Commun. Math. Phys. 127, 479–528 (1990)

Communications in Mathematical Physics

Periodic and Quasi-periodic Solutions of Nonlinear Wave Equations via KAM Theory

C. Eugene Wayne*
Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA

Abstract. In this paper the nonlinear wave equation

$$u_{tt} - u_{xx} + v(x)u(x,t) + \varepsilon u^3(x,t) = 0$$

is studied. It is shown that for a large class of potentials, $v(x)$, one can use KAM methods to construct periodic and quasi-periodic solutions (in time) for this equation.

I. Introduction

This paper studies the non-linear wave equation

$$u_{tt}(x,t) - u_{xx}(x,t) + v(x)u(x,t) + \varepsilon u^3(x,t) = 0,$$
$$0 \leqq x \leqq 1, \quad t \geqq 0; \quad u(0,t) = u(1,t) = 0; \quad v \in L^2[0,1]. \tag{1.1}$$

We show that for a large class of potentials, v, one can construct periodic and quasi-periodic solutions for (1.1), provided ε is small, using a variant of the Kolmogorov, Arnold, Moser [KAM] scheme. The method allows one to study more general non-linear terms than the cubic term in (1.1). For a discussion of the types of non-linearities that are permitted, see Sect. 2.

The existence of solutions, periodic in time, for non-linear wave equations has been studied by many authors. (See [B] for a review of these results. [BN] contains an extensive bibliography.) A wide variety of methods have been brought to bear on the problem, ranging from bifurcation theory, (see for example [H]), to variational techniques, pioneered by Rabinowitz [R], to ideas which exploit the hamiltonian structure of the problem.

The KAM techniques are somewhat complementary to these approaches. They are local methods in that they can only be applied if ε is small, (or equivalently to construct solutions $u(x,t)$ of small norm) whereas the variational methods often yield global results. On the other hand the variational techniques place very strong restrictions on the allowed periods of the solutions. The period, in time, must be

* Supported in part by NSF grants DMS-86-02001 and DMS-88-02118

a rational multiple of the length of the interval in x. This restriction results from the inability of these methods to deal with "small denominators" which arise when the period (in time) is irrationally related to the length of the interval in x. The KAM theory, on the other hand, was developed specifically to deal with such small denominator problems, and hence these restrictions on the period are absent in the present case.

In addition, previous techniques do not seem to give any information about the existence of quasi-periodic solutions, but using KAM ideas they become an easy extension of the periodic problem.

The KAM theory was developed to treat perturbations of integrable hamiltonian systems. Formal perturbation theories had existed for decades for such problems, but the convergence of these expansions was unresolved until Kolmogorov, Arnold and Moser showed how the convergence could be "accelerated." In its classical form the KAM theory only applied to systems with finitely many degrees of freedom. Recently, however, progress has been made in extending the theory to certain infinite dimensional cases. Thus far the systems studied have arisen mostly in condensed matter physics. In [FSW, BV, and P1], for instance, infinite dimensional invariant tori were constructed for some approximate models of anharmonic, disordered, crystals. More recently, in [AF and AFS], bifurcation theory ideas were combined with accelerated convergence techniques to prove the existence of periodic orbits in other, more realistic, models of these phenomena.

In this paper we seek to extend the KAM ideas to a different infinite dimensional setting, namely, perturbations of completely integrable partial differential equations. Nikolenko [N] has applied similar ideas in his study of normal forms for evolution equations, but here we wish to focus on the hamiltonian nature of the problem. We hope that this will shed some light on related, but more difficult, problems such as perturbations of the KdV equations. We use the KAM ideas to construct finite dimensional tori in the infinite dimensional phase space of this system. Our method was influenced by Eliasson's work [E] on constructing low dimensional tori for nearly integrable hamiltonian systems with finitely many degrees of freedom. Eliasson's work has been extended by Rüssmann [Ru] and Pöschel [P2]. In fact, since the results of this paper were presented at the Oberwolfach meeting in May 1987, Pöschel has been able to show that his construction of these low dimensional tori can be extended to treat (1.1). The results he obtained are apparently similar to those presented here, but an exact statement of them is not contained in [P2]. While it seems like a very interesting question to inquire whether or not there are infinite dimensional invariant tori in these systems, corresponding to quasi-periodic motions with infinitely many frequencies, I do not yet have any idea how to prove or disprove their existence.

Our method is based on a perturbation of the known solutions of the $\varepsilon = 0$ case of (1.1). As such it will clearly be important to know the eigenfunctions and eigenvalues of the Sturm-Liouville operator $L_v = d^2/dx^2 - v$, with Dirichlet boundary conditions. In particular we will need to know what sequences of numbers can occur as the eigenvalues of L_v. This constitutes the inverse spectral theory of such operators, a topic which has been studied in [Bo and GL], and recently cast in a clear form by Pöschel and Trubowitz, [PT], to which we shall often refer.

II. Statement of Results

We begin by remarking on why the particular form of Eq. (1.1) was chosen. Suppose we began with a general non-linear wave equation

$$u_{tt} - u_{xx} + f(x,u) = 0; \quad u(0,t) = u(1,t) = 0.$$

Since we will be perturbing about a solution of the linear problem, we expand $f(x,u)$ in a Taylor series about $u \equiv 0$. Thus,

$$f(x,u) = f(x,0) + f_u(x,0)u + \text{higher order terms}. \tag{2.1}$$

We assume $f(x,0) = 0$. Physically this means that there is no force acting when the string is at rest, tending to distort its equilibrium of $u \equiv 0$. We define $f_u(x,0) = v(x)$—we need not assume much smoothness on $v(x)$, so we merely require that it be in $L^2[0,1]$. The results we obtain below would not be changed if we assumed more smoothness for v.

Finally, we assume that the higher order terms in (2.1) have the specific form $\varepsilon u^3(x,t)$. The small parameter ε is introduced just by rescaling u. The choice of a u^3 non-linearity was motivated just by convenience. From the discussion that follows it is clear that any polynomial nonlinearity could equally well be dealt with. However, it is not clear how, or if, one could treat a general analytic interaction term.

Having now motivated the form of (1.1), let us state our results. We will consider cases where the potential, $v(x)$, lies in the subspace of L^2, given by

$$E_0 = \left\{ v \in L^2[0,1] \,\middle|\, \int_0^1 v(x)dx = 0,\ v(x) = v(1-x) \right\}.$$

There is nothing special about fixing the average of v to be zero. Changing the average of v just shifts the eigenfrequencies of the $\varepsilon = 0$ case of (1.1) by a fixed amount.

We restrict our attention further to even functions in an attempt to simplify the description slightly. We select the allowed potentials by placing restrictions on the eigenvalues of the associated operator L_v. In general the set of $v(x)$ giving rise to a chosen set of eigenvalues forms an analytic submanifold of $L^2[0,1]$([PT], Chap. 4). However, this manifold intersects E_0 in a unique point, and it is somewhat easier to deal with this point, rather than the whole submanifold.

We will, for reasons discussed below, restrict our attention to that subset of E_0 for which the spectrum of $L_v = d^2/dx^2 - v$ is negative. In Sect. 9 we construct a probability measure on this subset and show that our results hold for almost every v with respect to this measure.

Let $\{\psi^j\}_{j=1}^{\infty}$ be the eigenfunctions of $L_v = d^2/dx^2 - v$ (with Dirichlet boundary conditions) and $\{\mu_j\}_{j=1}^{\infty}$ the corresponding eigenvalues. Then $u_j^0(x,t) = \sin(\sqrt{|\mu_j|}t)\psi^j(x)$ is a periodic solution of (1.1) with $\varepsilon = 0$ provided $\mu_j < 0$, and $\varphi_N^0(x,t) = \sum_{j=1}^{N} u_j(x,t)$ is a quasi-periodic solution provided $\{\sqrt{|\mu_j|}\}_{j=1}^{N}$ are irrationally related to one another. We wish to prove that these solutions persist when $\varepsilon \neq 0$. That is the content of

Theorem 2.1. *There are sets $\mathscr{E}_0(j) \subset E_0$ such that if $v \in \mathscr{E}_0(j)$, there is a constant $\varepsilon_0(v,j) > 0$, such that whenever $|\varepsilon| < \varepsilon_0(v,j)$ Eq. (1.1) has a weak periodic solution $u_j^\varepsilon(x,t)$ whose frequency differs from $\sqrt{|\mu_j|}$ by $\mathcal{O}(\varepsilon)$. There is a natural probability measure, P, defined on the subset of E_0 for which the spectrum of $L_v = d^2/dx^2 + v$ is negative and $\mathscr{E}_0(j)$ has measure one with respect to P.*

For the quasi-periodic case one has an analogous result.

Theorem 2.2. *There are sets $\mathscr{F}_0(N) \subset E_0$ such that if $v \in \mathscr{F}_0(N)$, there is a constant $\varepsilon_0(v,N) > 0$ such that whenever $|\varepsilon| < \varepsilon_0(v,N)$ Eq. (1.1) has a weak, quasi-periodic solution $\varphi_N^\varepsilon(x,t)$ whose frequency vector differs from that of $\varphi_N^0(x,t)$ by $\mathcal{O}(\varepsilon)$. The set $\mathscr{F}_0(N)$ has measure one with respect to the probability measure, P, introduced in Theorem 2.1.*

Remark. By a weak, quasi-periodic solution of (1.1) we mean a non-zero, continuous function, quasi-periodic in time, which is a distributional solution of (1.1).

Remark. If $u(x,t)$ is a quasi-periodic function of t, with N independent frequencies, there is a function $\tilde{u}(x,\varphi):[0,1] \times \mathbb{T}^N \to \mathbb{R}$, and a vector $\Omega \in \mathbb{R}^N$ such that $u(x,t) = \tilde{u}(x,\Omega t)$. In this last expression we interpret Ωt as an element of $\mathbb{T}^N$ by taking each of its components (mod 1).) The vector Ω is the frequency vector of u. For $\varphi_N^0, \Omega = (\sqrt{|\mu_1|}, \ldots, \sqrt{|\mu_N|})$.

Unfortunately, the vector Ω is not uniquely defined. One can take different linearly independent combinations of its components, which defines a set of frequency vectors known as the frequency module, all of which correspond to the same quasi-periodic orbit. In our case, however, we are perturbing a fixed quasi-periodic orbit of the unperturbed problem, φ_N^0, so it will be clear to which element of the frequency module we are referring. Note that Theorem 2.1 does not establish that the solutions constructed there actually have periods that are irrational, as we claimed in the introduction. One can establish that by keeping track of the changes in frequency produced by the iterative scheme used to prove the theorem. However, the existence of solutions with irrational period is guaranteed by the following result which in addition proves that one obtains a large number of solutions, which is typical for KAM methods.

Theorem 2.3. *Suppose we consider (1.1). If we restrict attention to the subset of E_0 for which the spectrum of L_v is negative, then for almost every potential, v, (with respect to the measure P) there is a constant $\varepsilon_0(v) > 0$ such that if $|\varepsilon| < \varepsilon_0(v)$, there is an interval of the real line, $A(v)$, and a subset $B \subset A(v)$ with* $\operatorname{meas}(B) \geqq (1 - \mathcal{O}(1/|\log\varepsilon|))\operatorname{meas}(A(v))$ *such that for every point $\Omega \in B$, Eq. (1.1) has a periodic orbit with frequency Ω. Similarly, there exists a set of positive, N-dimensional Lebesgue measure such that for every point $\tilde{\Omega}$ in this set one has a quasi-periodic solution with frequency vector $\tilde{\Omega}$.*

Remark. Although Theorem 2.3 gives one many periodic solutions, one has no information about whether or not they occur in smooth families, or even whether there is one corresponding to each value of the energy as is the case when one has only finitely many degrees of freedom. One also has nothing like the results of Weinstein and Moser [W, M], which guarantee (again in the case of finitely many degrees of freedom) one periodic orbit in the non-linear problem corresponding to each periodic orbit in the linear problem.

III. Reduction to a Hamiltonian System with Countably Many Degrees of Freedom

In the present section we show how the results of the previous section can be reduced to questions about a hamiltonian system with countably many degrees of freedom. To this end, we will need to know the eigenvalues and eigenvectors of the differential operators

$$L_v = \frac{d^2}{dx^2} - v, \tag{3.1}$$

acting on $L^2[0,1]$, with Dirichlet boundary conditions at $x=0$ and $x=1$. The spectral theory of such operators has been extensively studied and is reviewed in a form particularly useful for our purposes in [PT].

Suppose we restrict ourselves to the set of potentials, E_0, defined in the previous section. Then combining Theorems 4 and 7 of Chap. 2 of [PT] we obtain

Theorem 3.1. *The eigenfunctions $\{\psi^n\}_{n\geqq 1}$ of L_v form a complete, orthogonal, set in $L^2[0,1]$. The eigenvalues $\{\mu_n\}_{n\geqq 1}$ form a decreasing sequence and obey the asymptotic estimate*

$$\mu_n = -n^2\pi^2 + l^2(n), \quad n = 1, 2, \ldots, \tag{3.2}$$

where $l^2(n)$ denotes the n^{th} component of an element in l^2.

More remarkably, any decreasing sequence of the form (3.2) arises as the spectrum of L_v, for some $v \in E_0$. Let S be the set of all such sequences. Let $\mu: E_0 \to S$ be the map which associates to each element, v, of E_0, the spectrum of L_v. Then Theorem 2 of Chapter 6 of [PT] gives

Theorem 3.2. *μ is a real analytic isomorphism between E and S.*

Let us now fix v, and take $\{\psi^n\}$ to be the normalized eigenvectors of L_v. Since $\{\psi^n\}$ are complete, we can write the solution of (1.1) as

$$u(x,t) = \sum_{n=1}^{\infty} q_n(t)\psi^n(x).$$

Inserting this expansion into (1.1), and assuming that we can interchange summation and differentiation at will, we find that (1.1) is equivalent to the infinite system of coupled, ordinary differential equations

$$\ddot{q}_j(t) = \mu_j q_j(t) - \varepsilon \sum_{j_1, j_2, j_3 \geqq 1} q_{j_1}(t) q_{j_2}(t) q_{j_3}(t) (\psi^j, \psi^{j_1}\psi^{j_2}\psi^{j_3}). \tag{3.3}$$

Here $(\cdot,\cdot)$ is the inner product in $L^2[0,1]$.

These are the equations of motion for a hamiltonian system with hamiltonian

$$H(p,q) = \sum_{j=1}^{\infty} \tfrac{1}{2}(p_j^2 - \mu_j q_j^2) - \frac{\varepsilon}{4} \sum_{j_1 j_2 j_3 j_4} q_{j_1} q_{j_2} q_{j_3} q_{j_4} (\psi^{j_1}, \psi^{j_2}\psi^{j_3}\psi^{j_4}). \tag{3.4}$$

This is the hamiltonian of an infinite set of harmonic oscillators, coupled together through the non-linear terms. Note that unlike the models studied in [FSW or VB], the nonlinear terms in (3.4) do not have a short range property. This introduces significant new difficulties into the analysis of these models. These difficulties are partially off-set, however, by the fact that the unperturbed frequencies in this model—namely the μ_j's, do not form a dense set as they did in previously studied models.

If we look at the $\varepsilon = 0$ case of (3.4) we find that it possesses two types of solutions. If $\mu_j < 0, p_j(t)$ and $q_j(t)$ are periodic with frequency $\sqrt{|\mu_j|}$. If, on the other hand, $\mu_j > 0$, the corresponding mode grows or decays exponentially. From the asymptotic formula (3.2) we see that at most finitely many of the μ_j's are positive. If we consider perturbing an N-dimensional invariant torus corresponding to a quasi-periodic solution of N-modes with negative eigenvalues, the modes with positive eigenvalues will mean that some of the directions normal to the torus will be unstable. Experience with finite dimensional hamiltonian systems has shown that such "unstable," or "hyperbolic" invariant tori are easier to treat than "stable" tori [G, Z, E], which in our terminology are those with $\mu_j < 0$. However, their presence complicates the construction below, so we exclude potentials, v, from (1.1) which have $\mu_1 \geqq 0$. It is relatively easy to derive sufficient conditions to ensure that this is satisfied. In particular, we have:

Proposition 3.3. *If* $\|v\|_{L^1} < 1$, *then* L_v *has strictly negative spectrum.*

The proof of this proposition is straightforward, but not particularly relevant to what follows, so we omit it.

The hamiltonian (3.4) has infinitely many periodic and quasi-periodic solutions when $\varepsilon = 0$. We will use KAM methods to show that these solutions continue to exist for ε small and non-zero, provided the eigenvalues μ_n satisfy certain non-resonance conditions.

Suppose that we wish to study the quasi-periodic (or periodic) solution corresponding to the modes $j_1, \ldots, j_N$ of the hamiltonian (3.4). (In the periodic case we just have $N = 1$.) We can assume, without loss of generality, that $\{j_1, \ldots, j_N\} = \{1, \ldots, N\}$, since we can always relabel finitely many of the eigenfunctions $\{\psi^j\}$ to make this so, and the asymptotic estimates on the $\{\psi^j\}$'s and the eigenvalues $\{\mu_j\}$ that we need below will not be affected by shuffling finitely many of the modes.

The relevant quantities for the non-resonance conditions are square roots of the eigenvalues μ_j, so we define $\omega_j = \sqrt{|\mu_j|}$, $j = 1, 2, \ldots$.

Non-resonance Conditions. Let $\Omega = (\omega_1, \ldots, \omega_N)$. We then require

(D.1) $$|n \cdot \Omega \pm j\pi| \geqq D_0^{(1)}[|n| + j]^{-\tau}, \quad \text{for} \quad n \neq 0, \quad n \in \mathbb{Z}^N, \quad j \geqq 0,$$

(D.2) $$|n \cdot \Omega \pm \omega_j| \geqq D_0^{(1)}[|n| + j]^{-\tau}, \quad \text{for} \quad j \geqq N+1, \quad n \in \mathbb{Z}^N,$$

(D.3) $$|n \cdot \Omega \pm (\omega_j \pm \omega_l)| \geqq D_0^{(2)}[|n| + |j - l|]^{-4\tau}, \quad \text{for} \quad j, l \geqq N+1, \quad n \in \mathbb{Z}^N,$$

for some constants $D_0^{(1)}, D_0^{(2)}$, and τ. (Given any vector x, $|x| \equiv \sum_j |x_j|$.)

