. Before we describe the norm, some more definitions and notations are useful.

The main ingredient of our perturbation theory is a nonnegative weight function

\[
[ \cdot ] : A \mapsto [A]
\]

defined on \( S \). The weight of a subset may reflect its size, its location or something else. Throughout this paper, we always use the following weight function

\[
[A] = 1 + \sum_{\theta \in A} \log^\theta (1 + |\theta|),
\] (2.2)

where \( \varrho > 2 \) is a constant.
In this paper, the frequency \( \omega = (\ldots, \omega_\lambda, \ldots) \) of real analytic almost periodic functions is not only rationally independent with \( |\omega|_\infty = \sup\{|\omega_\lambda| : \lambda \in \mathbb{Z}\} < +\infty \), but also satisfies the strongly nonresonant condition (2.5) below.

In a crucial way, the weight function determines the nonresonance conditions for the small divisors arising in this theory. As we will do later on by means of an appropriate norm, it suffices to estimate these small divisors from below not only in terms of the norm of \( k \),

\[
|k| = \sum_{\lambda \in \mathbb{Z}} |k_\lambda|,
\]

but also in terms of the weight of its support

\[
[[k]] = \min_{\text{supp} k \subseteq A \in \mathcal{S}} |A|.
\]

Then, the nonresonance conditions read

\[
|\langle k, \omega \rangle| \geq \frac{\alpha}{\Delta([k]) \Delta(|k|)}, \quad 0 \neq k \in \mathbb{Z}_{\mathcal{S}}^Z,
\]

where, as usual, \( \alpha \) is a positive parameter and \( \Delta \) some fixed approximation function as described in the following. One and the same approximation function is taken here in both places for simplicity, since the generalization is straightforward. A nondecreasing function \( \Delta : [0, \infty) \rightarrow [1, \infty) \) is called an approximation function, if

\[
\frac{\log \Delta(t)}{t} \searrow 0, \quad t \to \infty,
\]

and

\[
\int_0^\infty \frac{\log \Delta(t)}{t^2} \, dt < \infty.
\]

In addition, the normalization \( \Delta(0) = 1 \) is imposed for definiteness.

In the following we will give a criterion for the existence of strongly nonresonant frequencies. It is based on growth conditions on the distribution function

\[
N_i(t) = \text{card} \left\{ A \in \mathcal{S} : \text{card}(A) = i, \ [A] \leq t \right\}
\]

for \( i \geq 1 \) and \( t \geq 0 \).

**Lemma 2.1** There exists a constant \( N_0 \) and an approximation function \( \Phi \) such that

\[
N_i(t) \leq \begin{cases} 
0, & t < t_i, \\
N_0 \Phi(t), & t \geq t_i
\end{cases}
\]
with a sequence of real numbers \( t_i \) satisfying

\[
i \log^{\varrho-1} i \leq t_i \sim i \log^{\varrho} i
\]

for \( i \) large with some exponent \( \varrho - 1 > 1 \). Here we say \( a_i \sim b_i \), if there are two constants \( c, C \) such that \( ca_i \leq b_i \leq Ca_i \) and \( c, C \) are independent of \( i \).

For a rigorous proof of Lemma 2.1 the reader is referred to Pöschel (1990) and Huang et al. (2016), we omit it here.

According to Lemma 2.1, there exist an approximation function \( \Delta \) and a probability measure \( u \) on the parameter space \( \mathbb{R}^Z \) with support at any prescribed point such that

\[
|\langle k, \omega \rangle| \geq \frac{\alpha}{\Delta([|k|]) \Delta(|k|)}, \quad \alpha > 0, \quad \text{for all } k \in \mathbb{Z}^Z \setminus \{0\}
\]

(2.5)
is positive for a suitably small \( \alpha \), the proof can be found in Pöschel (1990), we omit it here.

Throughout this paper, we assume that the frequency \( \omega = (\ldots, \omega_\lambda, \ldots) \) satisfying the nonresonance condition (2.5).

### 2.2 The Space of Real Analytic Almost Periodic Functions

In order to find almost periodic solutions \( x \) for (1.3), we have to define a kind of real analytic almost periodic functions which admit a spatial series expansion similar to (1.2) with the frequency \( \omega = (\ldots, \omega_\lambda, \ldots) \).

We first define the space of real analytic quasi-periodic functions \( Q(\omega) \) as in Siegel and Moser (1997, chapter 3), here the \( m \)-dimensional frequency vector \( \omega = (\omega_1, \omega_2, \ldots, \omega_m) \) is rationally independent, that is, for any \( k = (k_1, k_2, \ldots, k_m) \neq 0 \),

\[
\langle k, \omega \rangle = \sum_{j=1}^{m} k_j \omega_j \neq 0.
\]

**Definition 2.2** A function \( f : \mathbb{R} \to \mathbb{R} \) is called real analytic quasi-periodic with the frequency \( \omega = (\omega_1, \omega_2, \ldots, \omega_m) \), if there exists a real analytic function

\[
F : \theta = (\theta_1, \theta_2, \ldots, \theta_m) \in \mathbb{R}^m \to \mathbb{R}
\]
such that \( f(t) = F(\omega_1 t, \omega_2 t, \ldots, \omega_m t) \) for all \( t \in \mathbb{R} \), where \( F \) is \( 2\pi \)-periodic in each variable and bounded in a complex neighborhood \( \Pi_r^m = \{ (\theta_1, \theta_2, \ldots, \theta_m) \in \mathbb{C}^m : |\text{Im } \theta_j| \leq r, j = 1, 2, \ldots, m \} \) of \( \mathbb{R}^m \) for some \( r > 0 \). Here we call \( F(\theta) \) the shell function of \( f(t) \).

We denote by \( Q(\omega) \) the set of real analytic quasi-periodic functions with the frequency \( \omega = (\omega_1, \omega_2, \ldots, \omega_m) \). Given \( f(t) \in Q(\omega) \), the shell function \( F(\theta) \) of \( f(t) \) admits a Fourier series expansion.
where $k = (k_1, k_2, \ldots, k_m)$, $k_j$ range over all integers and the coefficients $f_k$ decay exponentially with $|k| = |k_1| + |k_2| + \cdots + |k_m|$, then $f(t)$ can be represented as a Fourier series of the type from the definition,

$$ f(t) = \sum_{k \in \mathbb{Z}^m} f_k e^{i(k, \omega)t}. $$

In the following, we define the norm of the real analytic quasi-periodic function $f(t)$ through that of the corresponding shell function $F(\theta)$.

**Definition 2.3** For $r > 0$, let $Q_r(\omega) \subseteq Q(\omega)$ be the set of real analytic quasi-periodic functions $f$ such that the corresponding shell functions $F$ are bounded on the subset $\Pi^m_r$ with the supremum norm

$$ |F|_r = \sup_{\theta \in \Pi^m_r} |F(\theta)| = \sup_{\theta \in \Pi^m_r} \left| \sum_k f_k e^{i(k, \theta)} \right| < +\infty. $$

Thus we define $|f|_r := |F|_r$.

Similarly, we give the definition of real analytic almost periodic functions with the frequency $\omega = (\ldots, \omega_\lambda, \ldots)$. For this purpose, we first define analytic functions on some infinite dimensional space (see Dineen 1999).

**Definition 2.4** Let $X$ be a complex Banach space. A function $f : U \subseteq X \rightarrow \mathbb{C}$, where $U$ is an open subset of $X$, is called analytic if $f$ is continuous on $U$, and $f|_{U \cap X_1}$ is analytic in the classical sense as a function of several complex variables for each finite dimensional subspace $X_1$ of $X$.

**Definition 2.5** A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is called real analytic almost periodic with the frequency $\omega = (\ldots, \omega_\lambda, \ldots) \in \mathbb{R}^\mathbb{Z}$, if there exists a real analytic function $F : \theta = (\ldots, \theta_\lambda, \ldots) \in \mathbb{R}^\mathbb{Z} \rightarrow \mathbb{R}$, which admits a spatial series expansion

$$ F(\theta) = \sum_{A \in \mathcal{S}} F_A(\theta), $$

where

$$ F_A(\theta) = \sum_{\supp k \subseteq A} f_k e^{i(k, \theta)}, $$

such that $f(t) = F(\omega t)$ for all $t \in \mathbb{R}$, where $F$ is $2\pi$-periodic in each variable and bounded in a complex neighborhood $\Pi_r = \{ \theta = (\ldots, \theta_\lambda, \ldots) \in \mathbb{C}^\mathbb{Z} : |\Im \theta|_\infty \leq r \}$. 
for some $r > 0$, where $|\text{Im} \theta|_\infty = \sup_{\lambda \in \mathbb{Z}} |\text{Im} \theta^\lambda|$. Here $F(\theta)$ is called the shell function of $f(t)$.

Denote by $AP(\omega)$ the set of real analytic almost periodic functions with the frequency $\omega$ defined by Definition 2.5. As a consequence of the definitions of $S$, the support $\text{supp} k$ of $k$, and $\mathbb{Z}^*_S$, from Definition 2.5, the spatial series expansion of the shell function $F(\theta)$ has another form

$$F(\theta) = \sum_{A \in S} \sum_{\text{supp} k \subseteq A} f_k e^{i(k,\theta)} = \sum_{k \in \mathbb{Z}^*_S} f_k e^{i(k,\theta)}.$$ 

Hence $f(t)$ can be represented as a series expansion of the type

$$f(t) = \sum_{k \in \mathbb{Z}^*_S} f_k e^{i(k,\omega)t}.  \quad (2.6)$$

If we define

$$f_A(t) = \sum_{\text{supp} k \subseteq A} f_k e^{i(k,\omega)t},$$

then

$$f(t) = \sum_{A \in S} f_A(t).$$

From the definitions of the support $\text{supp} k$ and $A$, we know that $f_A(t)$ is a real analytic quasi-periodic function with the frequency $\omega_A = \{\omega_\lambda : \lambda \in A\}$. Therefore, the almost periodic function $f(t)$ can be represented the sum of countably many quasi-periodic functions $f_A(t)$ formally.

### 2.3 The Norms and Some Properties of Real Analytic Almost Periodic Functions

Now we can define the norm of the real analytic almost periodic function $f(t)$ through that of the corresponding shell function $F(\theta)$ just like in the quasi-periodic case.

**Definition 2.6** Let $AP_r(\omega) \subseteq AP(\omega)$ be the set of real analytic almost periodic functions $f$ such that the corresponding shell functions $F$ are real analytic and bounded on the set $\Pi_r$ with the norm

$$\|F\|_{m,r} = \sum_{A \in S} |F_A|_r e^{m|A|} = \sum_{A \in S} |f_A|_r e^{m|A|} < +\infty,$$

where $m > 0$ is a constant and

$$|F_A|_r = \sup_{\theta \in \Pi_r} |F_A(\theta)| = \sup_{\theta \in \Pi_r} \left| \sum_{\text{supp} k \subseteq A} f_k e^{i(k,\theta)} \right| = |f_A|_r.$$
Hence we define

\[ \| f \|_{m,r} := \| F \|_{m,r}. \]

In the following, some properties of real analytic almost periodic functions are given.

**Lemma 2.7** The following statements are true:

(i) Let \( f(t), g(t) \in AP(\omega) \), then \( f(t) \pm g(t), g(t + f(t)) \in AP(\omega) \);

(ii) Let \( f(t) \in AP(\omega) \) and \( \tau = \beta t + f(t) (\beta + f' > 0, \beta \neq 0) \), then the inverse relation is given by \( t = \beta^{-1} \tau + g(\tau) \) and \( g \in AP(\omega/\beta) \). In particular, if \( \beta = 1 \), then \( g \in AP(\omega) \).

The detail proofs of Lemma 2.7 can be seen in Huang et al. (2016); we omit it here.

### 3 The Hamiltonian Setting and the Main Result

Consider the following Hamiltonian

\[ H = h(y) + f(x, y, t), \tag{3.1} \]

where \( y \in D \subseteq \mathbb{R}^n \) with \( D \) being an open bounded domain, \( x \in \mathbb{T}^n \), \( f(x, y, t) \) is a real analytic almost periodic function in \( t \) with the frequency \( \omega = (\ldots, \omega_\lambda, \ldots) \).

After introducing two conjugate variables \( \theta \in \mathbb{T}^Z \) and \( J \in \mathbb{R}^Z \), the Hamiltonian (3.1) can be written in the form of an autonomous Hamiltonian as follows

\[ H = \langle \omega, J \rangle + h(y) + F(x, y, \theta), \tag{3.2} \]

where \( F(x, y, \theta) \) is the shell function of the almost periodic function \( f(x, y, t) \) with respect to \( t \). Thus, the perturbed motion of Hamiltonian (3.1) is described by the following equations

\[ \begin{align*}
\dot{\theta} &= HJ = \omega, \\
\dot{x} &= H_y(y) = h_y(y) + F_y(x, y, \theta), \\
\dot{J} &= -H_\theta = -F_\theta(x, y, \theta), \\
\dot{y} &= -H_x = -F_x(x, y, \theta).
\end{align*} \tag{3.3} \]

When \( f = 0 \), the unperturbed system (3.3) has invariant tori \( T_0 = \mathbb{T}^Z \times \mathbb{T}^n \times \{0\} \times \{y_0\} \) with the frequency \( \tilde{\omega} = (\omega, \tilde{\omega}) \), carrying an almost periodic flow \( \theta = \omega t, x(t) = x_0 + \tilde{\omega}t \), where \( \tilde{\omega} = h_y(y_0) \). The aim is to prove the persistence of invariant tori under small perturbations.

We now make the assumption that this system is nondegenerate in the sense that

\[ \det h_{yy} = \det \frac{\partial h_y}{\partial y} \neq 0. \]
on $D$. Then, $h_y$ is an open map, even a local diffeomorphism between $D$ and some open frequency domain $O \subseteq \mathbb{R}^n$.

As in Pöschel (2001), instead of proving the existence of invariant tori for the Hamiltonian system (3.2) directly, we are going to concerned with the existence of invariant tori of a family of linear Hamiltonians. This is accomplished by introducing the frequency as independent parameters and changing the Hamiltonian system (3.2) to a parameterized system. This approach was first taken in Moser (1967).

To this end we write $y = y_0 + z$ and expand $h$ around $y_0$ so that

$$h(y) = h(y_0) + \langle h_y(y_0), z \rangle + \int_0^1 (1 - t)\langle h_{yy}(y_t)z, z \rangle \, dt,$$

where $y_t = y_0 + tz$. By the assumptions, the frequency map is a diffeomorphism

$$h_y : D \to O, \quad y_0 \mapsto \tilde{\omega} = h_y(y_0).$$

Hence, instead of $y_0 \in D$ we may introduce the frequency $\tilde{\omega} \in O$ as independent parameters, determining $y_0$ uniquely. Incidentally, the inverse map is given as

$$g_{\tilde{\omega}} : O \to D, \quad \tilde{\omega} \mapsto y_0 = g_{\tilde{\omega}}(\tilde{\omega}),$$

where $g$ is the Legendre transform of $h$, defined by $g(\tilde{\omega}) = \sup_y (\langle y, \tilde{\omega} \rangle - h(y))$. See Arnold (1978) for more details on Legendre transforms.

Thus, we write

$$H = e(\tilde{\omega}) + \langle \omega, J \rangle + \langle \tilde{\omega}, z \rangle + F(x, g_{\tilde{\omega}}(\tilde{\omega}) + z, \theta) + \int_0^1 (1 - t)\langle h_{yy}(y_t)z, z \rangle \, dt$$

and the term $O(|z|^2)$ can be taken as a new perturbation, we obtain the family of Hamiltonians $H = N + P$ with

$$N = e(\tilde{\omega}) + \langle \omega, J \rangle + \langle \tilde{\omega}, z \rangle,$$

$$P(\theta, x, z; \tilde{\omega}) = F(x, g_{\tilde{\omega}}(\tilde{\omega}) + z, \theta) + \int_0^1 (1 - t)\langle h_{yy}(g_{\tilde{\omega}}(\tilde{\omega}) + tz)z, z \rangle \, dt.$$

They are real analytic in the coordinates $(\theta, x, J, z)$ in $T^Z \times T^n \times \mathbb{R}^Z \times B$, $B$ some sufficiently small ball around the origin in $\mathbb{R}^n$, as well as the frequency $\tilde{\omega}$ taken from a parameters domain $O$ in $\mathbb{R}^n$.

From the definition of the shell function of the almost periodic function, $P$ admits a spatial series expansion

$$P(\theta, x, z; \tilde{\omega}) = \sum_{A \in S} \sum_{\supp \vec{k} \subseteq A} P_{A, k}(x, z; \tilde{\omega}) e^{i(k, \theta)}.$$
Expanding $P_{A,k}(x, z; \tilde{\omega})$ into a Fourier series at $x \in T^n$, therefore $P$ admits a spatial series

$$P(\theta, x, z; \tilde{\omega}) = \sum_{A \in S} P_A(\theta, x, z; \tilde{\omega}) = \sum_{A \in S} \sum_{\text{supp} k \subseteq A} \sum_{k \in \mathbb{Z}^n} P_{A,k,\tilde{k}}(z; \tilde{\omega}) e^{i((k,\theta) + (\tilde{k}, x))},$$

where $\tilde{k} = (\tilde{k}_1, \tilde{k}_2, \ldots, \tilde{k}_n)$. Let $B = \{1, 2, \ldots, n\}$, and supp $\tilde{k} \subseteq B$ for all $\tilde{k} \in \mathbb{Z}^n$. Denote $\tilde{S} = \{A \times B : A \in S\}$, then $P$ can be represented as a spatial series of the type

$$P = \sum_{\tilde{A} \in \tilde{S}} P_{\tilde{A}} = \sum_{\tilde{A} \in \tilde{S}} \sum_{\text{supp}(k,\tilde{k}) \subseteq \tilde{A}} P_{A,k,\tilde{k}}(z; \tilde{\omega}) e^{i((k,x) + (\tilde{k},\theta))}, \quad (3.4)$$

where supp $(k, \tilde{k}) = \text{supp} k \times \text{supp} \tilde{k}$, $P_{A,k,\tilde{k}} = P_{A,k,\tilde{k}}(z; \tilde{\omega})$. Moreover, we define

$$\mathbb{Z}_{\tilde{S}}^\mathbb{Z} := \{(k, \tilde{k}) \in \mathbb{Z}^n \times \mathbb{Z}^n : \text{supp} (k, \tilde{k}) \subseteq \tilde{A}, \tilde{A} \in \tilde{S}\}$$

and

$$\lfloor ((k, \tilde{k})) \rfloor = \min_{\text{supp} (k, \tilde{k}) \subseteq \tilde{A} \in \tilde{S}} |\tilde{A}|.$$

To state the invariant tori theorem, we therefore single out the subsets $O_\alpha \subseteq O$ denoted the set of all $\tilde{\omega}$ satisfying

$$|\lfloor (k, \omega) + (\tilde{k}, \tilde{\omega}) \rfloor \geq \frac{\alpha}{\Delta(\lfloor (k, \tilde{k}) \rfloor) \Delta(\lfloor |k| + |\tilde{k}| \rfloor)}, \quad \text{for all } (k, \tilde{k}) \in \mathbb{Z}_{\tilde{S}}^\mathbb{Z} \setminus \{\tilde{k} = 0\} \quad (3.5)$$

with fixed $\omega$, where $\Delta$ is an approximation function.

**Remark 3.1** For the distribution function $N_i(t)$ in (2.4), by replacing $A, S$ with $\tilde{A}, \tilde{S}$, Lemma 2.1 is also true.

**Remark 3.2** For the frequency $\omega$ being fixed, there is an approximation function $\Delta$ such that the set $O_\alpha$ is a set of positive Lebesgue measure provided that $\alpha$ is small (see Theorem 7.5 in Sect. 7).

**Remark 3.3** From the proofs of Theorem 7.5 and the measure of the set of $\omega$ in Pöschel (1990), we know that there is the same approximation function $\Delta$ such that (2.5) and (3.5) hold simultaneously.

To state the basic result quantitatively, we need to introduce a few notations. Let

$$\mathcal{D}_{r,s} = \{(\theta, x, J, z) : |\text{Im} \theta| < r, |\text{Im} x| < r, |J| < |z| < s\}$$

$$\subseteq \mathbb{C}^2/2\pi \mathbb{Z}^2 \times \mathbb{C}^n/2\pi \mathbb{Z}^n \times \mathbb{C}^z \times \mathbb{C}^n$$
and
\[ \mathcal{O}_h = \{ \tilde{\omega} : |\tilde{\omega} - \mathcal{O}_\alpha| < h \} \subset \mathbb{C}^n, \]
denote complex neighborhoods of the \( T^Z \times T^n \times \{ 0, 0 \} \) and \( \mathcal{O}_\alpha \), respectively, where
\[ |\theta|_\infty = \sup_{\lambda \in \mathbb{Z}} |\theta_{\lambda}|, \quad |J|_w = \sum_{\lambda \in \mathbb{Z}} |J_{\lambda}| e^{w|\lambda|}, \]
|·| stands for the sup-norm of real vectors, where \( w \geq 0 \) is another parameter, and the weight at the individual lattice site is defined by \( |\lambda| = \min_{\lambda \in A} \mathbb{E}[A] \).

Its size is measured in terms of the weighted norm
\[ |||P|||_{m,r,s,h} = \sum_{\tilde{A} \in \tilde{S}} \|P_{\tilde{A}}\|_{r,s,h} e^{m|\tilde{A}|}, \]

where
\[ \|P_{\tilde{A}}\|_{r,s,h} = \sum_{\text{supp}(k,\tilde{k}) \subseteq \tilde{A}} |P_{\tilde{A},k,\tilde{k}}|_{s,h} e^{r(|k|+|\tilde{k}|)}, \]
the norm \( |·|_{s,h} \) is the sup-norm over \(|z| < s \) and \( \tilde{\omega} \in \mathcal{O}_h \).

Now, the Hamiltonian is
\[ H(\theta, x, J, z; \tilde{\omega}) = N + P = e(\tilde{\omega}) + \langle \omega, J \rangle + \langle \tilde{\omega}, z \rangle + P(\theta, x, z; \tilde{\omega}), \quad (3.6) \]
P is real analytic on \( D_{r,s} \times \mathcal{O}_h \). The corresponding Hamiltonian system (3.3) becomes
\[ \dot{\theta} = H_J = \omega, \]
\[ \dot{x} = H_z = \tilde{\omega} + P_z(\theta, x, z; \tilde{\omega}), \]
\[ \dot{J} = -H_\theta = -P_\theta(\theta, x, z; \tilde{\omega}), \]
\[ \dot{z} = -H_x = -P_x(\theta, x, z; \tilde{\omega}). \]

Thus, the persistence of invariant tori for nearly integrable Hamiltonian system (3.3) is reduced to the persistence of invariant tori for the family of Hamiltonian system (3.6) depending on the parameter \( \tilde{\omega} \). Our aim is to prove the persistence of the invariant torus
\[ \mathcal{T}_0 = T^Z \times T^n \times \{ 0, 0 \} \]
of maximal dimension together with its constant vector field \((\omega, \tilde{\omega})\).