We can now state our principal technical result.

Theorem 3.4. *Suppose the frequencies in the hamiltonian (3.4) satisfy* (D.1)–(D.3). *Suppose further that certain non-linear relationships between* $\omega_1, \dots, \omega_N$ *and* $\psi^1, \dots, \psi^N$ *(which are analogous to the "anisochronicity" requirements of the usual KAM theorems) are non-zero. (In Sect. 4 we define these relationships explicitly and show that they are always satisfied if* $N = 1, 2$ *or* 3, *and for arbitrary finite N they are satisfied for almost every set of frequencies, with respect to the probability measure introduced in Theorem 2.1.) Then there exists* $\varepsilon_0 > 0$, *such that if* $|\varepsilon| < \varepsilon_0$, *the hamiltonian (3.4) has either a periodic, or quasi-periodic trajectory. (Depending on whether* $N = 1$, *or* $N > 1$.) *The frequency,* $\tilde{\Omega}$, *of the periodic orbit, or the independent frequencies,* $\tilde{\Omega}$, *in the quasi-periodic case, satisfy* $|\Omega - \tilde{\Omega}| \sim \mathcal{O}(\varepsilon)$. *What is more, the functions* $q_j(t)$ *satisfy* $\sup_t |q_j(t)| \leqq K_2/j^{(3/2)-\eta}$, *where* $K_2 > 0$, *and* η *is a small positive constant, which we will choose to be* $1/10$. *(*η *can be made arbitrarily small by shrinking* ε_0).

Remark. One may wonder if there are any potentials, v, whose spectrum satisfies (D.1)–(D.4). In Sect. 9 we will show that there is a natural probability measure on the space of potentials with negative spectrum such that with probability one, a given potential gives rise to a set of frequencies that satisfy the non-resonance conditions.

We conclude this section by showing how Theorem 3.4 implies Theorems 2.1 and 2.2. We have already remarked that the hypotheses of Theorem 3.4 apply to almost every potential with purely negative spectrum, v, so Theorems 2.1 and 2.2 will follow if we show how to construct periodic, or quasi-periodic solutions $u(x,t)$ for (1.1), from the solutions $q(t)$, found in Theorem 3.4.

Set $u^M(x,t) = \sum_{j=1}^{M} q_j(t)\psi^j(x)$. Since $|\psi^j(x)|$ is uniformly bounded in j and x (assume ψ^j is normalized so that $\|\psi^j\|_{L^2} = 1$), the estimates on q_j imply that $u^M(x,t)$ converges uniformly to a continuous function $u(x,t)$ which is either periodic, or quasi-periodic in t, depending on whether or not $q(t)$ is periodic or quasi-periodic.

In fact we have even better estimates on $u(x,t)$. Recall that $\psi^j(x)$ is C^1, and in fact, by Theorem 4 of Chap. 2 of [PT],

$$(\psi^j)'(x) = \sqrt{2}\pi j \cos(\pi j x) + \mathcal{O}(1).$$

With this estimate we see that

$$\begin{aligned} u(x+\varepsilon,t) - u(x,t) &= \sum_{j=1}^{\infty} q_j(t)[\psi^j(x+\varepsilon) - \psi^j(x)] \\ &= \sum_{j=1}^{M} q^j(t)[\psi^j(x+\varepsilon) - \psi^j(x)] + \sum_{j=M+1}^{\infty} q_j(t)[\psi^j(x+\varepsilon) - \psi^j(x)]. \end{aligned}$$

The second of these terms is bounded by $C/M^{1/2-\eta}$, using the estimate on $q_j(t)$ in Theorem 3.4. The first sum is bounded by noting that $|\psi^j(x+\varepsilon) - \psi^j(x)| \leqq C \cdot j \cdot \varepsilon$, which when combined with the bound on q_j implies the first sum is bounded by $C \cdot \varepsilon \cdot M^{1/2-\eta}$. If we take $M =$ integer part of $[\varepsilon^{-1}]$, we see that $u(x,t)$ is Hölder in x, with exponent $(\frac{1}{2} - \eta)$. A similar calculation shows that if one fixes $x, u(x,t)$ is Hölder in t.

With this information, we can estimate the amount by which $u^M(x,t)$ fails to satisfy (1.1).

$$\begin{aligned}
&|\partial_t^2 u^M(x,t) - (\partial_x^2 - v)u^M(x,t) + \varepsilon(u^M(x,t))^3| \\
&= \left| \sum_{j=1}^{M} \psi^j(x) \left[\ddot{q}_j - \mu_j q_j + \varepsilon \sum_{j_1,j_2,j_3 \leqq M} q_{j_1} q_{j_2} q_{j_3} (\psi^j, \psi^{j_1}\psi^{j_2}\psi^{j_3}) \right] \right. \\
&\quad \left. + \varepsilon \sum_{j=M+1}^{\infty} \sum_{j_1,j_2,j_3 \leqq M} \psi^j(x) q_{j_1} q_{j_2} q_{j_3} (\psi^j, \psi^{j_1}, \psi^{j_2}, \psi^{j_3}) \right| . \qquad (3.5)
\end{aligned}$$

In this step we have expanded $(u^M)^3$ in its Fourier series with respect to $\{\psi^j\}$. We used, and will reuse below.

Lemma 3.5. *The expansions of $(u^M(x,t))^3$ and $(u(x,t))^3$ with respect to $\{\psi^j\}$ converge uniformly.*

This lemma is a corollary of Proposition 4.2, so we delay its proof until the next section.

If we now use the fact that $\{q_j\}$ satisfies (3.3) we can rewrite (3.5) as

$$\begin{aligned}
&\varepsilon \left| ((u^M(x,t))^3 - (u(x,t))^3 + \sum_{j=M+1}^{\infty} \sum_{j_1,j_2,j_3 \geqq 1} \psi^j(x) q_{j_1} q_{j_2} q_{j_3} (\psi^j, \psi^{j_1}\psi^{j_2}\psi^{j_3}) \right| \\
&= \varepsilon \left| ((u^M(x,t))^3 - (u(x,t))^3) + \sum_{j=M+1}^{\infty} \psi^j(x)(\psi^j, u^3) \right| .
\end{aligned}$$

By Lemma 3.5, this last sum converges uniformly to zero as $M \to \infty$, so we have

Proposition 3.6. $|\partial_t^2 u^M(x,t) - (\partial_x^2 - v)u^M(x,t) + \varepsilon(u^M(x,t))^3|$ *goes uniformly to zero as $M \to \infty$. Thus, $u(x,t)$ is a distributional solution of (1.1).*

Proposition (3.6) immediately implies Theorem 2.1 or Theorem 2.2 depending on whether we are in the periodic, or quasi-periodic case.

IV. The Iterative Scheme

In the present section we introduce new canonical variables for the hamiltonian (3.4). We also prove bounds on the interaction terms in the hamiltonian. For this task, we must introduce the domains on which we work, and norms for functions analytic on these domains.

We begin with some terminology. Let γ be a small positive constant. Define integers $N_k = 0$, for $k \leqq 0$, and N_k an increasing sequence defined in Sect. 5 for $k \geqq 1$, which increases like $c^{(1+\gamma)^k}$. We now partition the positive integers into subsets $\mathbb{I}_k$ defined by

$$\mathbb{I}_k = \{N_k + 1, N_k + 2, \ldots, N_{k+1}\}, \quad k \geqq 0.$$

Now take $\mathbb{L}^{(k)}$ a (possibly empty) set of positive integers.

Definition 4.1. The set $\mathbb{L}^{(k)}$ is *k-admissible* if

$$|\mathbb{L}^{(k)} \cap \mathbb{I}_j| \leqq \mathscr{L}(k,j), \quad \text{for} \quad j = 0, 1, 2, \ldots,$$

where

$$\mathscr{L}(k,j)=\begin{cases}2^3[M(j+1)-k] & \text{if } j>k-M\\ 0 & \text{if } j\leqq k-M.\end{cases}$$

(In this definition, M is a large positive integer whose value, like that of the constant γ will be fixed in the course of the proof, and $|\cdots|$ denotes the cardinality of the set enclosed.) A k-admissible set of integers is a "sparse" set of integers whose "sparseness" increases as k increases, and also as one gets farther from the origin.

Now let $\underline{t}$ and $\underline{T}$, be elements of $\mathbb{R}^{\mathbb{Z}^+}$, whose components are all positive. Let $\mathbb{L}^{(0)}$ be a zero admissible set. We define the domain

$$V(\underline{t};\mathbb{L}^{(0)},\underline{T})=\{q\in\mathbb{C}^{\mathbb{Z}^+}\mid |q_j|<T_j \text{ if } j\in\mathbb{L}^{(0)}; |q_j|<t_j \text{ otherwise}\}.$$

The motivation for choosing these rather peculiar domains is discussed later in this section. For the moment, however, we wish to concentrate on showing why, for a particular choice of $\underline{t}$ and $\underline{T}$, the interaction terms of (3.4) are bounded on V.

Proposition 4.2. *Let $\mathbb{L}^{(0)}$ be a 0-admissible set, $t_j=c(t)/j^{9/10}$ and $T_j=C(T)/j^{1/12}$. (Here $c(t)$ and $C(T)$ are arbitrary large constants.) Then there is a constant $K>0$ (independent of $\mathbb{L}^{(0)}$) such that*

$$\sum_{j_1,j_2,j_3,j_4\geqq 1}\ \sup_{V(t,\mathbb{L}^{(0)},T)} |q_{j_1}||q_{j_2}||q_{j_3}||q_{j_4}||(\psi^{j_1},\psi^{j_2}\psi^{j_3}\psi^{j_4})|<K. \tag{4.1}$$

The convergence of the sum in (4.1) depends on estimating $(\psi^{j_1},\psi^{j_2}\psi^{j_3}\psi^{j_4})$. To estimate this inner product we use

Lemma 4.3. *Let $\varphi^n(x)=\sqrt{2}\sin(n\pi x)$, (the normalized eigenfunctions of the $v=0$ problem). Then there is a constant c, depending only on v, such that*

$$|(\psi^j,\varphi^n)|\leqq\frac{c}{1+|j^2-n^2|},\quad \text{for all}\quad j,n=1,2,3,\ldots.$$

Proof. $\mu_j(\psi^j,\varphi^n)=(L_v\psi^j,\varphi^n)=(\psi^j,L_v\varphi^n)=-n^2\pi^2(\psi^j,\varphi^n)+(\psi^j,v\varphi^n)$. Using the asymptotic estimate $-\mu_j\sim(j\pi)^2$, the desired estimate follows immediately.

Since the $\{\varphi^n\}$ are complete we can write

$$\psi^j(x)=\sum_{n=1}^{\infty}a^j{}_n\varphi^n(x),$$

where

$$|a^j{}_n|\leqq\frac{c}{1+|j^2-n^2|},$$

and the Fourier series converges uniformly since $\psi^j\in C^1$.

Thus, to estimate the sum in Proposition 4.2 we need to estimate

$$\sum_{\substack{j_1,\ldots,j_4\\ n_1,\ldots,n_4}}|q_{j_1}|\cdots|q_{j_4}|[(1+|j_1^2-n_1^2|)\cdots(1+|j_4^2-n_4^2|)]^{-1}|(\varphi^{n_1},\varphi^{n_2}\varphi^{n_3}\varphi^{n_4})|. \tag{4.2}$$

The issue is further complicated by the fact that the size of $|q_j|$ depends on whether or not $j \in \mathbb{L}^{(0)}$. However, (4.2) may be rewritten as

$$\sum_{l=0}^{4} \binom{4}{l} \sum_{j_1 \cdots j_l \in L^{(0)}} |q_{j_1}| \cdots |q_{j_l}| \left(\sum_{j_{l+1} \cdots j_4 \notin L^{(0)}} |q_{j_{l+1}}| \cdots |q_{j_4}| (\psi^{j_1}, \ldots, \psi^{j_4}) \right)$$

$$\leqq c \times \sum_{l=0}^{4} \binom{4}{l} \sum_{j_1 \cdots j_l \in L^{(0)}} T_{j_1} \cdots T_{j_l}$$

$$\cdot \left(\sum_{\substack{j_{l+1} \cdots j_4 \notin L^{(0)} \\ n_1 \cdots n_4}} t_{j_{l+1}} \cdots t_{j_4} \frac{(\phi^{n_1}, \ldots, \phi^{n_4})}{(1 + |j_1^2 - n_1^2|) \cdots (1 + |j_4^2 - n_4^2|)} \right). \tag{4.3}$$

We bound the sums over $j \notin L^{(0)}$ by inserting the definitions of t_j and then estimating the resulting sums with the aid of the integrals

$$\int_1^{a-1} \frac{dx}{x^{9/10}|x^2 - a^2|} \leqq \frac{c}{(a-1)^{9/10}} \quad (a > 1)$$

and

$$\int_{a+1}^{\infty} \frac{dx}{x^{9/10}|x^2 - a^2|} \leqq \frac{c}{(a-1)^{9/10}}.$$

Thus the term in parentheses is bounded by

$$c \times \sum_{n_1, \ldots, n_4} \frac{(\phi^{n_1}, \ldots, \phi^{n_4})}{(1 + |j_1^2 - n_1^2|) \cdots (1 + |j_l^2 - n_l^2|) n_{l+1}^{(9/10)} \cdots n_4^{(9/10)}}.$$

(If $l = 0$, then no factors of $(1 + |j_i^2 - n_i^2|)$ appear, while if $l = 4$, no factors of $n_i^{9/10}$ appear.) We will show that this sum is bounded by a constant by first showing that

$$\sum_{n_{l+1} \cdots n_4} \frac{(\phi^{n_1} \cdots \phi^{n_4})}{n_{l+1}^{9/10} \cdots n_4^{9/10}}$$

is uniformly bounded, since one has immediately that

$$\sum_{n_1 \cdots n_l} \frac{1}{(1 + |j_1^2 - n_1^2|) \cdots (1 + |j_l^2 - n_l^2|)} < \text{constant}.$$

Because of the orthogonality of the trigonometric functions, $(\varphi^{n_1}, \ldots, \varphi^{n_4})$ vanishes unless $n_4 \pm n_3 \pm n_2 \pm n_1 = 0$ for some combination of plus and minus signs. We will estimate the sum in the case where $n_4 - (n_3 + n_2 + n_1) = 0$. The other five choices of signs are handled in exactly the same fashion. Since $n_4 = n_3 + n_2 + n_1$ (and all n_j's are positive),

$$\frac{1}{n_4^{9/10}} \leqq \frac{c}{(n_1 \cdot n_2 \cdot n_3)^{3/10}} \leqq \frac{c}{(n_{l+1}, \ldots, n_3)^{3/10}},$$

so the sum over $n_{l+1} \cdots n_4$ is bounded by

$$c \cdot \sum_{n_{l+1}, \ldots, n_3} \frac{1}{n_{l+1}^{12/10} \cdots n_3^{12/10}} \leqq \text{const.}$$

(as usual, the sum is set equal to one if $l + 1 > 3$).

Returning now to (4.3) we see that it is bounded by

$$c \times \sum_{l=0}^{4} \binom{4}{l} \sum_{j_1 \cdots j_l \in L^{(0)}} T_{j_1} \cdots T_{j_l}.$$

We bound the sums over $j \in L^{(0)}$ by,

$$\sum_{j \in L^{(0)}} T_j \leqq \sum_{j \in L^{(0)}} j^{-1/12} = \sum_{n=0}^{\infty} \sum_{j \in L^{(0)} \cap I_n} j^{-1/12} \leqq \sum_{n=0}^{\infty} \frac{2^3 M(n+1)}{(N_n+1)^{1/12}}$$
$$\leqq \sum_{n=0}^{\infty} 2^3 M(n+1)(1+\varepsilon_0^{-(1+\gamma)^n})^{-1/12} < \infty.$$

Thus (4.3) is bounded independent of $L^{(0)}$ and the proof of Proposition 4.2 is complete.

Note that as a rather straightforward extension of the proof of this proposition we have

Corollary 4.4. (Lemma 3.5.) *Let $u^M(x,t)$ and $u(x,t)$ be the approximate, and distributional solutions of (1.1), constructed in the previous section. Then the expansions of $(u^M)^3$ and u^3, with respect to $\{\psi^j\}$, converge uniformly.*

Proof. We will write out the details for $u^3(x,t)$—$(u^M(x,t))^3$ follows in like fashion. The expansion of $u^3(x,t)$ is

$$\sum_{j=1}^{\infty} \psi^j(x)(\psi, u^3) = \sum_{j,j_1,j_2,j_3} \psi^j(x) q_{j_1} q_{j_2} q_{j_3} (\psi^j, \psi^{j_1} \psi^{j_2} \psi^{j_3}).$$

We will show that this sum converges uniformly. Since we know it converges to u^3 in L^2, and since $u^3(x,t)$ is cont, this implies the sum converges uniformly to u^3.

Bounding $|q_j| < K/j^{3/2-\eta}$ with the help of Theorem 3.4, and inserting the expansion for the ψ^j's that comes from Lemma 4.3 we have

$$\sup_{x \in [0,1]} \left| \sum_{j=1}^{\infty} \psi^j(x)(\psi, u^3) \right|$$
$$\leqq c \leqq \sum_{\substack{j,j_1,j_2,j_3 \\ n,n_1,n_2,n_3}}$$
$$\cdot \frac{|(\varphi^n, \varphi^{n_1}\varphi^{n_2}\varphi^{n_3})|}{(j_1^{(3/2-\eta)} j_2^{(3/2-\eta)} j_3^{(3/2-\eta)})(1+|n^2-j^2|)(1+|n_1^2-j_1^2|)(1+|n_2^2-j_2^2|)(1+|n_3^2-j_3^2|)}$$
$$\leqq c' \sum_{n,n_1,n_2,n_3} \frac{|(\varphi^n, \varphi^{n_1}\varphi^{n_2}\varphi^{n_3})|}{n_1^{5/4} n_2^{5/4} n_3^{5/4}} < \infty,$$

so the expansion converges uniformly. Note that the last inequality used the orthogonality of the trigonometric functions to eliminate the sum over n.