The smallness condition of the following theorem is expressed in terms of two functions \( \Psi_0, \Psi_1 \) that are defined on the positive real axis entirely in terms of the approximation function \( \Delta \) and reflect the effect of the small divisors in solving the nonlinear problem. See Appendix A in Pöschel (1990) for their definitions.

Now we are in a position to state our main result.
Theorem 3.4 Suppose that $P$ admits a spatial expansion as in (3.4), is real analytic on $D_{r,s} \times O_h$ and satisfies the estimate

$$ s^{-1} |||P|||_{m,r,s,h} \leq \frac{\alpha \varepsilon_*}{\Psi_0(\mu) \Psi_1(\rho)} \leq \frac{h}{2^6} $$

for some $0 < \mu \leq m - w$ and $0 < \rho < r/2$, where $\varepsilon_* = 2^{-22}$ is an absolute positive constant, $\Psi_0(\mu), \Psi_1(\rho)$ is defined by (6.1). Then, there exists a transformation

$$ \mathcal{F} : D_{r-2\rho, s/2} \times O_\alpha \to D_{r,s} \times O_h, $$

that is real analytic and symplectic for each $\tilde{\omega}$ and uniformly continuous in $\tilde{\omega}$, such that

$$ (N + P) \circ \mathcal{F} = e_* + \langle \omega, J \rangle + \langle \tilde{\omega}, z \rangle + \ldots, $$

where the dots denote terms of higher-order in $z$. Consequently, the perturbed system has a real analytic invariant torus of maximal dimension and with a vector field conjugate to $(\omega, \tilde{\omega})$ for each frequency vector $\tilde{\omega}$ in $O_\alpha$. These tori are close of order $s^{-1} |||P|||_{m,r,s,h}$ to the torus $T_0$ with respect to the norm $| \cdot |_w$.

4 Outline of the Proof of Theorem 3.4

Theorem 3.4 is proven by the familiar KAM-method employing a rapidly converging iteration scheme (Arnold 1963; Kolmogorov 1954; Moser 1967; Pöschel 1990). At each step of the scheme, a Hamiltonian $H_j = N_j + P_j$ is considered, which is a small perturbation of some normal form $N_j$. A transformation $\mathcal{F}_j$ is set up so that

$$ H_j \circ \mathcal{F}_j = N_{j+1} + P_{j+1} $$

with another normal form $N_{j+1}$ and a much smaller error term $P_{j+1}$. For instance,

$$ |||P_{j+1}||| \leq C_j |||P_j|||^\kappa $$

for some $\kappa > 1$. This transformation consists of a symplectic change of coordinates $\Phi_j$ and a subsequent change $\varphi_j$ of the parameters $\tilde{\omega}$ and is found by linearizing the above equation. Repetition of this process leads to a sequence of transformations $\mathcal{F}_0, \mathcal{F}_1, \ldots$, whose infinite product transforms the initial Hamiltonian $H_0$ into a normal form $N_*$ up to first order.

Here is a more detailed description of this construction. To describe one cycle of this iterative scheme in more detail, we now drop the index $j$. 

Springer
Approximating the perturbation $P$ in a suitable way, we write

$$H = N + P = N + R + (P - R).$$

In particular, $R$ is chosen such that its spatial series expansion is finite; hence, all subsequent operations are finite dimensional.

The coordinate transformation $\Phi$ is written as the time-1-map of the flow $X^t_F$ of a Hamiltonian vector field $X_F$:

$$\Phi = X^t_F |_{t=1}.$$ 

This makes $\Phi$ symplectic. Moreover, we may expand $H \circ \Phi = H \circ X^t_F |_{t=1}$ with respect to $t$ at 0 using Taylor’s formula. Recall that

$$\frac{d}{dt} G \circ X^t_F = \{G, F\} \circ X^t_F,$$

the Poisson bracket of $G$ and $F$ evaluated at $X^t_F$. Thus, we may write

$$(N + R) \circ \Phi = N \circ X^t_F |_{t=1} + R \circ X^t_F |_{t=1}$$

$$= N + \{N, F\} + \int_0^1 (1 - t)\{\{N, F\}, F\} \circ X^t_F \, dt$$

$$+ R + \int_0^1 \{R, F\} \circ X^t_F \, dt$$

$$= N + R + \{N, F\} + \int_0^1 ((1 - t)\{N, F\} + R, F\} \circ X^t_F \, dt.$$ 

The last integral is of quadratic order in $R$ and $F$ and will be part of the new error term.

The point is to find $F$ such that $N + R + \{N, F\} = \hat{N}$ is a normal form. Equivalently, setting $N_+ = N + \hat{N}$, the linear equation

$$\{F, N\} + \hat{N} = R \tag{4.1}$$

has to be solved for $F$ and $\hat{N}$, when $R$ is given. Given such a solution, we obtain $(1 - t)\{N, F\} + R = (1 - t)\hat{N} + tR$ and hence $H \circ \Phi = N_+ + P_+$ with

$$P_+ = \int_0^1 \{(1 - t)\hat{N} + tR, F\} \circ X^t_F \, dt + (P - R) \circ \Phi.$$ 

Setting up the spatial expansions for $F$ and $\hat{N}$ of the same form as that for $R$, the linearized equation (4.1) breaks up into the component equations

$$i(\langle k, \omega \rangle + \langle \tilde{k}, \tilde{\omega} \rangle) \hat{F}^A + \hat{\hat{N}}^A = R^A.$$
Their solutions are well known and straightforward. These equations introduce the small divisors, which in our case are zero if and only if \((k, \tilde{k})\) is zero by the nonresonance conditions. It therefore suffices to choose 
\[
\hat{N}_{\tilde{A}} = [R_{\tilde{A}}],
\]
the mean value of \(R_{\tilde{A}}\) over \(\mathbb{T}_{\tilde{A}}\), and to solve uniquely
\[
i ((k, \omega) + (\tilde{k}, \tilde{\omega})) F_{\tilde{A}} = R_{\tilde{A}} - [R_{\tilde{A}}], \quad [F_{\tilde{A}}] = 0.
\]
We obtain
\[
F_{\tilde{A}} = \sum_{\text{supp} (k, \tilde{k}) \subseteq \tilde{A}, (k, \tilde{k}) \neq 0} \frac{R_{\tilde{A}, k, \tilde{k}}}{l((k, \omega) + (\tilde{k}, \tilde{\omega}))} e^{i((k, \theta) + (\tilde{k}, \tilde{\omega}))}, \quad (4.2)
\]
where \(R_{\tilde{A}, k, \tilde{k}}\) are the Fourier coefficients of \(R_{\tilde{A}}\).

The truncation of \(P\) will be chosen so that \(R\) is independent of \(J\) and of first order in \(z\). Hence, the same is true of each of the \(\hat{N}_{\tilde{A}}\) and so
\[
\tilde{N} = \sum_{\tilde{A} \in S} \hat{N}_{\tilde{A}} = \hat{e} + \langle v(\tilde{\omega}), z \rangle.
\]
It suffices to change parameters by setting
\[
\tilde{\omega}_+ = \tilde{\omega} + v(\tilde{\omega}) \quad (4.3)
\]
to obtain a new normal form \(N_+ = N + \tilde{N}\). This completes one cycle of the iteration.

By the same truncation, \(F\) is independent of \(J\) and of first order in \(z\). It follows that \(\Phi = X_F^t \big|_{t=1}\) has the form
\[
\theta = \theta_+, \quad x = U_1(\theta_+, x_+), \quad J = U_2(\theta_+, x_+) + U_3(\theta_+, x_+) z_+, \quad z = U_4(\theta_+, x_+) + U_5(\theta_+, x_+) z_+,
\]
where the dependence of all coefficients on \(\tilde{\omega}\) has been suppressed. This map is composed of the inverse \(\varphi\) of the parameter map (4.3) to obtain \(\mathcal{F}\).

Such symplectic transformations form a group under composition. So, if \(\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_{j-1}\) belong to this group, then so does \(\mathcal{F}^j = \mathcal{F}_0 \circ \mathcal{F}_1 \circ \cdots \circ \mathcal{F}_{j-1}\) and the limit transformation \(\mathcal{F}\) for \(j \to \infty\).

In this paper, we use the converging iteration scheme in Pöschel (1990), the notations and format in Sects. 5, 6 of this paper are the same as Sections 6, 7 of Pöschel (1990), respectively. But the content is not exactly the same because in this paper we consider a Hamiltonian of the form (3.6), and only intrinsic frequency \(\tilde{\omega}\) have some drifts at each
KAM step, the external frequency $\omega$ remains the same. There are a lot of differences in details between Sects. 5, 6 of this paper and Sections 6, 7 of Pöschel Pöschel (1990). Moreover, for the KAM theory, the KAM step, iteration and convergence are very important. It is necessary for us to retain the contents of Sects. 5, 6 in this paper.

5 The KAM Step

Before plunging into the details of the KAM construction, we observe that it suffices to consider some normalized value of $\alpha$, say

$$\tilde{\alpha} = 2.$$ 

Indeed, stretching the time scale by the factor $2/\alpha$ the Hamiltonians $H$ and $N$ are scaled by the same amount, and so is the frequency $\tilde{\omega}$. By a similar scaling of the action-variables $J, z$ the radius $s$ may also be normalized to some convenient value. We will not do this here.

5.1 The Set Up

Consider a Hamiltonian of the form

$$H = N + P = e(\tilde{\omega}) + (\omega, J) + (\tilde{\omega}, z) + P(\theta, x, z; \tilde{\omega}),$$

where

$$P = \sum_{\tilde{A} \in \tilde{S}} P_{\tilde{A}} = \sum_{\tilde{A} \in \tilde{S}} \sum_{\text{supp}(k, \tilde{k}) \subseteq \tilde{A}} P_{A, k, \tilde{k}} e^{i(\langle k, x \rangle + \langle \tilde{k}, \theta \rangle)}.$$ 

Assume that $P$ is real analytic on the complex domain

$$D_{r, s} \times O_h : |\text{Im } \theta|_\infty < r, |\text{Im } x| < r, |J|_w < s, |z| < s, |\tilde{\omega} - O_\ast| < h,$$

where $O_\ast$ is a closed subset of the parameter space $\mathbb{R}^n$ consisting of the frequency $\tilde{\omega}$ that satisfying

$$|\langle k, \omega \rangle + (\tilde{k}, \tilde{\omega})| \geq \frac{\tilde{\alpha}}{\Delta(||(k, \tilde{k})||)} \Delta(|k| + |\tilde{k}|),$$

for all $(k, \tilde{k}) \in \mathbb{Z}_B^2 \setminus \{\tilde{k} = 0\}$, (5.1)

where $\tilde{\alpha} = 2$. Moreover, assume that for some $m > w$,

$$|||H - N|||_{m, r, s, h} = |||P|||_{m, r, s, h} \leq \epsilon$$

(5.2)

is sufficiently small. The precise condition will be given later in the course of the iteration.
Unless stated otherwise the following estimates are uniform with respect to $\widetilde{\omega}$. Therefore, the index $h$ is usually dropped.

### 5.2 Truncating the Perturbation

Let $\mu$ and $\rho$ be two small and $K$ a large positive parameter to be chosen during the iteration process. The Fourier series of the $\widetilde{A}$-component $P_{\widetilde{A}}$ of the perturbation is truncated at order $(|k| + |\widetilde{k}|) \leq \langle \widetilde{A} \rangle$ which is the smallest nonnegative number satisfying

$$\mu[\widetilde{A}] + \rho(\langle \widetilde{A} \rangle) \geq K. \quad (5.3)$$

Thus, the larger $[\widetilde{A}]$, the more Fourier coefficients are discarded. If $[\widetilde{A}]$ is sufficiently large, the whole $\widetilde{A}$-component is dropped. The upshot is that for the remaining perturbation $Q$ one has

$$|||P - Q|||_{m-\mu,r-\rho,s} \leq e^{-K} |||P|||_{m,r,s}. \quad (5.4)$$

for $0 < \mu < m$, $0 < \rho < r$ and $0 < \eta < 1$. Moreover, the estimate

$$|||R|||_{m,r,s} \leq 2|||P|||_{m,r,s} \quad (5.5)$$

obviously holds.