It is clear from the foregoing discussion that if one had chosen in Eq. (1.1) not a cubic non-linearity but rather some more complicated polynomial, one would get a hamiltonian consisting of infinitely many harmonic oscillators coupled by a more complicated polynomial interaction. Nonetheless, one can bound the interaction terms on domains like those above (although one may have to choose the vectors

$\underline{t}$ and $\underline{T}$ to decay at slightly different rates) and the remainder of the proof goes through very much as in the present case. It is not clear, however, how one would deal with a general analytic interaction term in (1.1).

Having now gotten some control over the interaction terms in the hamiltonian we make a canonical change of variables which simplifies the perturbative argument used to prove Theorem 3.4. Recall that we wish to construct a quasi-periodic (or periodic) orbit for the hamiltonian corresponding to the unperturbed motion of $(p_1, q_1, \ldots, p_N, q_N)$. If we look at the equations of motion for $\varepsilon = 0$ we see that we get a quasi-periodic solution by taking $p_j(t) = q_j(t) = 0$ if $j \geqq N+1$. Thus we treat these two groups of variables somewhat differently. Define

$$q_j = \sqrt{I_j/\omega_j} \cos \varphi_j, \quad p_j = -\sqrt{I_j \omega_j} \sin \varphi_j, \quad j = 1, \ldots, N. \tag{4.4}$$

and

$$z_j = \sqrt{\frac{1}{2\omega_j}} (p_j - i\omega_j q_j), \quad \bar{z}_j = \sqrt{\frac{1}{2\omega_j}} (p_j + i\omega_j q_j), \quad j \geqq N+1.$$

This differentiation between the coordinates we wish to perturb, and the remaining degrees of freedom for the system, has been used in [G, Z, and E] to construct low dimensional invariant tori for nearly integrable hamiltonian systems with finitely many degrees of freedom. Note that this is not quite a canonical transformation since $dz_j \wedge d\bar{z}_j = i dp_j \wedge dq_j$, but as noted in [E], this factor of $\sqrt{-1}$ causes no problems.

In terms of these new variables, the hamiltonian (3.4) takes the form

$$H(I, \varphi; z, \bar{z}) = \sum_{j=1}^{N} \omega_j I_j + \sum_{j=N+1}^{\infty} \omega_j z_j \bar{z}_j + \frac{\varepsilon}{4} g(I, \varphi; z, \bar{z}).$$

We will look more closely at the exact form of g in a moment, but let us first note that we are interested in perturbing around the torus $\underline{I} = \underline{I}^0$, $z = \bar{z} = 0$. Defining $\underline{J} = \underline{I} - \underline{I}^0$, the hamiltonian becomes

$$H(J, \varphi; z, \bar{z}) = \text{const.} + \sum_{j=1}^{N} \omega_j J_j + \sum_{j=N+1}^{\infty} \omega_j z_j \bar{z}_j + \frac{\varepsilon}{4} f(I^0, J, \varphi; z, \bar{z}), \tag{4.5}$$

with $f(I^0, J, \varphi; z, \bar{z}) = g(I^0 + J, \varphi; z, \bar{z})$. This is a convenient set of variables with which to work since we are then perturbing around the point $J = z = \bar{z} = 0$.

The equations of motion for the hamiltonian (4.5) are

$$\dot{J}_j = -\frac{\varepsilon}{4} \frac{\partial f}{\partial \varphi_j}; \quad \dot{\varphi}_j = \omega_j + \frac{\varepsilon}{4} \frac{\partial f}{\partial J_j}, \quad j = 1, \ldots, N,$$

$$\dot{z}_j = -i\omega_j z_j - i \frac{\varepsilon}{4} \frac{\partial f}{\partial \bar{z}_j}; \quad \dot{\bar{z}}_j = i\omega_j \bar{z}_j + i \frac{\varepsilon}{4} \frac{\partial f}{\partial z_j}, \quad j \geqq N+1. \tag{4.6}$$

We will show below that f admits a power series expansion in J, z, and $\bar{z}$. If that expansion contained only terms of quadratic or higher powers of these variables, $J = z = \bar{z} = 0$ would still be an invariant torus for (4.6). It is this observation which motivates our approach to the problem. Instead of attempting to construct a canonical transformation which "kills" the term, f, in (4.5) completely, we will show that there is such a transformation which eliminates the low order terms in the power series for f. This is sufficient to conclude the existence of an invariant torus and it has the important consequence that the frequencies $\{\omega_j\}$ need to satisfy far fewer non-resonance conditions than would be needed if we tried to eliminate f entirely. It is not at all clear that the frequencies $\{\omega_j\}$ arising in this problem could be forced to satisfy these additional conditions. We remark that Nikolenko [N] does transform to linear equations of motion when he applies similar ideas to evolution equations and thus he is required to impose these more stringent non-resonance conditions on the linear part of his systems.

We now look somewhat closer at the function $f(I^0, J, \varphi; z, \bar{z})$, in (4.5). If we insert the change of variables into the interaction term in (3.4) we find

$$f(I^0, J, \varphi; z, \bar{z}) = \sum{}' \omega^{-(\alpha+\beta+\gamma)/2}(I^0 + J)^{\gamma/2}(\cos\varphi)^\gamma \frac{(z)^\alpha(-\bar{z})^\beta}{(2i)^{|\alpha+\beta|/2}}(\psi^{\alpha+\beta+\gamma}). \tag{4.7}$$

The multi-index notation is standard, except for $(\cos\varphi)^\gamma \equiv \prod_{j=1}^{N}(\cos\varphi_j)^{\gamma_j}$, and $(\psi^{\alpha+\beta+\gamma}) \equiv (\psi^{j_1}, \psi^{j_2}\psi^{j_3}\psi^{j_4})$, where we use the fact that the restrictions on $\sum'$ insure that $\psi^{\alpha+\beta+\gamma} = \prod_j (\psi^j)^{(\alpha+\beta+\gamma)_j}$ is a product of four ψ^j's. The restrictions on $\sum'$ are $4 \geqq \alpha_i, \beta_i, \gamma_i \geqq 0$ for all $i, \alpha_i, \beta_i = 0$ if $i \leqq N$, $\gamma_i = 0$ if $i \geqq N+1$, and $|\alpha+\beta+\gamma| \equiv \sum_i |(\alpha+\beta+\gamma)_i| = 4$.

If we fix some value of I^0 then Proposition 4.2 implies that for I near I^0, J near 0, and z_j and $\bar{z}_j$ going to zero quickly enough with j, the sum in (4.7) converges. We now define some domains, related to the domains, V, defined earlier, that allow us to discuss this convergence.

Definition 4.5. Let $\mathbb{L}^{(k)}$ be a k-admissible set of integers with $k \geqq 0$. Let $\underline{\tau}$ and $\underline{\mathscr{T}}$ be infinite dimensional vectors with positive components, let v, ρ, and σ be positive real numbers, and let $I^0 \in (\mathbb{R}^+)^N$. We define the domain

$$\begin{aligned} D(I^0, v; \rho, \sigma, \underline{\tau}; \mathbb{L}^{(k)}, \mathscr{I}) = \{&(I, J, \varphi; z, \bar{z}) \in \mathbb{C}^N \times \mathbb{C}^N \times \mathbb{C}^N \times \mathbb{C}^{\mathbb{Z}^+} \times \mathbb{C}^{\mathbb{Z}^+} \mid \\ &|I_j^0 - I_j| < v, |J_j| < \rho, \text{ and } |\operatorname{Im}\varphi_j| < \sigma \text{ for } j = 1, \ldots, N; \\ &|z_i| < \tau_i, |\bar{z}_i| < \tau_i \text{ if } i \notin \mathbb{L}^{(k)}; |z_i| < \mathscr{T}_i. |\bar{z}_i| < \mathscr{T}_i \text{ if } i \in \mathbb{L}^{(k)}\}. \end{aligned}$$

Let $f(I, J, \varphi; z, \bar{z})$ be a function analytic on the domain $D(I^0, v, \rho, \sigma, \underline{\tau}; \mathbb{L}^{(k)}, \underline{\mathscr{T}})$, which is 2π-periodic in φ. Then we can expand f in a Laurent–Fourier series

$$f(I, J, \varphi; z, \bar{z}) = \sum{}' \hat{f}(I, k, n; \alpha, \beta) J^k e^{in\cdot\varphi} z^\alpha \bar{z}^\beta, \tag{4.8}$$

where $\sum'$ means we restrict the sum to $k \in \mathbb{Z}^N$, $k_i \geqq 0$, $n \in \mathbb{Z}^N$, $\alpha, \beta \in (\mathbb{Z})^{\mathbb{Z}^+}$, $\alpha_i, \beta_i \geqq 0$.

Define

$$(\alpha_S)_i = \begin{cases} \alpha_i & \text{if } i \notin \mathbb{L}^{(k)} \\ 0 & \text{otherwise} \end{cases} \qquad (\alpha_B)_i = \begin{cases} \alpha_i & \text{if } i \in \mathbb{L}^{(k)} \\ 0 & \text{otherwise} \end{cases}$$

Clearly, $\alpha = \alpha_S + \alpha_B$. Define β_S and β_B analogously. We then define the norm of the function f, on the domain D to be

$$\| f \|_D = \sum{}' \left(\sup_{|I_j - I_j^0| < v} |\hat{f}(I, k, n; \alpha, \beta)| \right) \rho^{|k|} e^{\sigma |n|} \tau^{\alpha_S + \beta_S} \mathscr{T}^{\alpha_B + \beta_B}.$$

These norms are modeled after those used by Vittot and Bellisard ([V, VB]).

Note that one property of these norms that we will use repeatedly is that if f and g are analytic on D, then

$$\| fg \|_D \leqq \| f \|_D \| g \|_D.$$

The proof of this simple fact follows [V].

Note that as a corollary of Proposition 4.2, we can find positive constants v^0, ρ^0 and σ^0, such that if we take $\tau_j^0 = j^{-(2/5)}$, and $\mathscr{T}_j^0 = j^{5/12}$, we have

Corollary 4.6. *The function $f(I, J, \varphi; z, \bar{z})$ is analytic on $D(I^0, v^0, \rho^0, \sigma^0, \underline{\tau}^0; \mathbb{L}^{(0)}, \underline{\mathscr{T}}^0)$, for every 0-admissible set $\mathbb{L}^{(0)}$. Furthermore there is a constant $K > 0$ (independent of $\mathbb{L}^{(0)}$) such that*

$$\| f \|_D \leqq K.$$

The corollary follows, because if we take any point $(I, J, \varphi; z, \bar{z})$ in the domain D, and untangle the various canonical change of variables, we find that the corresponding points $q_{j_1}, \ldots, q_{j_4}$ lie inside the domain $V(t, \mathbb{L}^{(0)}, I)$, of Proposition 4.2 if we choose $c(t)$ and $C(T)$ large enough. We can allow the variables z_j and $\bar{z}_j$ to go to zero more slowly with j than the variables q_j did, because of the factors of $1/\sqrt{\omega_j}$ that appear in (4.4), since asymptotically, $(1/\sqrt{\omega_j}) \sim (1/\sqrt{\pi j})$.

The reason for choosing these rather unusual domains and norms is the following. If we write out explicitly the Taylor–Laurent series for the interaction term $f(I^0, J, \varphi; z, \bar{z})$ in (4.7) we find that each factor of z_j or $\bar{z}_j$ has a coefficient proportional to $1/\sqrt{\omega_j}$, again because of the transformation (4.4). Thus, $\partial f / \partial z_j$ or $\partial f / \partial \bar{z}_j$ can be bounded by $c \| f \| / \sqrt{j}$, and these factors of $1/\sqrt{j}$ are very useful in making the sums one encounters in the iterative process converge. The domains, D, and their attendent norms are constructed in an attempt to preserve as much of this decay as possible. To see how this occurs suppose that f is a function, analytic, and uniformly bounded on all domains $D(I^0, v, \rho, \sigma, \tau; \mathbb{L}^{(k)}, \underline{\mathscr{T}})$, for $\mathbb{L}^{(k)}$ any k-admissible set. If we shrink $\underline{\mathscr{T}}$, say to $\hat{\underline{\mathscr{T}}} = \frac{1}{2} \underline{\mathscr{T}}$ then $\partial f / \partial z_j$ is analytic on any domain $\hat{D}(I^0, v, \rho, \sigma, \tau; \hat{\mathbb{L}}^{(k)}, \hat{\underline{\mathscr{T}}})$ such that $\mathbb{L}^{(k)} = \{j\} \cup \hat{\mathbb{L}}^{(k)}$, and Cauchy's Theorem implies that

$$\left\| \frac{\partial f}{\partial z_j} \right\|_{\hat{D}} \leqq \frac{2 \| f \|_D}{\mathscr{T}_j} \sim \frac{\| f \|_D}{j^{5/12}}.$$

Note that we may choose $\hat{\mathbb{L}}^{(k)}$ to be any set of integers satisfying $\| \hat{L}^{(k)} \cap \mathbb{I}_j \| \leqq \max(0, \mathscr{L}(k, j) - 1)$. Thus, as the iterative process proceeds, we preserve most of the decay of derivatives with respect to j, but the number of points in $\mathbb{L}^{(k)}$ decreases.

Let us now look again at the hamiltonian (4.5). At the moment, the frequencies ω_j, corresponding to motion on the invariant torus are constants. As is typically the case in KAM type arguments, we will need to adjust these frequencies slightly as the iteration proceeds. We extract a term from the interaction, f, which does not depend on the angle variables, that allows us to do that.

We begin by defining the *order* of terms in a power series like (4.8). The order of a term $\hat{f}(I,k,n;\alpha,\beta)J^k e^{in\cdot\varphi} z^\alpha \bar{z}^\beta$ is

$$2|k|+|\alpha|+|\beta| \equiv 2\left(\left(\sum_{j=1}^{N}|k_j|\right)+\sum_{j=N+1}^{\infty}(|\alpha_j|+|\beta_j|)\right).$$

As usual in KAM arguments, the "unperturbed" hamiltonian (in (4.5) the $\varepsilon=0$ case) will be altered by the accumulation of terms that cannot be eliminated by the successive changes of variables. In the present case these terms are what allow us to adjust the frequencies of the (quasi-)periodic motion, so we compute them explicitly to lowest order in ε. As can be seen from the inductive argument in Sects. 6–8, (see in particular (8.4)), the terms which are not eliminated from f:

(i) are quadratic with respect to the ordering defined above,
(ii) are independent of φ,
(iii) involve z_j and $\bar{z}_j$ only as the product $(z_j\bar{z}_j)$.

Thus, using (4.7), we can write an explicit representation of them to lowest order; namely

$$\varepsilon\sum_{j=1}^{N}\left(\frac{3}{4}\right)\frac{I_jJ_j}{\omega_j^2}(\psi^j\psi^j,\psi^j\psi^j)+\varepsilon\sum_{\substack{j,l=1\\ j\neq l}}^{N}\left(\frac{1}{4}\right)\frac{I_jJ_l}{\omega_j\omega_l}(\psi^j\psi^j,\psi^l\psi^l)$$
$$\varepsilon\sum_{j=1}^{N}\sum_{l=N+1}^{\infty}\left(\frac{1}{4}\right)\frac{I_jz_l\bar{z}_l}{\omega_j\omega_l}(\psi^j\psi^j,\psi^l\psi^l). \tag{4.9}$$

We first note that this will change the frequencies of the quasi-periodic motion:

$$\omega_j\to\tilde{\Omega}_j(I)=\omega_j+\varepsilon\left(\frac{3}{4}\right)\left(\frac{I_j}{\omega_j^2}\right)(\psi^j\psi^j,\psi^j\psi^j)+\varepsilon\sum_{\substack{l=1\\ j\neq l}}^{N}\left(\frac{1}{4}\right)\left(\frac{I_l}{\omega_j\omega_l}\right)(\psi^j\psi^j\psi^l\psi^l). \tag{4.10}$$

Define the right-hand side of this equality to be $\tilde{f}_j^{(1)}(I)$.

A very important quantity in KAM arguments is the matrix

$$\left(\frac{\partial\tilde{\Omega}}{\partial I}\right)_{ij}=\delta_{ij}\left(\frac{3\varepsilon}{4\omega_j^2}\right)(\psi^j\psi^j,\psi^j\psi^j)+(1-\delta_{ij})\left(\frac{\varepsilon}{4\omega_i\omega_j}\right)(\psi^i\psi^i,\psi^j\psi^j). \tag{4.11}$$

As we shall see in Sect. 9, it is necessary for this matrix to be invertible in order to control the small denominators which arise in the iteration. In the case $N=1$ (periodic motion) this matrix is always invertible. For $N=2$, we must invert

$$\varepsilon\begin{pmatrix}\frac{3}{4\omega_1^2}(\psi^1\psi^1,\psi^1\psi^1) & \frac{1}{4\omega_1\omega_2}(\psi^1\psi^1,\psi^2\psi^2)\\ \frac{1}{4\omega_1\omega_2}(\psi^1\psi^1,\psi^2\psi^2) & \frac{3}{4\omega_2^2}(\psi^2\psi^2,\psi^2\psi^2)\end{pmatrix}.$$

The determinant of this matrix is

$$\frac{\varepsilon^2}{16\omega_1^2\omega_2^2}(9(\psi^1\psi^1,\psi^1\psi^1)(\psi^2\psi^2,\psi^2\psi^2)-(\psi^1\psi^1,\psi^2\psi^2)^2).$$

This quantity never vanishes, however, since

$$|(\psi^1\psi^1,\psi^2\psi^2)|^2 \leqq |(\psi^1\psi^1,\psi^1\psi^1)(\psi^2\psi^2,\psi^2\psi^2)|.$$

Similarly, for $N=3$, the matrix $\partial\tilde{\Omega}/\partial I$ is always invertible. I have not yet found an argument guaranteeing $(\partial\tilde{\Omega}/\partial I)$ is always invertible for arbitrary N, but one can make the following observations. If the potential $v=0$, the terms in $\partial\tilde{\Omega}/\partial I$ can all be explicitly computed and we have

$$\det\left(\frac{\partial\tilde{\Omega}}{\partial I}\right)=\frac{\varepsilon^N}{4^N\left(\prod_{j=1}^N\omega_j\right)^2}\det\begin{pmatrix}3(\psi^1\psi^1,\psi^1\psi^1), & (\psi^1\psi^1,\psi^2\psi^2),\ldots, & (\psi^1\psi^1,\psi^N\psi^N)\\ (\psi^2\psi^2,\psi^1\psi^1), & 3(\psi^2\psi^2,\psi^2\psi^2),\ldots, & (\psi^2\psi^2,\psi^N\psi^N)\\ \vdots & & \\ (\psi^N\psi^N,\psi^1\psi^1), & \ldots\ldots & 3(\psi^N\psi^N,\psi^N\psi^N)\end{pmatrix}$$

$$=\frac{\varepsilon^N}{4^N\pi^{2N}(N!)^2}\det\begin{pmatrix}9,1,1,\ldots,1\\1,9,1,\ldots,1\\ \vdots\\ 1\ldots\ldots\ldots1\end{pmatrix}\neq 0.$$

We now make two observations:

(1) Since ψ^j and ω^j are analytic functions of the potential, v, (Chap. 2. [PT]) $\det(\partial\tilde{\Omega}/\partial I)$ will be non-zero for all v near zero.
(2) Since the (ψ^j's) are analytic functions of the frequencies, $\{\omega_j\}$, this determinant will vanish only for a set of frequencies of zero measure with respect to the probability measure defined in Sect. 9. This in turn implies that it vanishes for a set of potentials of zero measure. I am indebted to J. Poschel for this remark.

Combining these remarks we have

Lemma 4.7. *The matrix $(\partial\tilde{\Omega}/\partial I)$ is always invertible if $N=1,2$, or 3. For arbitrary positive, N, it is invertible for all except a set of potentials of measure zero. Furthermore we have $\det(\partial\tilde{\Omega}/\partial I)=c\varepsilon^N$.*

Note that from (4.9) we can also compute the change in the frequencies ω_j, $j\geqq N+1$. It is

$$\omega_j\to\tilde{\omega}_j(I)=\omega_j+\frac{\varepsilon}{4}\sum_{l=1}^N\frac{I_l}{\omega_l\omega_j}(\psi^j\psi^j,\psi^l\psi^l). \tag{4.12}$$

Define the right-hand side of this quantity to be $\tilde{g}_j^{(1)}(I)$. (We will need this formula in Sects. 5 and 9.)

V. The Inductive Step

In the present section we state an iterative proposition that allows us to prove Theorem 3.4. We begin by defining a sequence of inductive constants.

1. Define $\varepsilon_0 = K\varepsilon^{3/2}$, $\varepsilon_1 = \varepsilon^{5/2}$, $\varepsilon_2 = \varepsilon^{9/2}$, $\varepsilon_3 = \varepsilon^{17/2}$ and $\varepsilon_{k+1} = \varepsilon_k^{(1+\gamma)}$ for $k \geqq 3$, with γ a positive constant, smaller than 1, defined implicitly in the course of the proof. (We could take it to be (1/8), for instance.) These constants govern the size of the interaction term in the hamiltonian after k-steps in the iteration. The constant K is that which appears in Corollary 4.6.
2. A sequence of length scales (which were used in Sect. 4),

$$N_k = 0;\ k \leqq 0 \quad \text{and} \quad N_k = [\varepsilon_k^{-1}];\ k \geqq 1,$$

(Here, [...] refers to the integer part of the number enclosed.)
3. The size of the analyticity domain in the action variables, J, is determined by

$$\rho^{(0)} = \varepsilon^{1/4}/2^4, \quad \rho^{(3)} = \varepsilon^{19/32}, \quad \text{and} \quad \rho^{(k+1)} = \rho^{(k)}/2^5 \quad \text{for} \quad k = 0, 1 \quad \text{and} \quad k \geqq 3.$$

4. The size of the analyticity domain in the variables, I, used to adjust the frequencies of the quasi-periodic motion is given by $\nu^{(k)} = 4\rho^{(k)}$; $k \geqq 0$. We choose $I^0 = (1, \ldots, 1)\varepsilon^{1/4}$, which determines the torus around which we perturb.
5. Similarly we set $\sigma^{(0)} = 1, \delta^{(k)} = \sigma^{(0)}/2^6(k+1)^2$, $k \geqq 0$, and define $\sigma^{(k)} = \sigma^{(k)} - 8\delta^{(k)}$. The constant $\sigma^{(k)}$ will determine the size of the domain in the angle variables.
6. At each iterative step we will consider only finitely many Fourier coefficients, that number being determined by

$$M_k \equiv \frac{2}{\delta^{(k)}} |\log \varepsilon_k|.$$

The next two definitions determine the size of the domains in the variables z_j and $\bar{z}_j$.

7. Define $\underline{\tau}^{(k)}$ by

$$\tau_j^{(k)} = \tau_j^{(k-1)}/(1 + 2^{-5}k^{-2})^5; \quad k \geqq 1, j \geqq 1,$$
$$\tau_j^{(0)} = \varepsilon^{1/8} j^{-(2/5)}.$$

8. Define $\zeta_0 = 5/12$, and $\zeta_k = (1 + 2^{-11}k^{-2})^{-2}\zeta_{k-1}$, $k \geqq 1$. Then define $\underline{\mathscr{T}}^{(k)}$ by

$$\mathscr{T}_j^{(k)} = \varepsilon^{1/8} \cdot j^{\zeta_k} \Big/ \prod_{l=0}^{k} (1 + (l+1)^{-2})^4.$$

(In particular $\mathscr{T}_j^{(0)} = \varepsilon^{1/8} j^{\zeta_0}$.)

Remark. The powers with which $\tau_j^{(k)}$ and $\mathscr{T}_j^{(k)}$ decay and grow in definitions 6 and 7 are not uniquely determined. One could set $\tau_j^{(0)} \sim j^{-\alpha}$, $\mathscr{T}_j^{(0)} \sim j^{\beta}$ and attempt to optimize the estimates with respect to α and β, but this would further burden the notation, and so we have chosen one set of constants that works. We remark that the choice of α and β will change if we choose a different non-linearity in (1.1)—say u^5 instead of u^3.

9. The constants $[D_k^{(1)}]$ and $[D_k^{(2)}]$ determine the size of the "small denominators" at each step of the iteration. $D_0^{(1)}$ and $D_0^{(2)}$ are large constants defined in Sect. 3. For $k \geqq 1$ set:

$$D_k^{(1)} = D_0^{(1)} k^2 |\log \varepsilon|^2 \quad \text{for} \quad k = 1, 2; \quad D_k^{(1)} = D_0^{(1)} (\nu^{(0)}\varepsilon)^{-1} k^2 |\log \varepsilon|^2 \quad \text{for} \quad k \geqq 3$$

and

$$D_k^{(2)} = D_0^{(2)} k^4 |\log \varepsilon|^4 \quad \text{for} \quad k = 1, 2; \quad D_k^{(2)} = D_0^{(2)} (\nu^{(0)}\varepsilon)^{-1} k^4 |\log \varepsilon|^4, \quad k \geqq 3.$$

We can now state our inductive proposition.

Proposition 5.1. *Consider the family of hamiltonians,* $H^{(0)}, H^{(1)}, \ldots, H^{(k)}$, *with*

$$H^{(k)}(I, J, \varphi; z, \bar{z}) = Q^{(k)}(I, J; z, \bar{z}) + S^{(k)}(I, J, \varphi; z, \bar{z}) + R^{(k)}(I, J, \varphi; z, \bar{z}) + \text{const.},$$

and suppose that there is a set $\mathcal{Q}^{(k)} \subset B(\nu^{(0)}, I^0) = \{I \in \mathbb{R}^N \,|\, |I - I^0| < \nu^0\}$ *such that* $H^{(k)}$ *is analytic on* $D(\tilde{I}, \nu^{(k)}, \rho^{(k)}, \sigma^{(k)}, \underline{\tau}^{(k)}; \mathbb{L}^{(k)}, \underline{\mathcal{T}}^{(k)})$ *for every k-admissible sequence* $\mathbb{L}^{(k)}$ *and every* $\tilde{I} \in \mathcal{Q}^{(k)}$.

The terms $Q^{(k)}$ *in* $H^{(k)}$ *are all of second order with respect to the notion of order of terms in a power series defined at the end of the previous section. All terms in* $S^{(k)}$ *are of third order or higher.*

Suppose further that the following inductive hypotheses hold: Recall that $\Omega = (\omega_1, \ldots, \omega_N)$. *Then*

$$\text{(k.1)} \qquad Q^{(k)}(I, J; z, \bar{z}) = \Omega \cdot J + f^{(k)}(I) \cdot J + \sum_{j \geqq N+1} \omega_j z_j \bar{z}_j + \sum_{j \geqq N+1} g_j^{(k)}(I) z_j \bar{z}_j.$$

The functions $f^{(k)}(I)$ *and* $g_j^{(k)}(I)$ *are defined for all* $I \in \{I' \in \mathbb{C}^N \,|\, \operatorname{dist}(I', \mathcal{Q}^{(k)}) < \nu^{(k)}\}$. *We assume that*

$$f_j^{(0)}(I) = g_j^{(0)}(I) = 0 \quad \text{while if} \quad k = 1,$$
$$\sup_I |f_j^{(1)}(I) - \tilde{f}_j^{(1)}(I)| < \varepsilon^{7/6}, \qquad j = 1, \ldots, N,$$
$$\sup_I |g_j^{(1)}(I) - \tilde{g}_j^{(1)}(I)| < \varepsilon^{7/6}/j^{2\zeta_k}, \quad j \geqq N + 1.$$

(Recall that $\tilde{f}^{(1)}$ *and* $\tilde{g}^{(1)}$ *were defined in Eqs. (4.10) and (4.12).) If* $k > 1$ *we require* $f_j^{(k)}(I) = \sum_{l=2}^{k} (f_j^{(l)}(I) - f_j^{(l-1)}(I))$ *and* $g_j^{(k)}(I) = \sum_{l=2}^{k} (g_j^{(l)} - g_j^{(l-1)}(I))$, *with*

$$\sup_I |f_j^{(l)}(I) - f_j^{(l-1)}(I)| < \varepsilon \varepsilon_l^{1/3}, \qquad j = 1, \ldots, N,$$
$$\sup_I |g_j^{(l)}(I) - g_j^{(l-1)}(I)| < \varepsilon \varepsilon_l^{1/3}/j^{2\zeta_k}, \quad j \geqq N + 1,$$

for $l = 2, \ldots, k$.

Suppose further that for any k-admissible set $\mathbb{L}^{(k)}$ *one has the estimates*

$$\text{(k.2)} \qquad \|S^{(k)}\| \leqq C_k \varepsilon_0,$$

$$\text{(k.3)} \qquad \|R^{(k)}\| \leqq \varepsilon_k$$

on the domain $D(\tilde{I}, \nu^{(k)}, \rho^{(k)}, \sigma^{(k)}, \underline{\tau}^{(k)}; \mathbb{L}^{(k)}, \underline{\mathcal{T}}^{(k)})$, *with* $C_k = \prod_{j=1}^{k+1} (1 + j^{-2})$.

Then if ε_0 *is sufficiently small, (this smallness conditions is independent of k) there exists a set* $\mathcal{Q}^{(k+1)} \subset \mathcal{Q}^{(k)}$ *with* $\operatorname{meas}(\mathcal{Q}^{(k+1)}) \geqq (1 - \mathcal{O}(1/(l+1)^2 |\log \varepsilon_0|)) \operatorname{meas}(\mathcal{Q}^{(k)})$ *such that if* $\tilde{I} \in \mathcal{Q}^{(k+1)}$, *there exists a canonical transformation,* $C^{(k)}$, *analytic on* $D(\tilde{I}, \nu^{(k+1)}, \rho^{(k+1)}, \sigma^{(k+1)}, \underline{\tau}^{(k+1)}; \mathbb{L}^{(k+1)}, \underline{\mathcal{T}}^{(k+1)}) \, (\equiv D_{k+1})$ *for any* $(k+1)$*-admissible set* $\mathbb{L}^{(k+1)}$, *and mapping this domain into* $D(\tilde{I}, \nu^{(k)}, \rho^{(k)}, \sigma^{(k)}, \underline{\tau}^{(k)}; \mathbb{L}^{(k+1)}, \underline{\mathcal{T}}^{(k)}) (\equiv D_k)$.

(Note that any $(k+1)$-admissible domain is automatically k-admissible.) Further, the hamiltonian $H^{(k+1)} = H^{(k)} \circ C^{(k)} \equiv Q^{(k+1)} + S^{(k+1)} + R^{(k+1)} + \text{const.}$ *satisfies* $(k+1.1) - (k+1.3)$.

Finally, we note that if z_j *and* $\bar{z}_j$ *are sequences in the domain* D_{k+1}, *which satisfy* $\max(|z_j|, |\bar{z}_j|) \leqq C_k/j^{9/10}$, *with* C_k *the constants defined above, then the transformation* $C^{(k)}$ *maps them into points in the domain* D_k, *obeying a similar estimate, but with the numerator replaced by* C_{k+1}.

The proof of this rather long proposition is given in the following sections. We conclude the present section by showing how it implies Theorem 3.4.

If we take the hamiltonian (4.5) to be $H^{(0)}$, and set $\mathscr{Q}^{(0)} = B(v^{(0)}, I^0)$, then Corollary 4.6, plus the definitions of $v^{(0)}, \rho^{(0)}, \underline{\tau}^{(0)}$ and $\underline{\mathscr{T}}^{(0)}$ show that the induction hypotheses (0.1)–(0.3) are satisfied.

The estimates on the set $\mathscr{Q}^{(k)}$ imply that $\mathscr{Q}^{(\infty)} = \bigcap_{k \geqq 0} \mathscr{Q}^{(k)}$ satisfies meas $(\mathscr{Q}^{(\infty)}) \geqq (1 - \mathscr{O}(1/|\log \varepsilon_0|) \operatorname{meas}(B(v^{(0)}, I^0))$. If $I \in \mathscr{Q}^{(\infty)}$, the induction step may be repeated infinitely many times. This has the effect of "killing" the remainder term $R^{(k)}$. In particular, if we look at the hamiltonian vector field associated with $H^{(k)}$, we see that in the limit $k \to \infty$, the point $J = z = \bar{z} = 0$ is left invariant by the associated flow. Thus, we get a quasi-periodic trajectory given by $\varphi(t) = \Omega^{\infty} t + \varphi(0)$ for some N-vector Ω^{∞}.

That this gives a quasi-periodic orbit for our original hamiltonian, $H^{(0)}$, follows from Sect. 7 (see the discussion following Corollary 7.5) where we prove that the $\lim_{n \to \infty} C^{(0)} \circ C^{(1)} \circ \dots \circ C^{(n)}(0, \Omega^{\infty} t + \varphi(0); 0, 0)$ exists. Furthermore, the last statement of Proposition 5.1 implies that the z_j, or $\bar{z}_j$ component of this limit must satisfy $\max(|z_j(t)|, |\bar{z}_j(t)|) \leqq c/j^{9/10}$. If we now reexpress this quasi-periodic orbit in terms of $q_j(t) = (i/\sqrt{2\omega_j})(z_j - \bar{z}_j)$ we see that $|q_j(t)| \leqq c/j^{14/10}$, verifying the final claim of Theorem 3.4.

In order to prove Theorem 2.3, we note that we get a quasi-periodic orbit for each point $I \in \mathscr{Q}^{(\infty)}$, whose frequency vector is $\Omega^{(\infty)}(I) = \lim_{k \to \infty} \Omega^{(k)}(I)$. (The existence of this limit follows from Proposition 5.1). In Sect. 9 we prove that the set $\{\Omega | \Omega = \Omega^{(\infty)}(I), I \in \mathscr{Q}^{(\infty)}\}$ has large Lebesgue measure, which completes the proof of Theorem 2.3.

VI. The Hamilton–Jacobi Equation

We construct the canonical transformation, $C^{(k)}$, of Proposition 5.1 as the time one map of the hamiltonian flow whose hamiltonian, χ, approximately solves the linearized Hamilton–Jacobi equation

$$H^{(k)} + \{\chi, H^{(k)}\} = H^{(k+1)}. \tag{6.1}$$

As mentioned in Sect. 4, we need only kill those terms in $H^{(k)}$ of low order. For this reason, we only have to retain terms of order two or less in the power series for χ. Thus, we write χ as

$$\chi = \chi_0 + \chi_1 + \chi_2,$$

where χ_i contains terms of order i.

Note that if F_m and G_n are two terms of a power series with orders m and n, respectively, then $\{F_m, G_n\}$ has order $m+n-2$. Thus, if we write $R^{(k)} = R_0^{(k)} + R_1^{(k)} + R_2^{(k)}$, and $S^{(k)}$ as $S^{(k)} = S_3^{(k)} + S_4^{(k)} + S_{\geqq}^{(k)}$, with $S_{\geqq}^{(k)} = \sum_{j \geqq 5} S_j^{(k)}$, we can study Eq. (6.1) order by order.