### 5.3 Extending the Small Divisor Estimate

We claim that if

$$h \leq \min_{A\in S} \frac{1}{\Delta([A]) \langle A \rangle \Delta(\langle A \rangle)} \quad (5.6)$$

with $\langle A \rangle$ as in the previous subsection 5.2, then the estimates

$$|(k, \omega) + (\widetilde{k}, \widetilde{\omega})| \geq \frac{1}{\Delta([((k, \widetilde{k}))]) \Delta(|k + \widetilde{k}|)},$$

$$\mu[[(k, \widetilde{k})]] + \rho(|k| + |\widetilde{k}|) \leq K, \quad (k, \widetilde{k}) \in \mathbb{Z}_{S}^{2}\backslash \{\widetilde{k} = 0\} \quad (5.7)$$

hold uniformly in $\widetilde{\omega}$ on the complex neighborhood $O_{h}$ of the set $O_{*}$. 

© Springer
The proof is simple. Given \( \tilde{\omega} \) in \( O_h \) there exists an \( \tilde{\omega}_0 \) in \( O_\ast \) such that \( |\tilde{\omega} - \tilde{\omega}_0| < h \). Given \( (k, \tilde{k}) \) there exists an \( \tilde{A} \) in \( \tilde{S} \) containing the support of \( (k, \tilde{k}) \) such that \( [[(k, \tilde{k})]] = [\tilde{A}] \). It follows that \( |k| + |\tilde{k}| \leq \langle \tilde{A} \rangle \) and hence

\[
|\langle k, \omega \rangle - \langle \tilde{k}, \omega_0 \rangle| \leq |\tilde{k}| \frac{1}{\Delta(\langle \tilde{A} \rangle) \Delta(\langle \tilde{A} \rangle)} \leq \frac{1}{\Delta([[k, \tilde{k}]] \Delta(|k| + |\tilde{k}|)}
\]

by the monotonicity of \( \Delta \). The claim follows from the estimate (5.1) for \( \langle \tilde{k}, \omega_0 \rangle \).

### 5.4 Solving the Linearized Equation

The KAM theorem is proven by the usual Newton-type iteration procedure, which involves an infinite sequence of coordinate changes and is described in some detail for example in Pöschel (1989). Each coordinate change \( \Phi \) is obtained as the time-1-map \( X_F|_{t=1} \) of a Hamiltonian vector field \( X_F \). Its generating Hamiltonian \( F \) as well as some correction \( \hat{N} \) to the given normal form \( N \) are a solution of the linearized equation

\[
\{F, N\} + \hat{N} = R,
\]

which is the subject of this subsection.

The linearized equation \( \{F, N\} + \hat{N} = R \) is broken up into the component equations

\[
i(\langle k, \omega \rangle + \langle \tilde{k}, \xi \rangle) F_{\tilde{A}} + \hat{N}_{\tilde{A}} = R_{\tilde{A}} \]

with supp \((k, \tilde{k}) \subseteq \tilde{A}, (k, \tilde{k}) \neq 0, \) and solved for \( F_{\tilde{A}} \) and \( \hat{N}_{\tilde{A}} \) as described in Sect. 4. Clearly, \( \hat{N}_{\tilde{A}} = [R_{\tilde{A}}], \) which is the mean value of \( R_{\tilde{A}} \) over \( \mathbb{T}^\tilde{A} \) and \( \|\hat{N}_{\tilde{A}}\|_{r,s} \leq \|R_{\tilde{A}}\|_{r,s}. \) Hence

\[
|||\hat{N}|||_{m,r,s} \leq |||R|||_{m,r,s} \tag{5.8}
\]

by putting pieces together.

The normalized Fourier series expansion of \( F_{\tilde{A}} \) is given by (4.2). By the extended small divisor estimate (5.7) and nonresonance condition (2.5) with \( \alpha = \tilde{\alpha} = 2, \)

\[
\|F_{\tilde{A}}\|_{r-\rho,s} \leq \sum_{\text{supp } (k, \tilde{k}) \subseteq \tilde{A} \atop (k, \tilde{k}) \neq 0} \Delta([[k, \tilde{k}]] \Delta(|k| + |\tilde{k}|) |R_{\tilde{A},k,\tilde{k}}| \leq \Gamma_0(\rho) \|R\|_{r,s},
\]

\( \square \) Springer
where $\Gamma_0(\rho) = \sup_{t \geq 0} \Delta(t)e^{-\rho t}$. Similarly, for the convenience of later estimates,

$$
\sum_{\lambda \in A} \| \partial_{\theta^A} F \|_{r-\rho,s} \leq \sum_{(k,\bar{k}) \in \tilde{A}} \Delta \left( \left| \sum_{\mu \in A} F_{A,k} \right| + \left| \sum_{\mu \in A} F_{A,k} \right| e^{(r-\rho)(|k| + |\bar{k}|)} \right) \\
\sum_{(k,\bar{k}) \in \tilde{A}} \Delta \left( \left| \sum_{\mu \in A} F_{A,k} \right| + \left| \sum_{\mu \in A} F_{A,k} \right| e^{(r-\rho)(|k| + |\bar{k}|)} \right) \\
\sum_{(k,\bar{k}) \in \tilde{A}} \Delta \left( \left| \sum_{\mu \in A} F_{A,k} \right| + \left| \sum_{\mu \in A} F_{A,k} \right| e^{(r-\rho)(|k| + |\bar{k}|)} \right) \\
\sum_{(k,\bar{k}) \in \tilde{A}} \Delta \left( \left| \sum_{\mu \in A} F_{A,k} \right| + \left| \sum_{\mu \in A} F_{A,k} \right| e^{(r-\rho)(|k| + |\bar{k}|)} \right)
$$

where $\Gamma_1(\rho) = \sup_{t \geq 0}(1 + t)\Delta(t)e^{-\rho t}$. Putting the spatial components together,

$$
\| F \|_{m-\mu,r-\rho,s} \leq \sum_{A \in \mathcal{A}} \Delta \left( \left| \sum_{\mu \in A} \right| F_{A} \| R \|_{r,s} e^{(m-\mu)[\tilde{A}]} \right) \\
\leq \Gamma_0(\mu) \Gamma_1(\rho) \| R \|_{m,r,s}
$$

and

$$
\sum_{\lambda} \| \partial_{\theta^A} F \|_{m-\mu,r-\rho,s} \| \partial_{\bar{k}} F \|_{m-\mu,r-\rho,s} \leq \Gamma_0(\mu) \Gamma_1(\rho) \| R \|_{m,r,s}
$$

for $0 < \mu < m$.

In view of the estimate $\Gamma_0(\rho) \leq \rho \Gamma_1(\rho)$ in Lemma 6 in Pöschel (1990), we may summarize these estimates by writing

$$
\rho^{-1} \| F \|_{m-\mu,r-\rho,s} \leq \Gamma_\mu \Gamma_\rho \| R \|_{m,r,s}, \\
\sum_{\lambda} \| \partial_{\theta^A} F \|_{m-\mu,r-\rho,s} \leq \Gamma_\mu \Gamma_\rho \| R \|_{m,r,s}, \\
\| \partial_{\bar{k}} F \|_{m-\mu,r-\rho,s} \leq \Gamma_\mu \Gamma_\rho \| R \|_{m,r,s}
$$

with

$$
\Gamma_\mu = \Gamma_0(\mu) = \sup_{t \geq 0} \Delta(t)e^{-\mu t}, \quad \Gamma_\rho = \Gamma_1(\rho) = \sup_{t \geq 0}(1 + t)\Delta(t)e^{-\rho t}.
$$

5.5 The Derivatives of $F$

On the domain $\mathcal{D}_{r-\rho,s}$ we obtain the estimate

$$
| \partial_{\theta^A} F \|_w = \sum_{\lambda} | \partial_{\theta^A} F | e^{w[\lambda]} \leq \sum_{\lambda} \sum_{A \ni \lambda} | \partial_{\theta^A} F \|_{r-\rho,s} e^{w[A]} \\
\leq \sum_{\lambda} \| \partial_{\theta^A} F \|_{w,r-\rho,s}
$$

 Springer
and
\[ |\partial_x F| \leq |||\partial_x F|||_{m-\mu,r-\rho,s}. \tag{5.12} \]
Similarly, on the domain \( D_{r-\rho,s/2} \) we obtain estimate
\[ |\partial_J F|_\infty = 0 \tag{5.13} \]
and
\[ |\partial_z F| \leq |||\partial_z F|||_{m-\mu,r-\rho,s/2} \leq \frac{2}{s}|||F|||_{m-\mu,r-\rho,s}. \tag{5.14} \]
Requiring that
\[ m - \mu \geq w \tag{5.15} \]
and recalling the estimates (5.2), (5.5), (5.9), (5.11), (5.12), (5.13), (5.14) we thus have
\[ \frac{1}{\rho} |\partial_J F|_\infty, \frac{1}{\rho} |\partial_z F|, \frac{2}{s} |\partial_J F|_w, \frac{2}{s} |\partial_z F| \leq \frac{2}{s} \Gamma_\mu \Gamma_\rho |||R|||_{m,r,s} \leq 4 \Gamma_\mu \Gamma_\rho \frac{\epsilon}{s} \]
uniformly on the domain \( D_{r-\rho,s/2} \).
These estimates are expressed more conveniently by means of a weighted phase space norm. Let
\[ |(\theta, x, J, z)|_\mathcal{P} = \max (|\theta|_\infty, |x|, |J|_w, |z|), \]
\[ W = \text{diag}(\rho^{-1}I_{2s}, \rho^{-1}I_n, 2s^{-1}I_{2s}, 2s^{-1}I_n). \]
Then, the above estimates are equivalent to
\[ |WX_F|_\mathcal{P} \leq 4 \Gamma_\mu \Gamma_\rho E, \quad E = \frac{\epsilon}{s} \]
on \( D_{r-\rho,s/2} \).

5.6 Transforming the Coordinates

The \( |W \cdot |_\mathcal{P} \) -distance of the domain
\[ D_\Sigma = D_{r-2\rho,s/4} \subset D_L = D_{r-\rho,s/2} \]
to the boundary of \( D_L \) is exactly one half. Hence, if \( 16 \Gamma_\mu \Gamma_\rho E \leq 1 \), then \( |WX_F|_\mathcal{P} \) is less than or equal one fourth on \( D_L \) and consequently
\[ X^t_F : D_\Sigma \to D_L, \quad 0 \leq t \leq 1. \]
In particular, the time-1-map $\Phi$ is a symplectic map from $D_L$ into $D_L$, for which the estimate

$$|W(\Phi - id)|_{\mathcal{P}; D_L} \leq 4 \Gamma_\mu \Gamma_\rho E$$

(5.16)

holds.