Looking just at the zero$^{\text{th}}$ order terms gives

$$R_0^{(k)} + \{\chi_0, Q^{(k)}\} + \{\chi_0, R_2^{(k)}\} + \{\chi_1, R_1^{(k)}\} + \{\chi_2, R_0^{(k)}\} = R_0^{(k+1)}. \tag{6.2}$$

We expect χ to be of the same order of magnitude as $R^{(k)}$, i.e. $\mathcal{O}(\varepsilon_k)$, and furthermore expect $R^{(k+1)} \sim \mathcal{O}(\varepsilon_k^2)$. Thus, without worsening the approximation made in linearizing the Hamilton–Jacobi equation we may ignore the last three terms on the left-hand side of this equation, and the term on the right-hand side, and choose χ_0 to satisfy

$$R_0^{(k)} + \{\chi_0, Q^{(k)}\} = 0. \tag{6.3}$$

If we repeat this process for terms of first and second order, we find that χ_1 and χ_2 should solve the equations

$$R_1^{(k)} + \{\chi_0, S_3^{(k)}\} + \{\chi_1, Q^{(k)}\} = 0 \tag{6.4}$$

and

$$R_2^{(k)} + \{\chi_0, S_4^{(k)}\} + \{\chi_1, S_3^{(k)}\} + \{\chi_2, Q_2^{(k)}\} = Q_2^{(k+1)} - Q_2^{(k)}. \tag{6.5}$$

We now expand all the functions in (6.3)–(6.5) in Fourier–Laurent series and (formally) solve for the coefficients of χ. Note that χ_0 must be a function of φ alone and an easy calculation shows

$$\chi_0(\varphi) = \sum_{\substack{n \in \mathbb{Z}^N \\ n \neq 0}} \frac{\hat{R}^{(k)}(I, 0, n; 0, 0) e^{in \cdot \varphi}}{in \cdot \Omega^{(k)}(I)}, \tag{6.6}$$

where $\Omega^{(k)}(I) = \Omega + f^{(k)}(I)$. We will suppress the dependence of χ on I to slightly simplify the notation. (Note that we have omitted the term $\hat{R}^{(k)}(I, 0, 0; 0, 0)$—this is absorbed in the constant term in the hamiltonian.) We return in a moment to discuss the convergence of this sum, but first we write down similar expressions for χ_1 and χ_2. The function χ_1 will depend linearly on z or $\bar{z}$ but must be independent of J. If we define $v^1(I, \varphi; z, \bar{z}) = \{\chi_0, S_3^{(k)}\}$ we find

$$\chi_1(\varphi; z, \bar{z}) = \sum_{\substack{n \in \mathbb{Z}^N \\ j \geqq N+1}} \left\{ \frac{[\hat{R}^{(k)}(I, 0, n; \delta_j, 0) + \hat{v}^1(I, 0, n; \delta_j, 0)] z_j e^{in \cdot \varphi}}{i(n \cdot \Omega^{(k)}(I) - \omega_j^{(k)}(I))} + \frac{[\hat{R}^{(k)}(I, 0, n; 0, \delta_j) + \hat{v}^1(I, 0, n; 0, \delta_j)] \bar{z}_j e^{in \cdot \varphi}}{i(n \cdot \Omega^{(k)}(I) + \omega_j^{(k)}(I))} \right\}. \tag{6.7}$$

In this expression $\delta_j = (0, \ldots, 0, 1, 0, \ldots)$, with a one in the j^{th} component and zeros elsewhere, and $\omega_j^{(k)}(I) = \omega_j + g_j^{(k)}(I)$.

Finally, defining $v^2(I, J, \varphi; z, \bar{z}) = \{\chi_0, S_4^{(k)}\}$ and $v^3(I, J, \varphi; z, \bar{z}) = \{\chi_1, S_3^{(k)}\}$, we can write down the (formal) expression for χ_2. To slightly simplify the results we note that χ_2 will consist of two pieces—one of which is linear in J and independent of z and $\bar{z}$, which we denote by χ_2^J, and a second which is quadratic in z and $\bar{z}$, but independent of I, which we denote by χ_2^z. Then

$$\chi_2^J(J,\varphi)=\sum_{\substack{n\in\mathbb{Z}^N\\ n\neq 0\\ l\cdot l_i\geqq 0,|l|=1}}\frac{[\hat{R}^{(k)}(I,l,n;0,0)+\hat{v}^2(I,l,n;0,0)+\hat{v}^3(I,l,n;0,0)]J^l e^{in\cdot\varphi}}{in\cdot\Omega^{(k)}(I)}, \tag{6.8}$$

while

$$\chi_2^z(\varphi;z,\bar{z})=\sum_{\substack{n\in\mathbb{Z}^N\\ \alpha_i,\beta_i\in\{0,1,2\}\\ |\alpha|+|\beta|=2\\ |n|+|\alpha-\beta|\neq 0}}\frac{[\hat{R}^{(k)}(I,0,n;\alpha,\beta)+\hat{v}^2(I,0,n;\alpha,\beta)+\hat{v}^3(I,0,n;\alpha,\beta)]}{i(n\cdot\Omega^{(k)}(I)-(\alpha-\beta)\cdot\omega^{(k)}(I))}\times z^\alpha\bar{z}^\beta e^{in\cdot\varphi}. \tag{6.9}$$

As a first step in bounding the various terms in the generating function we discuss the denominators of these expressions. We first remark that we can restrict the sums somewhat by noting that the numerators are the Fourier coefficients of analytic functions. Cauchy's theorem then allows us to estimate these coefficients by $c\cdot e^{-\sigma|n|}$, where σ is the width of the analyticity domain in the φ variables. Thus, terms with $|n|>M_k\sim\mathcal{O}(|\ln\varepsilon_k|)$ will be of the same order of magnitude as things we have already discarded, *so we will restrict all the sums to* $|n|\leqq M_k$. This reduces the number of small denominator conditions that we have to impose at each step.

We then have the following sequence of lemmas, bounding the various terms in the generating function.

Lemma 6.1. *If $\tilde{I}\in\mathcal{Q}^{(k)}$ and $\mathbb{L}^{(k)}$ is any k-admissible set, then on $D(\tilde{I},v^{(k)},\rho^{(k)},\sigma^{(k)}-\delta^{(k)},\underline{\tau}^{(k)};\mathbb{L}^{(k)},\underline{\mathcal{T}}^{(k)})$, one has*

$$\|\chi_0\|\leqq E_0(\varepsilon_k),$$

with $E_0(\varepsilon_k)=D_k^{(1)}\varepsilon_k(2\tau/\delta^{(k)})^\tau(1-e^{-\delta^{(k)}/2})^{-N}$.

Proof. By the estimates of Lemma 9.2 on the small denominators which appear in (6.6), and a Cauchy estimate which bounds $|\hat{R}^{(k)}(I,0,n;0,0)|\leqq\varepsilon_k e^{-\sigma^{(k)}|n|}$, we can bound $\|\chi_0\|$ by

$$\begin{aligned}\sum_{0\neq|n|\leqq M_k}D_k^{(1)}\varepsilon_k|n|^\tau e^{-\delta^{(k)}|n|}&\leqq D_k^{(1)}\varepsilon_k\left(\frac{2\tau}{\delta^{(k)}}\right)^\tau\times\sum_{n\neq 0}e^{-\delta^{(k)}|n|/2}\\&\leqq D_k^{(1)}\varepsilon_k(2\tau/\delta^{(k)})^\tau(1-e^{-\delta^{(k)}/2})^{-N}.\end{aligned}$$

We now turn our attention to a bound for χ_1. We see from (6.7), that in order to bound χ_1 we need estimates on $v^1=\{\chi_0,S_3^{(k)}\}=-\frac{\partial\chi_0}{\partial\varphi}\cdot\frac{\partial S_3^{(k)}}{\partial J}\left(=-\sum_{l=1}^N\frac{\partial\chi_0}{\partial\varphi_l}\frac{\partial S_3^{(k)}}{\partial J_l}\right)$. We can easily bound the terms in this sum on domains $D(\tilde{I},v^{(k)},\rho^{(k)}/2,\sigma^{(k)}-2\delta^{(k)},\underline{\tau}^{(k)};\mathbb{L}^{(k)},\underline{\mathcal{T}}^{(k)})$ with a pair of dimensional estimates and we find $\|v^1\|\leqq 2C_kN\varepsilon_0E_0(\varepsilon_k)/\rho^{(k)}\delta^{(k)}$. (We use the terms "Cauchy estimate" and "dimensional estimate" interchangeably.)

Let $\hat{\mathbb{L}}$ be a set of positive integers satisfying $|\hat{\mathbb{L}}\cap\mathbb{I}_j|\leqq\max(0,\mathcal{L}(k,j)-1)$. Then $\hat{\mathbb{L}}\cup\{j\}$ is k-admissible for all $j>N_{k-M}$, and we can bound $|\hat{R}^{(k)}(I,0,n;0,\delta_j)|+$

$\hat{v}^1(I, 0, n; 0, \delta_j)|$ by

$$\frac{[(2C_k N\varepsilon_0 E_0(\varepsilon_k)/\rho^{(k)}\delta^{(k)}) + \varepsilon_k] e^{-(\sigma^{(k)} - 2\delta^{(k)})|n|}}{t_j}$$

on $D(\tilde{I}, v^{(k)}, \rho^{(k)}/2, \sigma^{(k)} - 2\delta^{(k)}, \tau^{(k)}; \hat{\mathbb{L}}, \mathscr{T}^{(k)})$, where $t_j = \tau_j^{(k)}$ if $j \leqq N_{k-M}$ and $t_j = \mathscr{T}_j^{(k)}$ otherwise. The estimates of Corollary 9.4 imply that

$$|n \cdot \Omega^{(k)}(I) \pm \omega_j^{(k)}(I)|^{-1} \leqq \tilde{c} D_k^{(1)} |n|^{\tau+1} j^{-1},$$

where $\tilde{c}$ is some constant. Thus, we immediately have the bound

$$\begin{aligned} \left\| \frac{\partial \chi_1}{\partial z_j} \right\| + \left\| \frac{\partial \chi_1}{\partial \bar{z}_j} \right\| &\leqq \sum_{0 \neq |n| \leqq M_k} \frac{\tilde{c} D_k^{(1)} [(2C_k N\varepsilon_0 E_0(\varepsilon_k)/\rho^{(k)}\delta^{(k)}) + \varepsilon_k] |n|^{\tau+1} e^{-\delta^{(k)}|n|}}{j t_j} \\ &\leqq \frac{\tilde{c} D_k^{(1)} [(2C_k N\varepsilon_0 E_0(\varepsilon_k)/\rho^{(k)}\delta^{(k)}) + \varepsilon_k]}{j t_j} \left(\frac{4\tau}{\delta^{(k)}} \right)^{\tau+1} (1 - e^{-\delta^{(k)}/2})^N \\ &\equiv E_1(\varepsilon_k)/j t_j, \end{aligned} \tag{6.10}$$

on the domain $D(\tilde{I}, v^{(k)}, \rho^{(k)}/2, \sigma^{(k)} - 3\delta^{(k)}, \tau^{(k)}; \hat{\mathbb{L}}, \mathscr{T}^{(k)})$. Note that the sum over $|n|$ was bounded in the same way as in Lemma 6.1.

Remark. One may think of the various constants $E_j(\varepsilon_k)$ that we have defined (and will define below) as being roughly of order ε_k.

We bound $\|\chi_1\|$ on these same domains, by bounding the factor of z_j or $\bar{z}_j$ in the numerator by $\mathscr{T}_j$ if $j \in \hat{\mathbb{L}}$ and τ_j otherwise. We must then sum both over $0 < |n| \leqq M_k$ and $j \geqq N + 1$. The sum over n goes exactly as before—we don't comment further on that and we are left with

$$\|\chi_1\| \leqq 2E_1(\varepsilon_k) \times \left\{ \sum_{j \in \hat{\mathbb{L}}} \left(\frac{1}{j} \right) + \sum_{j \notin \hat{\mathbb{L}}} \frac{\tau_j^{(k)}}{j \mathscr{T}_j^{(k)}} \right\}. \tag{6.11}$$

The first of these sums is bounded by

$$\begin{aligned} \sum_{l=0}^{\infty} \sum_{j \in \hat{\mathbb{L}} \cap \mathbb{I}_l} \left(\frac{1}{j} \right) \leqq \sum_{l=0}^{\infty} \frac{\mathscr{L}(k, l)}{N_l} \leqq 2^4 M \sum_{l=0}^{\infty} (l+1) \varepsilon_0^{-(1+\gamma)^l} \\ \leqq 2^5 M, \text{ for } \varepsilon_0 \text{ sufficiently small.} \end{aligned}$$

The second sum is immediately bounded by noting that

$$(\tau_j^{(k)}/j t_j) \leqq \frac{(1 + N_{k-M})}{j^{3/2}}, \quad \text{so} \quad \sum_{j \notin \hat{\mathbb{L}}} \frac{\tau_j^{(k)}}{j t_j} \leqq \sum_{j=1}^{\infty} (1 + N_{k-M}) j^{-3/2} \leqq 4(1 + N_{k-M}).$$

Combining (6.10) and (6.11) we have

Lemma 6.2. *If $\tilde{I} \in \mathscr{Q}^{(k)}$, $\hat{\mathbb{L}}$ is as defined in the previous paragraph and ε_0 is sufficiently small, we have*

$$\left\| \frac{\partial \chi_1}{\partial z_j} \right\| + \left\| \frac{\partial \chi_1}{\partial \bar{z}_j} \right\| \leqq 2E_1(\varepsilon_k)/j t_j$$

and

$$\|\chi_1\| \leqq 2^6 E_1(\varepsilon_k)(M+1)(1+N_{k-M})$$

on $D(\tilde{I}, v^{(k)}, \rho^{(k)}/2, \sigma^{(k)} - 3\delta^{(k)}, \tau^{(k)}; \hat{\mathbb{L}}, \mathscr{T}^{(k)})$.

We now turn our attention to χ_2^J and χ_2^z. To bound these two functions we must first bound

$$v^2 = \{\chi_0, S_4^{(k)}\} = -\sum_{l=1}^{N} \frac{\partial \chi_0}{\partial \varphi_l} \frac{\partial S_4^{(k)}}{\partial J_l},$$

and

$$v^3 = \{\chi_1, S_3^{(k)}\} = -\sum_{l=1}^{N} \frac{\partial \chi_1}{\partial \varphi_l} \frac{\partial S_3^{(k)}}{\partial J_l} + i \sum_{l=N+1}^{\infty} \left(\frac{\partial \chi_1}{\partial z_l} \cdot \frac{\partial S_3^{(k)}}{\partial \bar{z}_l} - \frac{\partial \chi_1}{\partial \bar{z}_l} \cdot \frac{\partial S_3^{(k)}}{\partial z_l} \right).$$

We bound v^2 on $D(\tilde{I}, v^{(k)}, \rho^{(k)}/2, \sigma^{(k)} - 2\delta^{(k)}, \underline{\tau}^{(k)}; \mathbb{L}^{(k)}, \mathscr{T}^{(k)})$ with the help of Lemma 6.1 and a pair of dimensional estimates and find $\|v^2\| \leqq 2NC_k \varepsilon_0 E_0(\varepsilon_k)/\rho^{(k)}\delta^{(k)}$. The terms in the first sum in the definition of v^3 are bounded in the same way as v^2. To bound the terms in the second sum we use the estimates of Lemma 6.2 to control the factors of $\partial \chi_1/\partial z_j$ and $\partial \chi_1/\partial \bar{z}_j$. By the induction hypothesis, combined with a dimensional estimate we have $\|\partial S_3^{(k)}/\partial \bar{z}_l\| + \|\partial S_3^{(k)}/\partial z_l\| \leqq 2^2 C_k \varepsilon_0 (N_{k-M} + 1)/t_l$, on any domain $D(\tilde{I}, v^{(k)}, \rho^{(k)}/2, \sigma^{(k)} - 3\delta^{(k)}, \underline{\tau}^{(k)}; \hat{\mathbb{L}}, \mathscr{T}^{(k)}/(1+(k+1)^{-2}))$, where $\hat{\mathbb{L}}$ and t_l are defined as they were in Lemma 6.2. (We emphasize once again that we are using the fact that if $l \geqq N_{k-M}$ this domain is contained in $D(\tilde{I}, v^{(k)}, \rho^{(k)}/2, \sigma^{(k)} - 3\delta^{(k)}, \underline{\tau}^{(k)}; \hat{\mathbb{L}} \cup \{l\}, \mathscr{T}^{(k)}/(1+(k+1)^{-2}))$ and then estimating $\partial S_3^{(k)}/\partial z_l$ and $\partial S_3^{(k)}/\partial \bar{z}_l$ on this domain.) Since $1/lt_l^2 \leqq 2^8 \varepsilon^{-1/4}/l^{(9/8)}$, we have

$$\begin{aligned} \|v^3\| &\leqq 2^7 N C_k \varepsilon_0 E_1(\varepsilon_k)/\rho^{(k)}\delta^{(k)} + \sum_{l=N+1}^{\infty} 2^{12} C_k \varepsilon^{-1/4} \varepsilon_0 E_1(\varepsilon_k)/l^{9/8} \\ &\leqq 2^7 C_k \varepsilon^{-1/4} \varepsilon_0 E_1(\varepsilon_k)[(N/\rho^{(k)}\delta^{(k)}) + 2^8], \end{aligned} \tag{6.12}$$

on $D(\tilde{I}, v^{(k)}, \rho^{(k)}/2, \sigma^{(k)} - 3\delta^{(k)}, \underline{\tau}^{(k)}; \hat{\mathbb{L}}, \mathscr{T}^{(k)}/(1+(k+1)^{-2}))$.

With these estimates in hand it is easy to bound χ_2^J. We bound

$$|\hat{R}^{(k)}(I,l,n;0,0)| + |\hat{v}^2(I,l,n;0,0)| + \hat{v}^3(I,l,n;0,0)|$$

by dimensional estimates, $|J^l|$ by $\rho^{(k)}/2$, and the denominator by the estimates of Sect. 9. We find

Lemma 6.3. *If $\tilde{I} \in \mathscr{Q}^{(k)}$, $\hat{\mathbb{L}}$ is as in the paragraph preceding Lemma 6.2, and ε_0 is sufficiently small then,*

$$\begin{aligned} \|\chi_2^J\| &\leqq 2^4 \{\varepsilon_k + 2^7 C_k \varepsilon^{-1/4} \varepsilon_0 E_1(\varepsilon_k)[(N/\rho^{(k)}\delta^{(k)}) + 2^8] + 2NC_k \varepsilon_0 E_0(\varepsilon_k)/\rho^{(k)}\delta^{(k)}\} \\ &\quad \cdot D_k^{(2)} (4\tau/\delta^{(k)})^{2\tau} (1 - e^{-\delta^{(k)}/2})^{-N} \\ &\equiv E_2^J(\varepsilon_k), \end{aligned}$$

on $D(\tilde{I}, v^{(k)}, \rho^{(k)}/4, \sigma^{(k)} - 4\delta^{(k)}, \tau^{(k)}, \hat{\mathbb{L}}, \mathscr{T}^{(k)}/(1+(k+1)^{-2}))$.