In fact, under the present smallness condition on $E$ this statement holds as well for the larger domain $D_{r-\kappa \rho, \kappa \lambda / 4}$ instead of $D_L$, where $\kappa = 3/2$. The $|W \cdot |_{\mathcal{P}}$-distance of its boundary to $D_L$ is exactly one fourth. Applying the general Cauchy’s inequality of Appendix B in Pöschel (1990) to the last estimate, it follows that in addition,

$$|W(D\Phi - I)W^{-1}|_{\mathcal{P}; D_L} \leq 16 \Gamma_\mu \Gamma_\rho E,$$

where the norm of derivative is the operator norm induced by $| \cdot |_{\mathcal{P}}$, see Appendix A in Pöschel (1990) for its definition. Finally, if we require

$$4 \Gamma_\mu \Gamma_\rho E \leq \eta \leq \frac{1}{2},$$

then

$$X^t_F : D_\beta = D_{r-2\rho, \eta s / 2} \rightarrow D_\eta = D_{r-\rho, \eta s}, \quad 0 \leq t \leq 1$$

by the same arguments as before.

### 5.7 Transforming the Frequencies

To put $N_+ = N + \tilde{N}$ into normal form, the frequency parameters are transformed by setting $\tilde{\omega}_+ = \tilde{\omega} + v(\tilde{\omega})$. Proceeding just as in (5.8) the estimate for $\tilde{N}$ implies that $|v|_{h/2} = |\partial_z \tilde{N}|_{h/2} \leq 4E$. Referring to Lemma 11 in Pöschel (1990) or Lemma A.3 in Pöschel (2001) it follows that for

$$E \leq \frac{h}{16}$$

(5.17)

and the map $id + v$ has a real analytic inverse

$$\varphi : \mathcal{O}_{h/4} \rightarrow \mathcal{O}_{h/2}, \quad \tilde{\omega}_+ \mapsto \tilde{\omega}$$

with the estimate

$$|\varphi - id|, \quad \frac{h}{4} \left| D\varphi - Id \right| \leq 4E$$

(5.18)

uniformly on $\mathcal{O}_{h/4}$. 
5.8 Estimating the New Error Term

The new error term is

\[ P_+ = \int_0^1 \{ R_t, F \} \circ X_t^1 \, dt + (P - R) \circ X_t^1, \]

where \( R_t = (1 - t)\hat{N} + tR \). By Lemma 10 in Pöschel (1990) and the estimate (5.9),

\[ |||G \circ X_t^1|||_{m-\mu,r-2\rho,\eta s/2} \leq 2 |||G|||_{m-\mu,r-\rho,\eta s}, \quad 0 \leq t \leq 1, \]

provided that

\[ 4 C_0 \Gamma_\mu \Gamma_\rho E \leq \eta \leq \frac{1}{2}, \quad (5.19) \]

where \( C_0 = 8 \) is a constant. Hence, with this assumption, it follows that

\[ |||P_+|||_{m-\mu,r-2\rho,\eta s/2} \leq 2 |||\{ R_t, F \}|||_{m-\mu,r-\rho,\eta s} + 2 |||P - R|||_{m-\mu,r-\rho,\eta s}. \]

Obviously, \( |||R_t|||_{m,r,s} \leq 2\varepsilon \) for \( 0 \leq t \leq 1 \) by the estimates for \( \hat{N} \) and \( F \), and therefore by (5.2), (5.5), (5.9), we get

\[ \begin{align*}
|||\partial_x R_t|||_{m-\mu,r-\rho,\eta s} |||\partial_x F|||_{m-\mu,r-\rho,\eta s} & \leq \frac{2\varepsilon}{\rho} \frac{2}{s} |||F|||_{m-\mu,r-\rho,\eta s} \leq 8 \Gamma_\mu \Gamma_\rho E \varepsilon, \\
|||\partial_z R_t|||_{m-\mu,r-\rho,\eta s} |||\partial_z F|||_{m-\mu,r-\rho,\eta s} & \leq \frac{2\varepsilon}{s} |||\partial_z F|||_{m-\mu,r-\rho,\eta s} \leq 8 \Gamma_\mu \Gamma_\rho E \varepsilon.
\end{align*} \]

Hence

\[ |||\{ R_t, F \}|||_{m-\mu,r-\rho,\eta s} \leq |||\partial_x R_t|||_{m-\mu,r-\rho,\eta s} |||\partial_x F|||_{m-\mu,r-\rho,\eta s} + |||\partial_z R_t|||_{m-\mu,r-\rho,\eta s} |||\partial_z F|||_{m-\mu,r-\rho,\eta s} \leq 16 \Gamma_\mu \Gamma_\rho E \varepsilon \]

in view of (5.13) and \( R_t \) independent of \( J \). Combined with (5.2), (5.4) we altogether obtain

\[ |||P_+|||_{m-\mu,r-2\rho,\eta s/2} \leq 32 \Gamma_\mu \Gamma_\rho E \varepsilon + 2e^{-K} \varepsilon + 4\eta^2 \varepsilon \quad (5.20) \]

for the new error term.
6 Iteration and Convergence

6.1 The Iterative Construction

To iterate the KAM step infinitely often, we now choose sequences for the pertinent parameters. Let $a = 13$, $b = 4$, $c = 6$, $d = 8$, $e = 22$ and $\kappa = 3/2$. The choice of these integer constants will be motivated later in the course of the proof of the iterative lemma.

Given $0 < \mu \leq m - w$ and $0 < \rho < r/2$ there exist sequences $\mu_0 \geq \mu_1 \geq \cdots > 0$ and $\rho_0 \geq \rho_1 \geq \cdots > 0$ such that

$$
\Psi_0(\mu)\Psi_1(\rho) = \prod_{\nu=0}^{\infty} \Gamma_{\mu_\nu}^{\kappa_{\nu}} \Gamma_{\rho_\nu}^{\kappa_{\nu}} 
$$

(6.1)

with

$$
\sum_{\nu=0}^{\infty} \mu_\nu = \mu, \quad \sum_{\nu=0}^{\infty} \rho_\nu = \rho, \quad \kappa_\nu = \frac{\kappa - 1}{\kappa^{v+1}},
$$

where $\Gamma_{\mu_\nu} = \Gamma_0(\mu_\nu)$ and $\Gamma_{\rho_\nu} = \Gamma_1(\rho_\nu)$, $\Gamma_0, \Gamma_1$ is defined by (5.10). Fix such sequences, and for $j \geq 0$ set

$$
\Gamma_j = 2^{j+a} \Gamma_{\mu_j} \Gamma_{\rho_j}, \quad \Theta_j = \prod_{\nu=0}^{j-1} \Gamma_{\nu}^{\kappa_{\nu}}, \quad E_j = (\Theta_j E_0)^{\kappa^j},
$$

where $\Theta_0 = 1$. Furthermore, set

$$
m_j = m - \sum_{\nu=0}^{j-1} \mu_\nu, \quad r_j = r - 2 \sum_{\nu=0}^{j-1} \rho_\nu,
$$

$$
s_j = s \prod_{\nu=0}^{j-1} \eta_{\nu}^2, \quad h_j = 2^{j+c} E_j,
$$

(6.2)

where $\eta_j^2 = 4^{-b} \Gamma_j E_j$. Then $m_j \downarrow m - w$, $r_j \downarrow r - 2\rho$ and $s_j \downarrow 0$, $h_j \downarrow 0$. These sequences define the complex domains

$$
\mathcal{D}_j = \mathcal{D}_{r_j,s_j}, \quad \mathcal{O}_j = \mathcal{O}_{h_j}.
$$

Finally, we introduce an extended phase space norm,

$$
| (\theta, x, J, z, \tilde{\omega}) |_{\tilde{\mathcal{P}}} = \max \left( |\theta|_\infty, |x|, |J|_w, |z|, |\tilde{\omega}| \right),
$$

\(\text{Springer}\)
and the corresponding weight matrices,
\[ \bar{W}_j = \text{diag}(\rho_j^{-1}I_Z, \rho_j^{-1}I_n, 2s_j^{-1}I_Z, 2s_j^{-1}I_n, h_j^{-1}I_n). \]

Then, we can state the iterative lemma.

**Lemma 6.1** (iterative lemma) Suppose that
\[ s_j^{-1}|||H - N|||_{m,r,s,h} \leq \frac{\tilde{\alpha}_* \varepsilon_*}{\Psi_0(\mu)\Psi_1(\rho)} \leq \frac{h}{2^c}, \quad (6.3) \]
where $\tilde{\alpha} = 2$ and $\varepsilon_* = 2^{-c}$. Then for each $j \geq 0$ there exists a normal form $N_j$ and a real analytic transformation
\[ \mathcal{F}^j = \mathcal{F}_0 \circ \mathcal{F}_1 \circ \cdots \circ \mathcal{F}_{j-1} : \mathcal{D}_j \times \mathcal{O}_j \to \mathcal{D}_0 \times \mathcal{O}_0 \]
of the form described in Sect. 4, which is symplectic for each $\tilde{\omega}$, such that $H \circ \mathcal{F}^j = N_j + P_j$ with
\[ s_j^{-1}|||P_j|||_{m_j,r_j,s_j,h_j} \leq E_j. \quad (6.4) \]
Moreover,
\[ |ar{W}_0(\mathcal{F}^{j+1} - \mathcal{F}^j)|_{\tilde{\omega}} \leq 8 \max \left( 2^{-a-j} \Gamma_j E_j, E_j/h_j \right) \]
on $\mathcal{D}_{j+1} \times \mathcal{O}_{j+1}$.

Before giving the proof of iterative lemma 6.1 we collect some useful facts. The $\kappa_v$ satisfy the identities
\[ \sum_{v=0}^{\infty} \kappa_v = 1, \quad \sum_{v=0}^{\infty} v \kappa_v = \frac{1}{\kappa - 1}. \]
This and the monotonicity of the $\Gamma$ -function imply that
\[ \Gamma_j = \prod_{v=j}^{\infty} \Gamma_j^{\kappa_v \kappa_j^j} \leq \left( \prod_{v=j}^{\infty} \Gamma_j^{\kappa_v} \right)^{\kappa_j^j}. \]
Together with the definition of $E_j$ and (6.1), we obtain the estimate
\[ \Gamma_j E_j \leq \left( \prod_{v=0}^{\infty} \Gamma_v^{\kappa_v} E_0 \right)^{\kappa_j^j} \quad \left( 2^{2^j-a} \Psi_0 \Psi_1 E_0 \right)^{\kappa_j^j}. \quad (6.6) \]
Moreover,

\[ \Gamma_j^{\kappa-1} E_j^\kappa = E_{j+1} \]  

(6.7)

by a straightforward calculation.

**Proof of iterative lemma 6.1** Lemma 6.1 is proven by induction. Choosing \( F^0 = \text{id} \) and

\[ E^0 = \tilde{\alpha} \varepsilon_*(\mu) \Psi_0(\mu) \Psi_1(\rho), \]

there is nothing to prove for \( j = 0 \). Just observe that \( h_0 \leq h \) by the definitions of \( h_0 \) and \( E_0 \).

Let \( j \geq 0 \). To apply the KAM step to \( H_j = H \circ F^j \) and \( N_j \), we need to verify its assumptions (5.6), (5.15), (5.17) and (5.19). Clearly, \( m_j - \mu_j \geq w \) by construction, and \( E_j \leq h_j/16 \) in view of the definition of \( h_j \) and \( c \geq 4 \), so the second and third requirements are met. Taking squares, the fourth requirement is equivalent to

\[ 4^{2-a-j} C_0^2 \Gamma_j^2 E_j^2 \leq 4^{-b} \Gamma_j E_j \leq \frac{1}{4}. \]

This holds for all \( j \geq 0 \), since \( C_0 = 8 \),

\[ \Gamma_j E_j \leq 2^{3+a-e} \]

(6.8)

by (6.2), (6.3), (6.6) and \( a \geq b + 2, b \geq 0, e \geq a + 9 \).