The most difficult piece of the generating function to control is χ_2^z which we now study. We will estimate $\partial^2 \chi_2^z/\partial z_i \partial \bar{z}_j$, $\partial \chi_2^z/\partial z_j$, and χ_2^z. Other derivatives are handled in an analogous fashion. First note that

$$\frac{\partial^2 \chi_2^z}{\partial z_i \partial \bar{z}_j} = \sum_{\substack{|n| \leqq M_k \\ (|n| \neq 0 \text{ if } i=j)}} \frac{[\hat{R}^{(k)}(I,0,n;\delta_i,\delta_j) + \hat{v}^2(I,0,n;\delta_i,\delta_j) + \hat{v}^3(I,0,n;\delta_i,\delta_j)]e^{in\cdot\varphi}}{i[n\cdot\Omega^{(k)}(I) - (\omega_i^{(k)}(I) - \omega_j^{(k)}(I))]}.$$

The estimates of Corollary 9.4 imply that if $\tilde{I} \in \mathscr{Q}^{(k)}$; the denominator in this expression may be bounded by

$$|n\cdot\Omega^{(k)}(I) - (\omega_i^{(k)}(I) - \omega_j^{(k)}(I))|^{-1} \leqq \tilde{C} D_k^{(2)} |n|^{4\tau+2}(1+|i-j|)^{-1},$$

with $\tilde{C}$ a fixed constant. On the other hand, dimensional estimates allow us to bound the numerator by

$$2^6 k^2 [\varepsilon_k + \|v^2\| + \|v^3\|] e^{-(\sigma^{(k)} - 3\delta^{(k)})|n|}/t_i t_j, \tag{6.13}$$

where

$$t_i = \begin{cases} \tau_i^{(k)} & \text{if} \quad 1 \leqq i \leqq N_{k-M} \\ \mathscr{T}_i^{(k)} & \text{if} \quad i > N_{k-M} \end{cases}.$$

Suppose we define $\tilde{\mathbb{L}}$ to be a set of integers such that $|\tilde{\mathbb{L}} \cap \mathbb{I}_j| \leqq \max(\mathscr{L}(k,j) - 3, 0)$. Then the estimate in (6.13) holds on all domains $D(\tilde{I}, v^{(k)}, \rho^{(k)}/2, \sigma^{(k)} - 3\delta^{(k)}, \tau^{(k)}/(1+2^{-5}(k+1)^{-2})^2; \tilde{\mathbb{L}}, \mathscr{T}^{(k)}/(1+(k+1)^{-2})^2)$. Using estimate (6.13) and our previous estimate on $\|v^2\|$ and $\|v^3\|$ we immediately obtain

$$\left\| \frac{\partial^2 \chi_2^z}{\partial z_i \partial \bar{z}_j} \right\| \leqq 2^6 k^2 \{\varepsilon_k + 2^7 C_k \varepsilon^{-1/4} \varepsilon_0 E_1(\varepsilon_k)[(N/\rho^{(k)}\delta^{(k)}) + 2^8] + 2N C_k \varepsilon_0 E_0(\varepsilon_k)/\rho^{(k)}\delta^{(k)}\}$$
$$\cdot D_k^{(2)} (4\tau/\delta^{(k)})^{4\tau} (1 - e^{-\delta^{(k)}/2})^{-N}/t_i t_j (1+|i-j|). \tag{6.14}$$

For convenience we define the right-hand side of this expression to be $E_2^z(\varepsilon_k)/t_i t_j(1+|i-j|)$.

Turning now to $\partial \chi_z^2/\partial z_i$ we find

$$\frac{\partial^2 \chi_2^z}{\partial z_i} = \sum_{\substack{0 \leqq |n| \leqq M_k \\ j \geqq N+1}} \left\{ \frac{[\hat{R}^{(k)}(I,0,n;\delta_i,\delta_j) + \hat{v}^2(I,0,n;\delta_i,\delta_j) + \hat{v}^3(I,0,n;\delta_i,\delta_j)]\bar{z}_j e^{in\cdot\varphi}}{i[n\cdot\Omega^{(k)}(I) - (\omega_i^{(k)}(I) - \omega_j^{(k)}(I))]} \right.$$
$$\left. + \frac{[\hat{R}^{(k)}(I,0,n;\delta_i+\delta_j,0) + \hat{v}^2(I,0,n;\delta_i+\delta_j,0) + \hat{v}^3(I,0,n;\delta_i+\delta_j,0)]z_j e^{in\cdot\varphi}}{i[n\cdot\Omega^{(k)}(I) - (\omega_i^{(k)}(I) + \omega_j^{(k)}(I))]} \right\}.$$

The first term is, with the exception of the factor of $\bar{z}_j$, precisely the quantity we considered in deriving (6.14). We bound the factor of $|\bar{z}_j|$ by $\tau_j^{(k)}$ if $j \notin \tilde{\mathbb{L}}$ and $\tilde{\mathscr{T}}_j^{(k)}$ otherwise. Here, $\tilde{\mathscr{T}}_j^{(k)} = \mathscr{T}_j^{(k)}/j^{(2^{-12}(k+1)^{-2})}$. The second term is bounded in like fashion (the estimate on the denominator again follows from Sect. 9) and we find that on

$$D(\tilde{I}, v^{(k)}, \rho^{(k)}/2, \sigma^{(k)} - 4\delta^{(k)}, \tau^{(k)}/(1+2^{-5}(k+1)^{-2}); \tilde{\mathbb{L}}, \tilde{\mathscr{T}}^{(k)}/(1+(k+1)^{-2})^2),$$

$$\left\| \frac{\partial \chi_2^z}{\partial z_i} \right\| \leqq 2E_2^z(\varepsilon_k) \times \left\{ \sum_{j \in \tilde{\mathbb{L}}} \left[\frac{\tilde{\mathscr{T}}_j^{(k)}}{\mathscr{T}_j^{(k)}} \right] \frac{1}{t_i(1+|i-j|)} + \sum_{j \notin \tilde{\mathbb{L}}} \frac{\tau_j^{(k)}}{t_i t_j (1+|i-j|)} \right\}. \tag{6.15}$$

The first of these sums is bounded by taking into account the fact that points in

$\tilde{\mathbb{L}}$ are very sparse. Thus, we have

$$\sum_{j\in\tilde{\mathbb{L}}} \frac{1}{t_i(1+|i-j|)}\left[\frac{\tilde{\mathscr{T}}_j^{(k)}}{\mathscr{T}_j^{(k)}}\right] \leqq \frac{1}{t_i}\sum_{l=0}^{\infty}\sum_{j\in\tilde{\mathbb{L}}\cap\mathbb{I}_l} j^{-(1/2^{12}(k+1)^2)} \leqq \frac{1}{t_i}\sum_{l=0}^{\infty}\mathscr{L}(k,l)N_l^{-(1/2^{12}(k+1)^2)}$$
$$\leqq \frac{2^4M}{t_i}\sum_{l=0}^{\infty}(l+1)\varepsilon_0^{(1+\gamma)^l/2^{12}(k+1)^2} \leqq \frac{2^{18}M(k+1)^2}{t_i(1+\gamma)} \quad (6.16)$$

if ε_0 is sufficiently small.

We bound the second sum in (6.15) by noting that if $j \leqq N_{k-M}$, $\tau_j^{(k)}/t_j = 1 \leqq (1+N_{k-M})/j^{3/4}$, while if $j > N_{k-M}$ and $\tau_j^{(k)}/t_j = \tau_j^{(k)}/\mathscr{T}_j^{(k)} \leqq c/j^{3/4}$. Thus, this sum is bounded by $c\cdot((1+N_{k-M})/t_i)\sum_{j\notin\tilde{\mathbb{L}}} 1/(1+|i-j|)j^{3/4} \leqq (1+N_{k-M})/t_i\sqrt{i}$, using the elementary estimate $\sum_{j=1}^{\infty} 1/j^{3/4}(1+|i-j|) \leqq c/\sqrt{i}$. This completes our estimate of $\|\partial\chi_z^2/\partial z_i\|$.

It remains for us to estimate $\chi_2^{\bar{z}}$. Reconsidering the expression for $\chi_2^{\bar{z}}$ we see that all terms in the sum over α and β must satisfy $|\alpha|+|\beta|=2$. This means that $\alpha+\beta = \pm\delta_i \pm \delta_j$, for some choice of i and j, which when combined with a pair of Cauchy estimates allows us to bound

$$|\hat{R}^{(k)}(I,0,n;\alpha,\beta) + v^2(I,0,n;\alpha,\beta) + \hat{v}^3(I,0,n;\alpha,\beta)|$$
$$\leqq 2^6k^2[\varepsilon_k + \|v^2\| + \|v^3\|]e^{-(\sigma^{(k)}-3\delta^{(k)})|n|}/t_it_j.$$

Similarly, the factor of $|z^\alpha||\bar{z}^\beta|$ is bounded by $\tau_i^{(k)}\tau_j^{(k)}$ if $i,j\notin\tilde{\mathbb{L}}$, $\tau_i^{(k)}\tilde{\mathscr{T}}_j^{(k)}$ if $j\in\tilde{\mathbb{L}}$, $i\notin\tilde{\mathbb{L}}$, $\tau_j^{(k)}\tilde{\mathscr{T}}_i^{(k)}$ if $i\in\tilde{\mathbb{L}}$ and $j\notin\tilde{\mathbb{L}}$, and $\tilde{\mathscr{T}}_i^{(k)}\tilde{\mathscr{T}}_j^{(k)}$ if $i,j\in\tilde{\mathbb{L}}$.

In order to bound the denominators we again note that the restrictions on the sum over α and β imply

$$|n\cdot\Omega^{(k)}(I) - (\alpha-\beta)\cdot\omega^{(k)}(I)|^{-1} = |n\cdot\Omega^{(k)}(I) \pm (\omega_i^{(k)}(I) \pm \omega_j^{(k)}(I))|^{-1}.$$

The discussion of Sect. 9 shows that this quantity is largest for the case $\omega_i^{(k)} - \omega_j^{(k)}$, and allows us to bound

$$|n\cdot\Omega^{(k)}(I) - (\alpha-\beta)\cdot\omega^{(k)}(I)|^{-1} \leqq \tilde{c}D_k^{(2)}|n|^{4\tau+2}/(1+|i-j|),$$

if $I\in\mathbb{C}^N$ and $\operatorname{dist}(I,\mathscr{Q}^{(k)}) < v^{(k)}$. Combining these remarks we see that $\|\chi_2^{\bar{z}}\|$ can be bounded on $D(\tilde{I}, v^{(k)}, \rho^{(k)}/2, \sigma^{(k)}-4\delta^{(k)}, \tau^{(k)}/(1+2^{-5}(k+1)^{-2}); \tilde{\mathbb{L}}, \tilde{\mathscr{T}}^{(k)}/(1+(k+1)^{-2})^2)$ by

$$2\tilde{c}E_2^{\bar{z}}(\varepsilon_k)\cdot\left\{\sum_{\substack{i\in\tilde{\mathbb{L}}\\ j\in\tilde{\mathbb{L}}}}\frac{\tilde{\mathscr{T}}_j^{(k)}\tilde{\mathscr{T}}_i^{(k)}}{t_it_j(1+|i-j|)} + \sum_{\substack{i\in\tilde{\mathbb{L}}\\ j\notin\tilde{\mathbb{L}}}}\frac{\tilde{\mathscr{T}}_i^{(k)}\tau_j^{(k)}}{t_it_j(1+|i-j|)} + \sum_{\substack{i\notin\tilde{\mathbb{L}}\\ j\in\tilde{\mathbb{L}}}}\frac{\tilde{\mathscr{T}}_j^{(k)}\tau_i^{(k)}}{t_it_j(1+|i-j|)} + \sum_{\substack{i\notin\tilde{\mathbb{L}}\\ j\notin\tilde{\mathbb{L}}}}\frac{\tau_i^{(k)}\tau_j^{(k)}}{t_it_j(1+|i-j|)}\right\}. \quad (6.17)$$

The first of these sums is bounded by $(2^{18}M(k+1)^2/(1+\gamma))^2$, as we see by comparing it with

$$\left(\sum_{i\in\tilde{\mathbb{L}}}\frac{\tilde{\mathscr{T}}_i^{(k)}}{t_i}\right)^2 \leqq \left(\sum_{l=0}^{\infty}\sum_{i\in\tilde{\mathbb{L}}\cap I_l} i^{-(1/2^{12}(k+1)^2)}\right)^2,$$

and then proceeding exactly as in (6.16). The second sum is bounded by $(2^{18}M(k+1)^2/(1+\gamma))(CN_{k-M})$. To see this, note that for $i\in\tilde{\mathbb{L}}$, $\mathcal{T}_i^{(k)}/t_i \leqq i^{-(1/2^{12}(k+1)^2)}$, while if $j\notin\tilde{\mathbb{L}}$, $\tau_j^{(k)}/t_j \leqq C\cdot(1+N_{k-M})/j^{3/4}$, as we saw above. Thus, this sum is bounded by

$$\begin{aligned} &C(1+N_{k-M})\sum_{\substack{i\in\tilde{\mathbb{L}}\\ j\notin\tilde{\mathbb{L}}}} i^{-(1/2^{12}(k+1)^2)}\frac{1}{j^{3/4}(1+|i-j|)} \\ &\leqq C(1+N_{k-M})\cdot\sum_{i\in\tilde{\mathbb{L}}} i^{-(1/2^{12}(k+1)^2)}/\sqrt{i} \\ &\leqq C(1+N_{k-m})\cdot\sum_{i\in\tilde{\mathbb{L}}} i^{-(1/2^{12}(k+1)^2)} \leqq C(1+N_{k-M})\left(\frac{2^{18}M(1+k)^2}{(1+\gamma)}\right). \end{aligned}$$

The last inequality follows just as in (6.16). Note that the third sum in (6.17) is bounded in precisely the same fashion—it differs only in that the roles of i and j are interchanged. The final sum in (6.17) is bounded by

$$(C(1+N_{k-M})^2)\sum_{\substack{i\notin\tilde{\mathbb{L}}\\ j\notin\tilde{\mathbb{L}}}}\frac{1}{(ij)^{3/4}(1+|i-j|)} \leqq (C(1+N_{k-M}))^2\sum_{i\notin\tilde{\mathbb{L}}}\frac{1}{i^{5/4}} \leqq (C(1+N_{k-M}))^2,$$

where the first of these expressions used our now familiar bound on $(\tau_j^{(k)}/\mathcal{T}_j^{(k)})$.

Combining these estimates we have

Lemma 6.4. *If $\tilde{I}\in\mathcal{Q}^{(k)}$, $\tilde{L}$ is as defined following Eq. (6.13), and ε_0 is sufficiently small, then on $D(\tilde{I},\nu^{(k)},\rho^{(k)}/2,\sigma^{(k)}-4\delta^{(k)},\tau^{(k)}/(1+2^{-5}(k+1)^{-2})^2;\tilde{\mathbb{L}},\mathcal{T}^{(k)}/(1+(k+1)^2)^2)$ we have:*

$$\left\|\frac{\partial^2\chi_2^z}{\partial z_i\partial z_j}\right\| \leqq E_2^z(\varepsilon_k)/t_it_j(1+|i-j|).$$

$$\left\|\frac{\partial\chi_2^z}{\partial z_i}\right\| \leqq 2E_2^z(\varepsilon_k)\cdot\left\{\frac{2^{18}M(k+1)^2}{(1+\gamma)}+\frac{C\cdot(1+N_{k-M})}{\sqrt{i}}\right\}\Big/ t_i$$

$$\begin{aligned} \|\chi_2^z\| \leqq 2E_2^z(\varepsilon_k)\Bigg\{&\left(\frac{2^{18}M(k+1)^2}{(1+\gamma)}\right)^2 + 2C(1+N_{k-M}) \\ &\cdot\left(\frac{2^{18}M(k+1)^2}{(1+\gamma)}\right) + (C(1+N_{k-M}))^2\Bigg\} \equiv \hat{E}_2^z(\varepsilon_k). \end{aligned}$$

VII. The Canonical Change of Variables

We now consider the canonical change of variables given by the time one map of the flow of the hamiltonian vector field, whose hamiltonian is the generating function constructed in the previous section. We must show that this transformation maps $D(\tilde{I},\nu^{(k+1)},\rho^{(k+1)},\sigma^{(k+1)},\underline{\tau}^{(k+1)};\mathbb{L}^{(k+1)},\mathcal{T}^{(k+1)})$ into $D(\tilde{I},\nu^{(k)},\rho^{(k)},\sigma^{(k)},\underline{\tau}^{(k)};\mathbb{L}^{(k+1)},\underline{\mathcal{T}}^{(k)})$. This task is simplified by the special form of the generating function. The equations of motion are:

$$\dot{I}_j = -\frac{\partial\chi}{\partial\varphi_j}, \quad \dot{\varphi}_j = \frac{\partial\chi}{\partial J_j} = \frac{\partial\chi_2^J}{\partial J_j},$$

$$\dot{z}_j = -i\frac{\partial\chi}{\partial\bar{z}_j}, \quad \dot{\bar{z}}_j = i\frac{\partial\chi}{\partial z_j}. \tag{7.1}$$