As to the first requirement, define \( K_j \) by \( e^{-K_j} = 2^{-d} \Gamma_j E_j \) and subsequently \( \langle \cdot \rangle \) as in (5.3). For arbitrary \( \tilde{A} \) in \( \tilde{S} \) with \( \langle \tilde{A} \rangle > 0 \) we then have

\[ \frac{1}{\langle \tilde{A} \rangle \Delta(\langle \tilde{A} \rangle) \Delta(\langle \tilde{A} \rangle)} = \frac{e^{-\rho_j(\tilde{A})} e^{-\mu_j(\tilde{A})}}{\langle \tilde{A} \rangle \Delta(\langle \tilde{A} \rangle) e^{-\rho_j(\tilde{A})} \Delta(\langle \tilde{A} \rangle) e^{-\mu_j(\tilde{A})}} \geq \frac{e^{-K_j}}{\Gamma_j \mu_j \Gamma_j \rho_j} = \frac{2^{-d} \Gamma_j E_j}{2^{-j+2} \Gamma_j} = 2^{j+a-d} E_j \geq h_j, \]

(6.9)

since \( a \geq c+d \). This estimate holds even more when \( \langle \tilde{A} \rangle = 0 \). Hence, also requirement (5.6) is satisfied.

The KAM construction now provides a normal form \( N_{j+1} \), a coordinate transformation \( \Phi_j \) and a parameter transformation \( \varphi_j \). By the definition of \( r_j \) and \( s_j \), \( \Phi_j \) maps \( D_{j+1} \) into \( D_j \), while \( \varphi_j \) maps \( O_{j+1} \) into \( O_j \), since

\[ \frac{h_{j+1}}{h_j} = \frac{2E_{j+1}}{E_j} = 2(\Gamma_j E_j)^{\kappa-1} \leq 2^{1+(3+a-e)/2} \leq \frac{1}{4}, \]

in view of (6.8) and \( \kappa = 3/2, e \geq a + 9 \). Setting

\[ F_{j+1} = F_j \circ F_j, \quad F_j = \Phi_j \circ \varphi_j, \]
we obtain a transformation $\mathcal{F}^{j+1}$ from $\mathcal{D}_{j+1} \times \mathcal{O}_{j+1}$ into $\mathcal{D}_0 \times \mathcal{O}_0$. For the new error term

$$P_{j+1} = H \circ \mathcal{F}^{j+1} - N_{j+1} = H_j \circ F_j - N_{j+1},$$

we obtain

$$|||P_{j+1}|||_{j+1} \leq 32 \Gamma_{\mu_j} \Gamma_{\rho_j} E_j \varepsilon_j + 2e^{-K_j} \varepsilon_j + 4\eta_j^2 \varepsilon_j$$

$$\leq (2^{5-a} + 2^{1-d} + 2^{2-2b}) \Gamma_j E_j \varepsilon_j.$$ 

Dividing by $s_{j+1} = \eta_j s_j / 2$ this yields

$$s_{j+1}^{-1} |||P_{j+1}|||_{j+1} \leq 2^{1+b} (2^{5-a} + 2^{1-d} + 2^{2-2b}) \Gamma_j^{k-1} E_j^k$$

$$= (2^{6-a+b} + 2^{2+b-d} + 2^{3-b}) E_{j+1}$$

$$\leq E_{j+1},$$

since $\eta_j^2 = 4^{-b} \Gamma_j E_j$, $E_j = \varepsilon_j / s_j$, $a \geq b + 8$, $b \geq 4$, $d \geq b + 4$, $\kappa = 3/2$ and (6.7).

To prove the first of the estimates, write

$$|\tilde{W}_0(\mathcal{F}^{j+1} - \mathcal{F}^j)|_{j+1} = |\tilde{W}_0(\mathcal{F}^j \circ \mathcal{F}_j - \mathcal{F}^j)|_{j+1}$$

$$\leq |\tilde{W}_0 \tilde{D} \mathcal{F}^j \tilde{W}_j^{-1}|_j |\tilde{W}_j(\mathcal{F}_j - id)|_{j+1},$$

(6.10)

where $| \cdot |_j \equiv | \cdot |_{\tilde{p}_j, \mathcal{D}_j \times \mathcal{O}_j}$, where $\tilde{D}$ denotes differentiation with respect to $(x, \theta, z, J, \tilde{\omega})$. By (5.16), (5.18) and the definition of $\Gamma_j$,

$$|\tilde{W}_j(\mathcal{F}_j - id)|_{j+1} \leq \max \left( |W_j(\Phi_j - id)|_p, h_j^{-1} |\varphi_j - id| \right)$$

$$\leq \max \left( 2^{2-a-j} \Gamma_j E_j, 4E_j / h_j \right),$$

(6.11)

It remains to show that the first factor is bounded by 2. By the inductive construction, $\mathcal{F}^j = \mathcal{F}_0 \circ \mathcal{F}_1 \circ \cdots \circ \mathcal{F}_{j-1}$, and

$$|\tilde{W}_v \tilde{D} \mathcal{F}^j \tilde{W}^{-1}_v|_{v+1} \leq \max \left( |W_v \tilde{D} \Phi_v W_v^{-1}|_p + h_v |W_v \phi_v \Phi_v|_p, |\tilde{\partial} \Phi_v| \right)$$

$$\leq \max \left( 1 + 2^{5-a-v} \Gamma_v E_v, 1 + 16E_v / h_v \right)$$

$$\leq 1 + 2^{4-c-v}$$

by (5.16), (5.18), (6.8). Since the weights of $\tilde{W}_v^{-1}$ do not decrease as $v$ decreases, and since $c \geq 6$, we obtain

$$|\tilde{W}_0 \tilde{D} \mathcal{F}^j \tilde{W}_j^{-1}|_j \leq \prod_{v=0}^{j-1} |\tilde{W}_v \tilde{D} \mathcal{F}_v \tilde{W}^{-1}_v|_{v+1} \leq \prod_{v=0}^{\infty} (1 + 2^{4-c-v}) \leq 2.$$ (6.12)
By (6.10), (6.11), (6.12), the conclusion (6.5) holds. This completes the proof of the iterative lemma 6.1.

\[ \Box \]

### 6.2 Convergence

By the estimates of the iterative lemma the \( F_j \) converge uniformly on

\[ \bigcap_{j \geq 0} D_j \times O_j = D_\ast \times O_\ast, \quad D_\ast = D_{r-2\rho,0} \]

to mappings \( F_\ast \) that are real analytic in \( x, \theta \) and uniformly continuous in \( \tilde{\omega} \). Moreover,

\[ |\bar{W}_0(F_\ast - id)|_{\bar{P}} \leq \frac{1}{2} \]

on \( D_\ast \times O_\ast \) by the usual telescoping argument.

But by construction, the \( F_j \) are affine linear in each fiber over \( \mathbb{T}^Z \times \mathbb{T}^n \times O_\ast \).

Therefore, they indeed converge uniformly on any domain \( D_{r-2\rho,\sigma} \times O_\ast \) with \( \sigma > 0 \)
to a map \( F_\ast \) that is real analytic and symplectic for each \( \tilde{\omega} \). In particular,

\[ F_\ast : D_{r-2\rho,s/2} \times O_\ast \to D_{r,s} \times O_h \]

by piecing together the above estimates.

Going to the limit in (6.4) and using Cauchy’s inequality, we finally obtain

\[ H \circ F_\ast = e_\ast + \langle \omega, J \rangle + \langle \tilde{\omega}, z \rangle + \cdots. \]

This completes the proof of Theorem 3.4.

### 6.3 Estimates

The scheme so far provides only a very crude estimate of \( F_\ast \) since the actual size of the perturbation is not taken into account in the estimates of the iterative lemma. But nothing changes when all inequalities are scaled down by the factor \( \varepsilon / E \leq 1 \), where

\[ \varepsilon = s^{-1} \|||H - N|||_{m,r,s,h} \leq E = \frac{\alpha e_\ast}{\Psi_\mu \Psi_\rho}. \]

It follows that

\[ |\bar{W}_0(F_\ast - id)|_{\bar{P}} \leq \frac{\varepsilon}{E} \]

uniformly on \( D_{r-2\rho,s/2} \times O_\ast \).
7 The Measure Estimate

In this section, the measure estimate of the frequency \( \tilde{\omega} \) satisfying the inequalities (3.5) will be given. Firstly, we give some useful lemmas.

**Lemma 7.1** (Lemma 2 in Pöschel (1990)) *For every given approximation function \( \Theta \), there exists an approximation function \( \Delta \) such that*

\[
\sum_{\tilde{A} \in \tilde{S}, \text{card}(\tilde{A}) = i} \frac{1}{\Delta(|\tilde{A}|)} \leq \frac{2N_0}{\Theta(t_i)}, \quad i \geq 1,
\]

*where \( N_0, t_i \) are given in Lemma 2.1.*

**Lemma 7.2** (Lemma 4 in Pöschel (1990)) *There is an approximation function \( \Delta \) such that*

\[
\sum_{\ell \in \mathbb{Z} \setminus \{0\}} \frac{1}{\Delta(|\ell|)} \leq M i \log \log i
\]

*for all sufficiently large \( i \) with some constant \( M \), where \( \ell = (\ell_1, \ell_2, \ldots, \ell_i) \) and \( |\ell| = |\ell_1| + |\ell_2| + \cdots + |\ell_i| \).*

**Remark 7.3** Of course, Lemma 7.2 also gives a bound for all small \( i \), since the left hand side is monotonically increasing with \( i \).

**Remark 7.4** From the proofs of Lemma 7.1 and Lemma 7.2, we know that there is the same approximation function \( \Delta \) such that Lemmas 7.1 and 7.2 hold simultaneously. The detail proofs of Lemmas 7.1 and 7.2 can be found in Pöschel (1990).

**Theorem 7.5** *There is an approximation function \( \Delta \) such that for suitable \( \alpha \), the set \( O_\alpha \) of \( \tilde{\omega} \) satisfying (3.5) has positive measure.*

**Proof** Choose the frequency \( \omega = (\ldots, \omega_\lambda, \ldots) \) satisfying the nonresonance condition (2.5). For any bounded \( O \in \mathbb{R}^n \), let \( O_\alpha \subseteq O \) denote the set of all \( \tilde{\omega} \) satisfying (3.5) with fixed \( \omega \). Then, the complement of the open dense set \( O_\alpha \) is \( \mathcal{R}_\alpha \), where

\[
\mathcal{R}_\alpha = \bigcup_{(k, \tilde{k}) \in \mathbb{Z}^n_S \setminus \{\tilde{k} = 0\}} \mathcal{R}^{k, \tilde{k}}_\alpha,
\]

\[
:= \bigcup_{(k, \tilde{k}) \in \mathbb{Z}^n_S \setminus \{\tilde{k} = 0\}} \left\{ \tilde{\omega} \in \mathbb{R}^n : \left| \langle k, \omega \rangle + \langle \tilde{k}, \tilde{\omega} \rangle \right| < \frac{\alpha}{\Delta([[(k, \tilde{k})]]) \Delta(|k| + |\tilde{k}|)} \right\}.
\]

Now we estimate the measure of the set \( \mathcal{R}^{k, \tilde{k}}_\alpha \). Since \( (k, \tilde{k}) \in \mathbb{Z}^n_S \setminus \{\tilde{k} = 0\} \), then \( \tilde{k} \neq 0 \), set \( |\tilde{k}_{\text{max}}| = \max_{1 \leq i \leq n} |\tilde{k}_i| \neq 0 \), then there exists some \( 1 \leq j \leq n \) such that \( |\tilde{k}_j| = |\tilde{k}_{\text{max}}| \). Therefore, we have

\[
\mathcal{R}^{k, \tilde{k}}_\alpha = \left\{ \tilde{\omega} \in \mathbb{R}^n : \left| \langle k, \omega \rangle + \langle \tilde{k}, \tilde{\omega} \rangle \right| < \frac{\alpha}{\Delta([[(k, \tilde{k})]]) \Delta(|k| + |\tilde{k}|)} \right\}
\]
= \left\{ \tilde{\omega} \in \mathbb{R}^n : |\tilde{k}_{\max} \tilde{\omega}_j + \sum_{i \neq j} \tilde{k}_i \tilde{\omega}_i + \langle k, \omega \rangle | < \frac{\alpha}{\Delta(|\{(k, \tilde{k})\}|) \Delta(|k| + |\tilde{k}|)} \right\}

= \left\{ \tilde{\omega} \in \mathbb{R}^n : |\tilde{\omega}_j + b_{k, \tilde{k}}| < \delta_{k, \tilde{k}} \right\}

= \left\{ \tilde{\omega} \in \mathbb{R}^n : -b_{k, \tilde{k}} - \delta_{k, \tilde{k}} < \tilde{\omega}_j < -b_{k, \tilde{k}} + \delta_{k, \tilde{k}} \right\},

where \( b_{k, \tilde{k}} = \frac{1}{\tilde{k}_{\max}} \left( \sum_{i \neq j} \tilde{k}_i \tilde{\omega}_i + \langle k, \omega \rangle \right) \) and \( \delta_{k, \tilde{k}} = \frac{\alpha}{\Delta(|\{(k, \tilde{k})\}|) \Delta(|k| + |\tilde{k}|)} \frac{1}{\tilde{k}_{\max}}. \)

Obviously, for any bounded domain \( \mathcal{O} \subset \mathbb{R}^n \), we have the Lebesgue measure estimate

\[
\text{meas}(\mathcal{R}_{\alpha}^{k, \tilde{k}} \cap \mathcal{O}) \leq C_0 \delta_{k, \tilde{k}} = \frac{C_0 \alpha}{\Delta(|\{(k, \tilde{k})\}|) \Delta(|k| + |\tilde{k}|)} \frac{1}{|\tilde{k}_{\max}|}
\]

with some positive constant \( C_0. \)

Since \( |\tilde{k}_{\max}| \neq 0, \tilde{k}_{\max} \in \mathbb{Z} \), which means \( |\tilde{k}_{\max}| \geq 1 \), then we have the following measure estimate

\[
\alpha^{-1} \text{meas}(\mathcal{R}_{\alpha}^{k, \tilde{k}} \cap \mathcal{O}) \leq \frac{C_0}{\Delta(|\{(k, \tilde{k})\}|) \Delta(|k| + |\tilde{k}|)}.
\]

Next we estimate the measure of the set \( \mathcal{R}_{\alpha}. \) From the definition of \( \mathbb{Z}_\mathcal{S} \), there exists a nonempty set \( \tilde{A} \in \tilde{\mathcal{S}} \) such that \( \text{supp}(k, \tilde{k}) \subseteq \tilde{A} \), we get

\[
\alpha^{-1} \text{meas}(\mathcal{R}_{\alpha} \cap \mathcal{O}) \leq \sum_{\tilde{A} \in \mathcal{S}} \sum_{\tilde{k} \neq 0} \alpha^{-1} \text{meas}(\mathcal{R}_{\alpha}^{k, \tilde{k}} \cap \mathcal{O})
\]

\[
\leq C_0 \sum_{\tilde{A} \in \mathcal{S}} \sum_{\tilde{k} \neq 0} \frac{1}{\Delta(|\{(k, \tilde{k})\}|) \Delta(|k| + |\tilde{k}|)}
\]

\[
\leq C_0 \sum_{\tilde{A} \in \mathcal{S}} \left( \frac{1}{\Delta(|\tilde{A}|)} \sum_{\text{supp}(k, \tilde{k}) \subseteq \tilde{A}} \frac{1}{\Delta(|k| + |\tilde{k}|)} \right)
\]

\[
\leq C_0 \sum_{\tilde{A} \in \mathcal{S}} \left( \frac{1}{\Delta(|\tilde{A}|)} \sum_{\text{card}(\tilde{A})=i} \sum_{\ell \in \mathbb{Z} \setminus \{0\}} \frac{1}{\Delta(|\ell|)} \right)
\]

Thus, the sum is broken up with respect to the cardinality and the weight of the spatial components of \( \tilde{\mathcal{S}} \). Each of these factors is now studied separately.
By applying Lemmas 7.1, 7.2 and Remark 7.3, we arrive at
\[ \alpha^{-1} \text{meas}(\mathcal{R}_\alpha \cap \mathcal{O}) \leq C + C \sum_{i=i_0}^{+\infty} \frac{M_i \log \log i}{\Theta(t_i)} \]
with some constant \( C \) and \( i_0 \) so large that \( t_i \geq i \log \log i \) for \( i \geq i_0, \varrho > 2 \) by hypotheses. Here we are still free to choose a suitable approximation function \( \Theta \), and choose
\[ \Theta(t) = \exp \left( \frac{t}{\log t \log^{\varrho-1} \log t} \right), \quad t > e, \varrho > 2, \]
the infinite sum does converge. Thus, there is an approximation function \( \Delta \) such that
\[ \alpha^{-1} \text{meas}(\mathcal{R}_\alpha \cap \mathcal{O}) < +\infty. \]
Hence,
\[ \text{meas}(\mathcal{R}_\alpha \cap \mathcal{O}) \leq O(\alpha) \]
and
\[ \mathcal{O}_\alpha \to \mathcal{O} \quad \text{as} \quad \alpha \to 0. \]
This completes the proof of Theorem 7.5. \qed

8 Application

In this section, we will apply Theorem 3.4 to the differential equation
\[ \ddot{x} + x^{2l+1} = \varepsilon \sum_{j=0}^{2l} p_j(t)x^j, \quad (8.1) \]
where \( p_0, p_1, \ldots, p_{2l} \) are real analytic almost periodic functions with the frequency \( \omega = (\ldots, \omega_k, \ldots) \) and admit the spatial series expansion similar to (2.6).

8.1 Rescaling

We first rescale the time variable \( t \) and the space variable \( x \) to get a slow system. Let \( u = \varepsilon x, \tau = \varepsilon^{-l} t \). Then, Equation (8.1) becomes
\[ \ddot{u} + u^{2l+1} = \varepsilon \sum_{j=0}^{2l} \varepsilon^{2l-j} p_j(\tau)u^j, \quad (8.2) \]
where \( \ddot{u} \) stands for \( \frac{d^2u}{d\tau^2} \), \( p_0, p_1, \ldots, p_{2l} \) are real analytic almost periodic functions in \( \tau \) with the frequency \( \hat{\omega} = \varepsilon \omega \). Without causing confusion, in the following we still use \( t \) instead of \( \tau \). Equation (8.2) is equivalent to the following Hamiltonian system

\[
\begin{cases}
\dot{u} = v, \\
\dot{v} = -u^{2l+1} + \varepsilon \sum_{j=0}^{2l} \varepsilon^{2l-j} p_j(t)u^j,
\end{cases}
\] (8.3)

and the corresponding Hamiltonian is

\[
h(u, v, t) = \frac{1}{2} v^2 + \frac{1}{2l + 2} u^{2l+2} - \varepsilon \sum_{j=0}^{2l} \frac{\varepsilon^{2l-j}}{j+1} p_j(t)u^{j+1}. \] (8.4)

It is obvious that (8.4) is a perturbation of the integrable Hamiltonian

\[
h_0(u, v) = \frac{1}{2} v^2 + \frac{1}{2l + 2} u^{2l+2} \] (8.5)

for \( \varepsilon > 0 \) small. Our aim is to construct, for every sufficiently small \( \varepsilon > 0 \), invariant cylinders tending to the infinity for (8.4) close to \( \{h_0(u, v) = C\} \times \mathbb{R} \) in the extended phase space, which prohibit any solution from going to the infinity. For this purpose, we will introduce the action-angle variables first.

### 8.2 Action and Angle Variables

We consider the following integrable Hamiltonian system

\[
\begin{cases}
\dot{u} = v, \\
\dot{v} = -u^{2l+1}
\end{cases}
\] (8.6)

with the Hamiltonian (8.5). Suppose \( (C(t), S(t)) \) is the solution of (8.6) satisfying the initial condition \( (C(0), S(0)) = (1, 0) \). Let \( T_\ast > 0 \) be its minimal period, which is a constant. Then, these analytic functions \( C(t), S(t) \) satisfy

(i) \( C(t + T_\ast) = C(t), S(t + T_\ast) = S(t) \) and \( C(0) = 1, S(0) = 0 \);
(ii) \( \dot{C}(t) = S(t), \dot{S}(t) = -C^{2l+1}(t) \);
(iii) \( (l + 1)S^2(t) + C^{2l+2}(t) = 1 \);
(iv) \( C(-t) = C(t), S(-t) = -S(t) \).

The action and angle variables are now defined by the map \( \Psi : \mathbb{R}^+ \times \mathbb{T} \to \mathbb{R}^2 \setminus \{0\} \) via \( (\varrho, \phi) = \Psi(u, v) \) is given by the formula

\[
\Psi : \begin{align*}
u &= c_1 \frac{1}{\sqrt{2\pi}} \varrho^{1/2} C\left(\frac{T_\ast \phi}{2\pi}\right), \\
v &= c_1 \frac{1}{\sqrt{2\pi}} \varrho^{1/2} S\left(\frac{T_\ast \phi}{2\pi}\right)
\end{align*}
\]
with \( c_1 = \frac{2\pi(l+2)}{l} > 0, \varrho > 0 \) and \( \phi \in \mathbb{T} \). We can check that \( \Psi \) is a symplectic diffeomorphism from \( \mathbb{R}^+ \times \mathbb{T} \) onto \( \mathbb{R}^2 \setminus \{0\} \). Under this transformation, the Hamiltonian (8.4) becomes

\[
H(\varrho, \phi, t) = h(\Psi(\varrho, \phi), t)
= \frac{1}{2l + 2} c_1 \frac{2l+2}{l+2} \varrho \frac{2j+2}{l+2} - \varepsilon \sum_{j=0}^{2l} \frac{\varepsilon^{2l-j} c_1}{j+1} \frac{l+1}{l+2} \varrho^{l+1} C^{j+1} \left( T_s \phi \right) P_j(t). \tag{8.7}
\]

After introducing two conjugate variables \( \theta \in \mathbb{T}^\mathbb{Z} \) and \( J \in \mathbb{R}^\mathbb{Z} \), the Hamiltonian (8.7) can be written in the form of an autonomous Hamiltonian as follows

\[
H = \langle \dot{\omega}, J \rangle + \frac{1}{2l + 2} c_1 \frac{2l+2}{l+2} \varrho \frac{2j+2}{l+2} - \varepsilon \sum_{j=0}^{2l} \frac{\varepsilon^{2l-j} c_1}{j+1} \frac{l+1}{l+2} \varrho^{l+1} C^{j+1} \left( T_s \phi \right) P_j(\theta),
\]

where \( P_j(\theta) \) is the shell function of the almost periodic function \( p_j(t) \).