Note that the $\dot{\varphi}$ equation is easy to deal with. We estimate the right-hand side of that equation using Lemma 6.3, and a dimensional estimate. Just as important is to note that $\partial\chi_2^J/\partial J_j$ is a function only of φ—it has no dependence on J, z, or $\bar{z}$. (It does depend on I, but I is not one of the dynamical variables.) Thus,

$$\dot{\varphi}_j = \frac{\partial\chi_2^J}{\partial J_j}(\varphi),$$

which implies $\varphi_j(t=1) = \varphi_j(t=0) + \Phi_j\,(\varphi(t=0))$, with

$$\|\Phi_j\| \leqq 2^4 E_2^J(\varepsilon_k)/\rho^{(k)}, \tag{7.2}$$

on $D(\tilde{I}, v^{(k)}, \rho^{(k)}/8, \sigma^{(k)} - 5\delta^{(k)}, \underline{\tau}^{(k)}; \hat{\mathbb{L}}, \mathcal{T}^{(k)}/(1+(k+1)^{-2}))$. This implies $\varphi_j(t=1)$ is in $D(\tilde{I}, v^{(k)}, \rho^{(k)}/8, \sigma^{(k)} - 4\delta^{(k)}, \underline{\tau}^{(k)}; \hat{\mathbb{L}}, \mathcal{T}^{(k)}/(1+(k+1)^{-2}))$, provided

$$2^4 E_2^J(\varepsilon_k)/\rho^{(k)} < \delta^{(k)}. \tag{7.3}$$

We now look at the equations of motion for z and $\bar{z}$,

$$\dot{z}_j = -i\frac{\partial\chi}{\partial\bar{z}_j} = -i\frac{\partial\chi_1}{\partial\bar{z}_j} - i\frac{\partial\chi_2^{\bar{z}}}{\partial\bar{z}_j}.$$

(The factors of i in this equation result from the fact the change of variables (4.4) is not canonical.) Note that $\partial\chi_1/\partial\bar{z}_j$ is a function only of φ, while $\partial\chi_2^{\bar{z}}/\partial\bar{z}_j$ is a linear function of z and $\bar{z}$, in addition to its dependence on φ. Analogous remarks hold for the equation for $\dot{\bar{z}}_j$, which is $\dot{\bar{z}}_j = i(\partial\chi_1/\partial z_j) + i(\partial\chi_2^{\bar{z}}/\partial z_j)$. Thus, the equations of motion for $d/dt\begin{pmatrix} z(t) \\ \bar{z}(t) \end{pmatrix}$ are linear, (but non-autonomous) and have the form

$$\frac{d}{dt}\begin{pmatrix} z \\ \bar{z} \end{pmatrix} = V(\varphi(t)) + M(\varphi(t))\begin{pmatrix} z \\ \bar{z} \end{pmatrix}, \tag{7.4}$$

where $V(\varphi(t))$ is an infinite dimensional vector, and $M(\varphi(t))$ is an $\infty \times \infty$ dimensional matrix. Lemmas 6.2 and 6.4 provide strong control on $\sup_t \|V_j(\varphi(t))\|$ and $\sup_t \|M_{ij}(\varphi(t))\|$ which we exploit by using the following:

Lemma 7.1. *Suppose $x(t)$ satisfies the first order system of* ODE's

$$\dot{x} = V(t) + M(t)x.$$

Suppose further that $\bar{V}_j > \sup_t |V_j(t)|$ and $\bar{M}_{ij} \geqq \sup_t |M_{ij}(t)|$, for all i, j. Then if $y_j = \sup_{0 \leqq t \leqq 1} |x_j(t)|$,

$$y_j \leqq [(1-\bar{M})^{-1}(\bar{V} + |x(0)|)]_j, \tag{7.5}$$

provided $(1-\bar{M})^{-1}$ exists.

Proof. We have

$$x_j(t) - x_j(0) = \int_0^t V_j(\tau)d\tau + \int_0^t \sum_k M_{jk}(\tau)x_k(\tau)d\tau.$$

Thus

$$|x_j(t)| \leqq |x_j(0)| + t\cdot\left(\sup_t |V_j(t)|\right) + t\cdot\sum_k\cdot\left(\sup_t |M_{jk}(t)|\right)\left(\sup_{0\leqq\tau\leqq t} |x_k(\tau)|\right).$$

Taking the supremum of both sides for $t\in[0,1]$ gives

$$y_j \leqq |x_j(0)| + \bar{V}_j + \sum_k \bar{M}_{jk} y_k,$$

which immediately yields (7.5) if $(1-M)^{-1}$ is invertible.

Remark. If the vectors $x(t)$ and y are infinite dimensional, as they are in the situation we are interested in, we assume that $x_j(t)$ exists for all $j=1,2,\ldots$ and $t\in[0,1]$. This is easy to check in our example.

We now turn to the estimation of $\bar{V}$ and $\bar{M}$ in our particular case. There is a slight notational complication which arises from the fact that $V(t)$ is in our case a concatenation of two infinite dimensional vectors— $-i(\partial\chi_1/\partial\bar{z}_j)$ and $i(\partial\chi_1/\partial z_j)$. We will write $V(t)=\begin{pmatrix} V^1(t) \\ V^2(t)\end{pmatrix}$, where $V_j^1 = -i(\partial\chi_1/\partial\bar{z}_j)$ and $V_j^2 = i(\partial\chi_1/\partial z_j)$, and similarly $\bar{V} \equiv \begin{pmatrix} \bar{V}^1 \\ \bar{V}^2\end{pmatrix}$. Then Lemma 6.2 implies that we can pick $\bar{V}_j^l = 4E_1(\varepsilon_k)/jt_j$, $l=1,2$ and $j=N+1,N+2,\ldots$. A similar notational problem arises with the matrices $M(t)$ and $\bar{M}$, which we solve by writing $M(t)=\begin{pmatrix} M^{11}(t) & M^{12}(t) \\ M^{21}(t) & M^{22}(t)\end{pmatrix}$, with $M_{ij}^{11} = -i(\partial^2\chi_2^z/\partial\bar{z}_i\partial z_j)$, and similarly for M^{12}, M^{21}, and M^{22}. Lemma 6.4 then implies that we can take

$$\bar{M}_{ij}^{11} = \bar{M}_{ij}^{12} = \bar{M}_{ij}^{21} = \bar{M}_{ij}^{22} = E_2^z(\varepsilon_k)/t_it_j(1+|i-j|), \quad i,j=N+1,N+2,\ldots.$$

Since we wish to estimate our change of variables by comparing it with $(1-\bar{M})^{-1}$, we note that this matrix will also have a block form, namely

$$(1-\bar{M})^{-1} = \begin{pmatrix} 1+A^{11} & A^{12} \\ A^{21} & 1+A^{22}\end{pmatrix}, \tag{7.6}$$

which because of the form of $\bar{M}$, satisfies $A^{11}=A^{12}=A^{21}=A^{22}$. The key step in estimating A^{ij} is

Lemma 7.2. *There exists a constant c, such that for* $n=1,2,\ldots,$

$$(\bar{M}^n)_{ij}^{lm} \leqq [c\varepsilon^{-1/4}E_2^z(\varepsilon_k)(1+N_{k-M})]^n/(1+|i-j|)t_it_j,$$

for $l,m=1,2$, *and* $i,j=N+1,N+2,\ldots.$

Proof. The case $n=1$ is certainly true. We now check it for $n+1$. (We will verify the case $l=m=1$—the other three cases are identical.)

$$(\bar{M}^{n+1})_{ij}^{11} = \sum_l (\bar{M}_{il}^n \bar{M}_{lj})^{11} = 2\sum_l (\bar{M}^n)_{il}^{11}(\bar{M})_{lj}^{11}$$

$$\leqq 2E_2^z(\varepsilon_k)[(c\varepsilon^{-1/4}E_2^z(\varepsilon_k)(1+N_{k-M})]^n \sum_l \frac{1}{(1+|i-l|)(1+|j-l|)t_i t_j t_l^2}$$

$$\leqq \frac{2\varepsilon^{-1/4}E_2^z(\varepsilon_k)[c\varepsilon^{-1/4}E_2^z(\varepsilon_k)(1+N_{k-M})]^n}{(1+|i-j|)t_i t_j} \times (1+N_{k-N})$$

$$\cdot \sum_l \frac{(1+|i-j|)}{(1+|i-l|)(1+|j-l|)l^{4/5}}.$$

This last inequality used two facts—first, if $l \leqq N_{k-M}$, then $1/t_l^2 = (1/\tau_l^{(k)})^2 \leqq 2((1+N_{k-M})/\varepsilon^{1/4}l^{4/5})$, and second, if $l > N_{k-M}$, $1/t_l^2 = (1/\mathscr{T}_l^{(k)})^2 \leqq 2/\varepsilon^{1/4}l^{4/5}$. Since $\sum_l (1+|i-j|)/(1+|i-l|)(1+|j-l|)l^{4/5} < c$, (as one readily sees by bounding the sum by the corresponding integral), the lemma follows. An immediate corollary of the lemma is

Corollary 7.3. *If* $c\varepsilon^{-1/4}E_2^z(\varepsilon_k)(1+N_{k-M}) < 1/2$, *then the matrices* A^{lm} *defined in* (7.6) *satisfy* $|A_{ij}^{lm}| \leqq 2c\varepsilon^{-1/4}E_2^z(\varepsilon_k)(1+N_{k-M})/(1+|i-j|)t_i t_j$.

Proof. This follows immediately from the Neumann series for $(1-\bar{M})^{-1}$.

We can now derive bounds on the canonical transformation.

Proposition 7.4. *Suppose* $(J(0), \varphi(0), z(0), \bar{z}(0)) \in D(\tilde{I}, v^{(k)}, \rho^{(k)}/4, \sigma^{(k)} - 6\delta^{(k)}, \underline{\tau}^{(k)}/(1+2^{-5}(k+1)^{-2})^2; \tilde{\mathbb{L}}, \tilde{\mathscr{T}}^{(k)}/(1+(k+1)^{-2})^2)$, $(\equiv D_i)$. *(Here,* $\tilde{\mathbb{L}}$ *is as in Lemma 6.4). Let* $(J(t), \varphi(t), z(t), \bar{z}(t))$ *be solutions of Eqs. (7.1), defining the canonical transformation, (with initial conditions* $(J(0), \varphi(0), z(0), \bar{z}(0))$*). Then*

$$(J(1), \varphi(1), z(1), \bar{z}(1)) \in D(\tilde{I}, v^{(k)}, \rho^{(k)}/2, \sigma^{(k)} - 5\delta^{(k)}, \underline{\tau}^{(k)}/(1+2^{-5}(k+1)^{-2}), \tilde{\mathbb{L}}, \tilde{\mathscr{T}}^{(k)}/(1+(k+1)^{-2})),$$

$(\equiv D_f)$ *provided*

(i) $2^3 E_2^J(\varepsilon_k) < \delta^{(k)}\rho^{(k)}$,

(ii) $c\varepsilon^{-1/4}E_2^z(\varepsilon_k)(1+N_{k-M}) < 1/2$ *and*

$$\left\{2E_1(\varepsilon_k) + 2\tilde{c}E_1(\varepsilon_k)E_2^z(\varepsilon_k)\varepsilon^{-1/2}(1+N_{k-M}) + \tilde{c}\varepsilon^{-1/4}E_2^z(\varepsilon_k)(1+N_{k-M})^2 \cdot\left[1 + \frac{k^2\tilde{c}}{(1+\gamma)|\log\varepsilon_0|}\right]\right\} \times (1+N_{k-M}) < \frac{2^{-6}\varepsilon^{1/4}}{(k+1)^2},$$

and

(iii) $[E_0(\varepsilon_k) + E_2^J(\varepsilon_k) + 2^6E_1(\varepsilon_k)(M+1) + \hat{E}_2^z(\varepsilon_k)] < \delta^{(k)}\rho^{(k)}/4$

are satisfied.

(*Remark.* Using the inductive definitions of Sect. 5, it is easy to check that these conditions can all be satisfied if ε_0 is sufficiently small.)

Proof. That $\varphi(1)$ is in D_f follows immediately from (7.2), (7.3), and (i).

That $z(1)$ and $\bar{z}(1)$ are in D_f follows from Corollary 7.3 (which is applicable becase of (ii)) and the following argument: By Lemma 7.1 and (7.6) we have

$$\begin{aligned}\sup_{0\leqq t\leqq 1} |z_j(t)| &\leqq \bar{V}_j + |z_j(0)| + \sum_{k=N+1}^{\infty} \{[A_{jk}^{11}(\bar{V}_k + |z_k(0)|)] + [A_{jk}^{12}(\bar{V}_k + |\bar{z}_k(0)|)]\} \\ &\leqq \bar{V}_j + |z_j(0)| + 2\sum_{k=N+1}^{\infty} A_{jk}^{11}\cdot \bar{V}_k + \sum_{k=N+1}^{\infty} A_{jk}^{11}|z_k(0)| + \sum_{k=N+1}^{\infty} A_{jk}^{12}|\bar{z}_k(0)|.\end{aligned} \tag{7.6'}$$

Lemma 6.2 implies $\bar{V}_j \leqq 2E_1(\varepsilon_k)/jt_j$. Thus,

$$\begin{aligned}\sum_{k=N+1}^{\infty} A_{jk}^{11}\bar{V}_k &\leqq 2^2 c E_1(\varepsilon_k) E_2^z(\varepsilon_k)\varepsilon^{-1/4}(1+N_{k-M})\cdot \sum_{k=N+1}^{\infty} \frac{2}{(1+|j-k|)t_j t_k^2 k} \\ &\leqq \tilde{c} E_1(\varepsilon_k) E_2^z(\varepsilon_k)\varepsilon^{-1/2}(1+N_{k-M})/t_j j.\end{aligned}$$

The last inequality followed from the fact that $t_k^{-1} < c(1+N_{k-M})\varepsilon^{-1/4}k^{-4/10}$, and the elementary estimate $\sum_{k=N+1}^{\infty} 1/(1+|j-k|)k^{18/10} < c/j$.

The last two sums are bounded in a similar way. Since $z(0)$ is in D_i, we have

$$\begin{aligned}\sum_{l=N+1}^{\infty} A_{jl}^{11}|z_l(0)| &\leqq 2c\varepsilon^{-1/4}E_2^z(\varepsilon_k)(1+N_{k-M}) \\ &\cdot\left\{\sum_{l\in\tilde{\mathbb{L}}} \frac{\tilde{\mathscr{T}}_l^{(k)}}{(1+|l-j|)t_l t_j} + \sum_{l\notin\tilde{\mathbb{L}}} \frac{\tau_l^{(k)}}{(1+|l-j|)t_l t_j}\right\}.\end{aligned}$$

The second of these sums is bounded by

$$\frac{1}{t_j}\sum_{l=N+1}^{\infty} \frac{c(1+N_{k-M})}{(1+|l-j|)l^{2/3}} \leqq \frac{c(1+N_{k-M})}{t_j},$$

while the first is bounded by

$$\sum_{l\in\tilde{\mathbb{L}}}^{\infty} \frac{c(1+N_{k-M})}{t_j(1+|j-l|)}\left(\frac{\tilde{\mathscr{T}}_l^{(k)}}{\mathscr{T}_l^{(k)}}\right) \leqq \frac{\tilde{c}(1+N_{k-M})k^2}{(1+\gamma)|\log\varepsilon_0|t_j},$$

the last inequality coming from (6.16). The sum over $A_{jl}^{12}|\bar{z}_l(0)|$ is bounded in an identical fashion. Putting this altogether with (7.6′) we have

$$\begin{aligned}\left|\sup_{t\in[0,1]} |z_j(t)| - |z_j(0)|\right| &\leqq \left\{\frac{2E_1(\varepsilon_k)}{j} + \frac{2\tilde{c}E_1(\varepsilon_k)E_2^z(\varepsilon_k)(1+N_{k-M})}{\sqrt{\varepsilon}\, j^{7/10}}\right. \\ &\quad \left. + \tilde{c}\varepsilon^{-1/4}E_2^z(\varepsilon_k)(1+N_{k-M})^2\left[1 + \frac{k^2\tilde{c}}{(1+\gamma)|\log\varepsilon_0|}\right]\right\} \\ &\quad \cdot(1+N_{k-M})/\mathscr{T}_j^{(k)}.\end{aligned} \tag{7.7}$$

If the right-hand side of this inequality is bounded by

$$\tau_j^{(k)}/(1+2^{-5}(k+1)^{-2}) - \tau_j^{(k)}/(1+2^{-5}(k+1)^{-2})^2 = \frac{2^{-5}(k+1)^{-2}}{(1+2^{-5}(k+1)^{-2})^2}\tau_j^{(k)},$$

then $z_j(1)$ will be in D_f. Since $(1/\tau_j^{(k)}\mathscr{T}_j^{(k)}) < \varepsilon^{-1/4}$, for all j, this condition will be satisfied if

$$\Big\{2E_1(\varepsilon_k) + 2\tilde{c}E_1(\varepsilon_k)E_2^z(\varepsilon_k)\varepsilon^{-1/2}(1+N_{k-M}) + \tilde{c}\varepsilon^{-1/4}E_2^z(\varepsilon_k)(1+N_{k-M})^2 \cdot\left[1+\frac{k^2\tilde{c}}{(1+\gamma)|\log\varepsilon_0|}\right]\Big\}(1+N_{k-M}) < \frac{2^{-6}}{(k+1)^2}\varepsilon^{1/4},$$

which is the second half of hypothesis (ii).

Finally, we bound the change in the variable $J(t)$. We know $\dot{J}_i = -\partial\chi/\partial\varphi_i(J,\varphi;z,\bar{z})$. Lemmas 6.1–6.4 allow us to bound $\partial\chi/\partial\varphi_i$ on D_f so we have

$$\|\dot{J}_i\| \leqq \left\|\frac{\partial\chi}{\partial\varphi_i}\right\| \leqq \{E_0(\varepsilon_k) + E_2^J(\varepsilon_k) + 2^6E_1(\varepsilon_k)(M+1) + \hat{E}_2^z(\varepsilon_k)\}/\delta^{(k)}.$$

Thus, if $J(0)$ is in D_i, $J(t)$ will remain in D_f for all $t\in[0,1]$ provided the right-hand side of this inequality is bounded by $\rho^{(k)}/4$. This is insured by (iii), and the proof of the proposition is complete.

Note that we have actually proved more than stated in Proposition 7.4. Denote the canonical transformation constructed above by $C^{(k)}$. The analyticity and invertibility of $C^{(k)}$ come for free in this method because of the analytic dependence of solutions on initial conditions for differential equations with analytic vector fields.