Let \([C_1, C_2] \subseteq \mathbb{R}^+\) be any bounded interval without 0, not depending on \( \varepsilon \). For any \( \varrho_0 \in [C_1, C_2] \), we denote \( \varrho = \varrho_0 + I \) and do Taylor expansion at \( \varrho_0 \) for \( |I| < \varrho_0/2 \). Then, we have

\[
H(\theta, \phi, J, I) = \langle \dot{\omega}, J \rangle + \frac{1}{2l + 2} c_1 \frac{2l+2}{l+2} \varrho_0 \frac{2j+2}{l+2} + \frac{1}{l+2} c_1 \frac{2l+2}{l+2} \varrho_0^{l+1} I + O(I^2)
- \varepsilon \sum_{j=0}^{2l} \frac{\varepsilon^{2l-j} c_1}{j+1} \frac{l+1}{l+2} \varrho_0^{l+1} C^{j+1} \left( T_s \phi \right) P_j(\theta). \tag{8.8}
\]

Denote \( \tilde{\omega} = \frac{1}{l+2} c_1 \frac{2l+2}{l+2} \varrho_0^{l+1} I \), for any \( \varrho_0 \in [C_1, C_2] \), we get

\[
\frac{\partial \tilde{\omega}}{\partial \varrho_0} = \frac{l}{(l+2)^2} c_1 \frac{2l+2}{l+2} \varrho_0^{l+1} \neq 0.
\]

We therefore single out the subsets \( \mathcal{O}_\alpha \subseteq \mathcal{O} := \left[ \frac{1}{l+2} c_1 \frac{2l+2}{l+2} \varrho_0^{l+1}, \frac{1}{l+2} c_1 \frac{2l+2}{l+2} \varrho_0^{l+1} \right] \) which consists of \( \tilde{\omega} \) satisfying

\[
|\langle k, \varepsilon \omega \rangle + \tilde{k} \tilde{\omega} | \geq \frac{\alpha}{\Delta([|k|+1])} \Delta(|k|+1)\Delta(k) \for all (k, \tilde{k}) \in \mathbb{Z}^2 \setminus \{ \tilde{k} \neq 0 \}
\]

with fixed \( \omega \). From the measure estimate in Sect. 7 with \( n = 1 \), it follows that the set \( \mathcal{O}_\alpha \) is a set of positive Lebesgue measure provided that \( \alpha \) is small. Let \( \mathcal{O}_h : |\tilde{\omega} - \mathcal{O}_\alpha| < h \) denote complex neighborhoods of \( \mathcal{O}_\alpha \).

Similar to Sect. 3, we can rewrite the Hamiltonian (8.8) as

\[
H(\theta, \phi, J, I; \tilde{\omega}) = N + P
= e(\tilde{\omega}) + \langle \tilde{\omega}, J \rangle + \tilde{\omega} I + P(\theta, \phi, I; \tilde{\omega}), \tag{8.9}
\]
where
\[ P(\theta, \phi; \tilde{\omega}) = \frac{1}{2l+2} c_{\frac{l}{l+2}}^{2l+2} \left( (q_0 + I)^{\frac{2l+2}{l+2}} - q_0^{\frac{2l+2}{l+2}} - \frac{2l+2}{l+2} q_0^{\frac{l}{l+2}} I \right) \\
- \varepsilon \sum_{j=0}^{2l} \frac{\varepsilon^{2l-j}}{j+1} c_{\frac{l}{l+2}}^{j+1} (q_0 + I)^{j+1} C^{j+1} \left( \frac{T_s \phi}{2\pi} \right) P_j(\theta) \]

with \( q_0 = g_{\tilde{\omega}}(\tilde{\omega}) \).

Hence, \( P \) is periodic in \( \theta, \phi \) with the period \( 2\pi \), and real analytic in \( (\theta, \phi, I) \in \mathbb{T}^2 \times \mathbb{T} \times \mathbb{R} \). Then, there exists \( r > 0 \), such that \( P \) admits analytic extension in the complex neighborhood \( \{ (\theta, \phi) : |\text{Im}\ \theta| < r, |\text{Im}\ \phi| < r \} \) of \( \mathbb{T}^2 \). Taking \( s = \varepsilon^{1/2} \), then there exists \( C_* > 0 \), depending on \( l, T, r, h \), but not on \( \varepsilon \), such that for any \( |\text{Im}\ \theta| < r, |\text{Im}\ \phi| < r, |I| < s, \tilde{\omega} \in \mathcal{O}_h \), we have \( |||P|||_{m, r, s, h} \leq C_* \varepsilon \). Without losing of generality, we can assume \( |||P|||_{m, r, s, h} \leq \varepsilon \), which means \( s^{-1} |||P|||_{m, r, s, h} \leq \varepsilon^{1/2} \).

### 8.3 The Main Result

**Theorem 8.1** Every solution of (8.1) with real analytic almost periodic functions \( p_0, p_1, \ldots, p_{2l} \in AP_r(\omega) \), \( \omega \) satisfying the nonresonance condition (2.5) is bounded. Moreover, (8.1) has infinitely many almost periodic solutions.

**Proof** If the conditions of Theorem 8.1 hold and \( 0 < \varepsilon < \left( \frac{\alpha \varepsilon_*}{\Psi_0(\mu) \Psi_1(\rho)} \right)^2 \), then
\[ s^{-1} |||P|||_{m, r, s, h} \leq \varepsilon^{1/2} < \frac{\alpha \varepsilon_*}{\Psi_0(\mu) \Psi_1(\rho)} \]

for some \( 0 < \mu \leq m - w \) and \( 0 < \rho < r/2 \), where \( \varepsilon_* = 2^{-22} \) is an absolute positive constant, \( \Psi_0(\mu) \Psi_1(\rho) \) is defined by (6.1), therefore the assumptions of Theorem 3.4 are met. Hence, the existence of the invariant tori of the Hamiltonian system (8.9) is guaranteed by Theorem 3.4, the Hamiltonian system (8.9) has a real analytic invariant torus of maximal dimension and with a vector field conjugate to \( (\varepsilon^l \omega, \tilde{\omega}) \) for each frequency vector \( \tilde{\omega} \in \mathcal{O}_h \). Each of these tori produces a family of almost periodic solution with the frequency \( (\varepsilon^l \omega, \tilde{\omega}) \), all solutions with initial datum lie in the interior of some invariant cylinders, which implies that all solutions are bounded for all time. Then, system (8.1) has infinitely many almost periodic solutions as well as the boundedness of solutions.

**Remark 8.2** It follows from the proof of Theorem 8.1 that if the conditions of Theorem 8.1 hold, then system (8.1) has infinitely many almost periodic solutions with the frequency \( \{ \omega, \tilde{\omega} \} \) for each frequency vector \( \tilde{\omega} \in \varepsilon^{-l} \mathcal{O}_\alpha \).
References

Arnold, V.I.: Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian. Uspekhi Mat. Nauk 18, 13–40 (1963)
Arnold, V.I.: Mathematical Methods of Classical Mechanics. Springer, Berlin (1978)
Cheng, C.Q., Sun, Y.S.: Existence of KAM tori in degenerate Hamiltonian systems. J. Differ. Equ. 114, 288–335 (1994)
Chow, S.-N., Li, Y., Yi, Y.: Persistence of invariant tori on submanifolds in Hamiltonian systems. J. Nonlinear Sci. 12, 585–617 (2002)
Dieckerhoff, R., Zehnder, E.: Boundedness of solutions via the twist theorem. Ann. Sci. Norm. Sop. Pisa. Cl. Sci. 14, 79–95 (1987)
Dineen, S.: Complex Analysis on Infinite Dimensional Spaces. Springer, Berlin (1999)
Eliasson, L.H.: Perturbations of stable invariant tori for Hamiltonian systems. Ann. Sc. Norm. Super. Pisa 15, 115–147 (1988)
Huang, P., Li, X., Liu, B.: Invariant curves of smooth quasi-periodic mappings. Discrete Contin. Dyn. Syst. 38, 131–154 (2018)
Huang, P., Li, X., Liu, B.: Quasi-periodic solutions for an asymmetric oscillation. Nonlinearity 29, 3006–3030 (2016)
Huang, P., Li, X., Liu, B.: Invariant curves of almost periodic twist mappings. Preprint arXiv:1606.08938 (2016)
Huang, P., Li, X., Liu, B.: Almost periodic solutions for an asymmetric oscillation. J. Differ. Equ. 263, 8916–8946 (2017)
Kolmogorov, A.N.: On the conservation of conditionally periodic motions for a small change in Hamilton’s function. Dokl. Akad. Nauk SSSR 98, 527–530 (1954). (in Russian)
Littlewood, J.E.: Some Problems in Real and Complex Analysis. Heath, Lexington (1968)
Levi, M.: Quasiperiodic motions in superquadratic time-periodic potentials. Commun. Math. Phys. 143, 43–83 (1986)
Levi, M., Zehnder, E.: Boundedness of solutions for quasiperiodic potentials. SIAM J. Math. Anal. 26, 1233–1256 (1995)
Laederich, S., Levi, M.: Invariant curves and time-dependent potentials. Ergodic Theory Dyn. Syst. 11, 365–378 (1991)
Liu, B., You, J.: Quasiperiodic solutions of Duffing’s equations. Nonlinear Anal. 33, 645–655 (1998)
Morris, G.R.: A case of boundedness in Littlewood’s problem on oscillatory differential equations. Bull. Austral. Math. Soc. 14, 71–93 (1976)
Moser, J.: On invariant curves of area-preserving mappings of an annulus. Nachr. Akad. Wiss. Göttingen Math. Phys. II, 1–20 (1962)
Moser, J.: Convergent series expansions for quasi-periodic motions. Math. Ann. 169, 136–176 (1967)
Moser, J.: Stable and Random Motions in Dynamical Systems. Annals of Mathematics Studies, 47. Princeton University Press, Princeton (1973)
Norris, J.: Boundedness in periodically forced second order conservative systems. J. Lond. Math. Soc. 45, 97–112 (1992)
Pöschel, J.: On elliptic lower dimensional tori in Hamiltonian systems. Math. Z. 202, 559–608 (1989)
Pöschel, J.: Small divisors with spatial structure in infinite dimensional Hamiltonian systems. Commun. Math. Phys. 127, 351–393 (1990)
Pöschel, J.: A lecture on the classical KAM theorem. Proc. Symp. Pure Math. 69, 701–732 (2001)
Rüssmann, H.: Nondegeneracy in the perturbation theory of integrable dynamical systems. In: Stochastics Algebra and Analysis in Classical and Quantum Dynamics (Marseille 1988) Math. Appl. vol. 59, pp. 211–223. Kluwer Academic Publishers, Dordrecht (1990)
Rüssmann, H.: Invariant tori in non-degenerate nearly integrable Hamiltonian systems. Regul. Chaotic Dyn. 6, 119–204 (2001)
Sevryuk, M.B.: KAM-stable Hamiltonians. J. Dyn. Control Syst. 1, 351–366 (1996)
Sevryuk, M.B.: Partial preservation of frequencies in KAM theory. Nonlinearity 19, 1099–1140 (2006)
Siegel, C., Moser, J.: Lectures on Celestial Mechanics. Springer, Berlin (1997)
Wang, J., You, J.: Boundedness of solutions for non-linear quasi-periodic differential equations with Liouvillean frequency. J. Differ. Equ. 261, 1068–1098 (2016)
Xu, J., You, J., Qiu, Q.: Invariant tori for nearly integrable Hamiltonian systems with degeneracy. Math. Z. 226, 375–387 (1997)
Xu, J., You, J.: Persistence of the non-twist torus in nearly integrable Hamiltonian systems. Proc. Am. Math. Soc. 138, 2385–2395 (2010)
You, J.: Invariant tori and Lagrange stability of pendulum-type equations. J. Differ. Equ. 85, 54–65 (1990)
Yuan, X.: Invariant tori of Duffing-type equations. J. Differ. Equ. 142, 231–262 (1998)
Zhang, D., Cheng, R.: On invariant tori of nearly integrable Hamiltonian systems with quasiperiodic perturbation. Fixed Point Theory Appl. 17, 2385–2395 (2010)