Furthermore, if we write $(\tilde{J},\tilde{\varphi},\tilde{z},\tilde{\bar{z}}) = C^{(k)}(J,\varphi,z,\bar{z})$, then since all the estimates on the vector field whose flow defines $C^{(k)}$ were in terms of the norm $\|\cdot\|$ on D_f, we have actually proved bonds on $(\tilde{J},\tilde{\varphi},\tilde{z},\tilde{\bar{z}})$ as functions of $(J,\varphi,z,\bar{z})$. We collect these results in the following.

Corollary 7.5. *The transformation $C^{(k)}$ is analytic and invertible on D_i. Furthermore, if we regard $(\tilde{J},\tilde{\varphi},\tilde{z},\tilde{\bar{z}}) = C^{(k)}(J,\varphi,z,\bar{z})$, as functions of $J,\varphi,z,\bar{z}$, we have*

$$\|\tilde{J}_l - J_l\| \leqq 2\{E_0(\varepsilon_k) + E_2^J(\varepsilon_k) + 2^6E_1(\varepsilon_k)(M+1) + \hat{E}_2^z(\varepsilon_k)\}/\delta^{(k)},$$

$$\|\tilde{\varphi}_l - \varphi_l\| \leqq 2^4E_2^J(\varepsilon_k)/\rho^{(k)},$$

$$\|\tilde{z}_l - z_l\| + \|\tilde{\bar{z}}_l - \bar{z}_l\| \leqq 2^2\Big\{2E_1(\varepsilon_k) + 2\tilde{c}\varepsilon^{-1/2}E_2^z(\varepsilon_k)(1+N_{k-M}) + \tilde{c}\varepsilon^{-1/4}E_2^z(\varepsilon_1)(1+N_{k-M})^2\left[1+\frac{k^2\tilde{c}}{(1+\gamma)|\log\varepsilon_0|}\right]\Big\}(1+N_{k-M})/\mathscr{T}_j^{(k)}.$$

Proof. This follows by using Gronwall's inequality plus the estimates above.

Note that from these estimates it follows easily that $\lim_{n\to\infty} C^{(0)}\circ\cdots\circ C^{(n)}(0,\Omega^\infty t + \varphi(0);0,0)$ exists, and this gives a quasiperiodic orbit for our original hamiltonian as discussed in Sect. 5.

We will need one more estimate on this change of variables. Suppose the variables $(J,\varphi,z,\bar{z})$ satisfy $|z_j| + |\bar{z}_j| < 1/j^{14/10}$. Then from Lemma 7.1, Corollary 7.3, Eq. (7.6) and the estimates following it, we can estimate the decay of $\tilde{z}_j$ and $\tilde{\bar{z}}_j$. Thus we have

Lemma 7.6. *Assume that the hypotheses of Proposition 7.4 hold and that in addition* $|z_j| + |\bar{z}_j| < C_k/j^{14/10}$. *Then if* $(\tilde{J}, \tilde{\varphi}, \tilde{z}, \tilde{\bar{z}}) = C^{(k)}(J, \varphi, z, \bar{z})$ *we have*

$$\max(|\tilde{z}_j| - |z_j|, |\tilde{\bar{z}}_j| - |\bar{z}_j|) \leqq \{2E_1(\varepsilon_k) + \tilde{c}E_1(\varepsilon_k)E_2^z(\varepsilon_k)\varepsilon^{-1/2}(1 + N_{k-M}) + \bar{c}\varepsilon^{-1/4}E_2^z(\varepsilon_k)(1 + N_{k-M})\}C_k/j^{14/10}.$$

Thus, the transformed variables decay as fast as original ones, and this verifies the last claim in Proposition 5.1.

VIII. The Transformed Hamiltonian

In the present section we study the transformed hamiltonian $H^{(k+1)} = H^{(k)} \circ C^{(k)}$ and show that it satisfies the estimates $(k + 1.1) - (k + 1.3)$ of Proposition 5.1. Recall that $C^{(k)}$ is the time one map of a hamiltonian flow (which we denote F^t) with hamiltonian χ. Thus,

$$\begin{aligned} H^{(k+1)} &= H^{(k)} \circ F^{t=1} = H^{(k)} + H^{(k)} \circ F^1 - H^{(k)} \circ F^0 \\ &= H^{(k)} + \int_0^1 \frac{d}{dt}(H^{(k)} \circ F^t)dt \\ &= H^{(k)} + \int_0^1 \{\chi, H^{(k)}\} \circ F^t dt. \end{aligned} \tag{8.1}$$

In order to bound this quantity we will need several further pieces of information. In particular we will need to know how to bound the Poisson bracket of functions with χ—the generating function constructed in Sect. 6, and we will need to know how to bound the composition of functions with the flow, F^t. We address these questions with the following series of lemmas.

Lemma 8.1. *Suppose G is a function analytic and uniformly bounded on* $D(\tilde{I}, v^{(k)}, \rho^{(k)}, \sigma^{(k)}, \underline{\tau}^{(k)}; \mathbb{L}^k, \mathcal{T}^{(k)})$, *for every k-admissible sequence* $\mathbb{L}^{(k)}$, *and* $\tilde{I} \in \mathcal{Q}^{(k)}$. *Let* $\chi = \chi_0 + \chi_1 + \chi_2$ *be the generating function defined in Sect. 6. Then on the smaller domain* $D(\tilde{I}, v^{(k)}, \rho^{(k)}/4, \sigma^{(k)} - 7\delta^{(k)}, \underline{\tau}^{(k)}/(1 + 2^{-5}(k+1)^{-2})^2; \mathbb{L}', \underline{\mathcal{T}}^{(k)}/(1 + 2^{-5}(k+1)^{-2})^2)$ *one has*

(i) $\|\{\chi_0, G\}\| \leqq (2^2 N E_0(\varepsilon_k)/\rho^{(k)}\delta^{(k)}) \|G\|$,
(ii) $\|\{\chi_1, G\}\| \leqq (2^8 N(M + 1)/\rho^{(k)}\delta^{(k)} + 2^5(1 + N_{k-M})^2 \varepsilon^{-1/4}) E_1(\varepsilon_k) \|G\|$,

and

(iii) $\|\{\chi_2, G\}\| \leqq [2^2(E_2^J(\varepsilon_k))/\rho^{(k)}\delta^{(k)} + c(1 + k^2)^2(1 + N_{k-M})^3 M\varepsilon^{-1/4}E_2^z(\varepsilon_k)] \|G\|$,

provided ε_0 *is sufficiently small.*

Here, $\mathbb{L}'$ is any sequence of integers satisfying $|\mathbb{L}' \cap \mathbb{I}_l| \leqq \max(0, \mathcal{L}(k, l) - 4)$, for $l = 0, 1, 2, \ldots$.

Note that combining (i), (ii), and (iii) we have $\|\{\chi, G\}\| \leqq K(\varepsilon_k)\|G\|$, where $K(\varepsilon_k)$ is obtained by summing the right-hand sides of the above inequalities.

Proof. Since $\{\chi_0, G\} = -\sum_{j=1}^{N} \frac{\partial \chi_0}{\partial \varphi_j}\frac{\partial G}{\partial I_j}$, a pair of dimensional estimates, combined with Lemma 6.1 yields (i).

Next note that $\{\chi_1, G\} = -\sum_{j=1}^{N} \frac{\partial \chi_1}{\partial \varphi_j}\cdot\frac{\partial G}{\partial I_j} + i \sum_{j=N+1}^{\infty}{}' \left(\frac{\partial \chi_1}{\partial z_j}\cdot\frac{\partial G}{\partial \bar{z}_j} - \frac{\partial \chi_1}{\partial \bar{z}_j}\frac{\partial G}{\partial z_j}\right)$. A pair of dimensional estimates again bounds the first sum by $2^8 N(M+1)\|G\| E_1(\varepsilon_k)/\rho^{(k)}\delta^{(k)}$. To bound the second sum we bound the derivatives of G using dimensional estimates as $\|\partial G/\partial z_j\| \leqq 2\|G\|/t_j \leqq 2^2\|G\|(1+N_{k-M})/\varepsilon^{1/8} j^{4/10}$, and similarly for $\|\partial G/\partial \bar{z}_j\|$. Lemma 6.2 bounds $\|\partial \chi_1/\partial z_j\|$ and $\|\partial \chi_1/\partial \bar{z}_j\|$, and one can then perform the sum over j, so that the second sum in $\{\chi_1, G\}$ is bounded by $2^5(1+N_{k-M}^2)^2 \varepsilon^{-1/4}\|G\| E_1(\varepsilon_k)$. Combining this with the previous estimate gives (ii).

Finally, we bound

$$\{\chi_2, G\} = \sum_{l=1}^{N}\left[\frac{\partial \chi_2}{\partial J_l}\cdot\frac{\partial G}{\partial \varphi_l} - \frac{\partial \chi_2}{\partial \varphi_l}\cdot\frac{\partial G}{\partial J_l}\right] + i \sum_{l=N+1}^{\infty}\left[\frac{\partial \chi_2}{\partial z_l}\cdot\frac{\partial G}{\partial \bar{z}_l} - \frac{\partial \chi_2}{\partial \bar{z}_l}\frac{\partial G}{\partial z_l}\right].$$

Using Lemmas 6.3 and 6.4, and a pair of Cauchy estimates the first sum is easily bounded by

$$2^2 N(E_2^J(\varepsilon_k) + \hat{E}_2^z(\varepsilon_k))\|G\|/\rho^{(k)}\delta^{(k)}.$$

To bound the second sum over l we first note that if we define $\hat{\mathbb{L}}$ as we did in the paragraph preceding (6.10), then on $D(\tilde{I}, v^{(k)}, \rho^{(k)}/4, \sigma^{(k)} - 7\delta^{(k)}, \underline{\tau}^{(k)}/(1+2^{-5}(k+1)^{-2}); \hat{\mathbb{L}}, \mathscr{T}^{(k)}/(1+2^{-5}(k+1)^{-2}))$ one can bound $\|\partial G/\partial z_l\|$ and $\|\partial G/\partial \bar{z}_l\|$ by $2^6(1+k)^2\|G\|/t_l \leqq 2^6(1+k)^2\|G\|(1+N_{k-M})/\mathscr{T}_l^{(k)}$, using the fact that $1/t_l \leqq (1+N_{k-M})/\mathscr{T}_l^{(k)}$. (The quantity t_l was defined in the paragraph preceding (6.10).) We must bound the factors of $\|\partial \chi_2/\partial z_l\|$ and $\|\partial \chi_2/\partial \bar{z}_l\|$ by using (6.15), rather than Lemma 6.3, as this estimate is quite delicate. Inserting the bounds from (6.15) we have:

$$\left\|\sum_{l=N+1}^{\infty} \frac{\partial \chi_2}{\partial z_l}\cdot\frac{\partial G}{\partial \bar{z}_l} - \frac{\partial \chi_2}{\partial \bar{z}_i}\cdot\frac{\partial G}{\partial z_l}\right\| \leqq c(1+k^2)(1+N_{k-M})E_2^z(\varepsilon_k)\|G\|$$
$$\cdot \sum_{l=N+1}^{\infty}\left\{\sum_{j\in\tilde{\mathbb{L}}}\left[\frac{\tilde{\mathscr{T}}_j^{(k)}}{\mathscr{T}_j^{(k)}}\right]\frac{1}{\mathscr{T}_l^{(k)} t_l(1+|j-l|)} + \sum_{j\notin\tilde{\mathbb{L}}}\frac{\tau_j^{(k)}}{\mathscr{T}_l^{(k)} t_l t_j(1+|l-j|)}\right\}. \tag{8.2}$$

Here, $\tilde{\mathbb{L}}$ is as defined in the paragraph following (6.13). We now use the fact that $(1/\mathscr{T}_l^{(k)} t_l) \leqq (c(1+N_{k-M})\varepsilon^{-1/4}/l^{3/4})$, and $\tau_j^{(k)}/t_j \leqq c(1+N_{k-M})/j^{3/4}$, whereby we see that the first of these two sums is bounded by

$$\sum_{l=N+1}^{\infty}\sum_{j\in\tilde{\mathbb{L}}}[\tilde{\mathscr{T}}_j^{(k)}/\mathscr{T}_j^{(k)}]\cdot\frac{C(1+N_{k-M})\varepsilon^{-1/4}}{l^{3/4}(1+|j-l|)}$$
$$\leqq C(1+N_{k-M})\varepsilon^{-1/4} \times \sum_{j\in\tilde{\mathbb{L}}}\left\{[\tilde{\mathscr{T}}_j^{(k)}/\mathscr{T}_j^{(k)}] \times \sum_{l=N+1}^{\infty}\frac{1}{l^{3/4}(1+|j-l|)}\right\}$$
$$\leqq C(1+N_{k-M})\varepsilon^{-1/4} \times \sum_{j\in\tilde{\mathbb{L}}}[\tilde{\mathscr{T}}_j^{(k)}/\mathscr{T}_j^{(k)}] \leqq C(1+N_{k-M})M(1+k)^2\varepsilon^{-1/4}/(1+\gamma), \tag{8.3}$$

where the sum over l was bounded by comparing it to an integral, and the sum over j was bounded exactly as in (6.16).

We bound the second sum in (8.2) by

$$c(1+N_{k-M})^2\varepsilon^{-1/4}\sum_{l=N+1}^{\infty}\sum_{j\notin\tilde{\mathbb{L}}}\frac{1}{j^{3/4} l^{3/4}(1+|j-l|)} \leqq c(1+N_{k-M})^2\varepsilon^{-1/4}. \tag{8.4}$$

Here, the double sum over l and j was bounded by comparing it with the double integral

$$\int_2^\infty \left\{ \int_1^\infty \frac{dy}{(1+|x-y|)y^{3/4}} \right\} \frac{dx}{x^{3/4}} \leqq \int_2^\infty \left\{ \frac{dy}{(1+|x-y|)\sqrt{y}} \right\} \frac{dx}{x^{3/4}}.$$

The integral over y can now be performed exactly and it is bounded by $c/(x-1)^{1/3}$ when $x \geqq 2$, so the integral over x is also bounded by a constant. Combining (8.3) and (8.4) gives (iii).

A further fact we will need in bounding the hamiltonian is the following. Suppose $R(J, \varphi; z, \bar{z})$ is analytic on $D(I^0, v, \rho, \sigma, \tau; \mathbb{L}, \mathcal{T}) \equiv D_i$. Define

$$R^+(J, \varphi; z, \bar{z}) = \sum_{\substack{n \in \mathbb{Z}^N \\ |n| > M \\ l \in (\mathbb{Z}^+)^N \\ \alpha, \beta}} \hat{R}(l, n; \alpha, \beta) e^{in \cdot \varphi} J^l z^\alpha \bar{z}^\beta.$$

R^+ is obtained from R just by omitting those Fourier coefficients with small indices. Define $R^- = R - R^+$. We then have

Lemma 8.2. *On $D_f \equiv D(I^0, v, \rho, \sigma - \delta, \tau; \mathbb{L}, \mathcal{T})$ we have*

(i) $\| R^- \|_{D_f} \leqq \| R \|_{D_i}$,
(ii) $\| R^+ \|_{D_f} \leqq e^{-\delta M} \| R \|_{D_i}$.

Proof. The first inequality is immediate. The second follows by noting that

$$\begin{aligned} \| R^+ \|_{D_f} &= \sum_{\substack{n \in \mathbb{Z}^N \\ |n| > M \\ l \in (\mathbb{Z}^+)^N \\ \alpha, \beta}} |\hat{R}(l, n; \alpha, \beta)| e^{|n|(\sigma - \delta)} \rho^{|l|} \tau^{\alpha_S + \beta_S} \mathcal{T}^{\alpha_B + \beta_B} \\ &\leqq e^{-\delta M} \cdot \sum_{\substack{n \in \mathbb{Z}^N \\ |n| > M \\ l \in (\mathbb{Z}^+)^N \\ \alpha, \beta}} |\hat{R}(l, n; \alpha, \beta)| e^{|n|\sigma} \rho^{|l|} \tau^{\alpha_S + \beta_S} \mathcal{T}^{\alpha_B + \beta_B} \leqq e^{-\delta M} \| R \|_{D_i}. \end{aligned}$$

We remark that when we apply this lemma below we will choose the constant M in the definition of R^+ to equal M_k, the constant defined in Sect. 5.

The last ingredient we will need is a way of bounding quantities like $\| \{\chi, R^{(k)}\} \circ F^t \|$ in terms of $\| \{\chi, R^{(k)}\} \|$. The key ingredient in this estimate is Corollary 7.5. The basic requirement is that the difference between the transformed variables and the new variables should be small with respect to the norm in which we wish to measure $\| \{\chi, R\} \circ F^t \|$.

Lemma 8.3. *Let $(J(t), \varphi(t), z(y), \bar{z}(t)) = F^t(J, \varphi, z, \bar{z})$. Let $D_i = D(\tilde{I}, v, \rho/2, \sigma - \delta, \underline{\tau}/c; \mathbb{L}, \mathcal{T}/c')$, and $D_f = D(\tilde{I}, v, \rho, \sigma, \underline{\tau}; \mathbb{L}, \underline{\mathcal{T}})$. (We assume c and c' are greater than 1.) Suppose*

$$\| J_l(t) - J_l \|_{D_i} \leqq \rho/2; \quad \| z_l(t) - z_l \|_{D_i} \leqq (c-1)\tau_l/c \quad \text{if} \quad l \notin \mathbb{L},$$

$$\| \varphi_l(t) - \varphi_l \|_{D_i} \leqq \delta; \quad \| z_l(t) - z_l \|_{D_i} \leqq (c'-1)\mathcal{T}_l/c' \quad \text{if} \quad l \in \mathbb{L},$$

and

$$\| \bar{z}_l(t) - \bar{z}_l \|_{D_i} \leqq (c-1)\tau_l/c \quad \text{if} \quad l \notin \mathbb{L},$$

$$\| \bar{z}_l(t) - \bar{z}_l \|_{D_i} \leqq (c'-1)\mathcal{T}_l/c' \quad \text{if} \quad l \in \mathbb{L}.$$

Then if G is analytic on D_f,

$$\| G \circ F^t \|_{D_i} \leqq \| G \|_{D_f}.$$

Proof. We can simply the notation slightly by combining the variables $(I, z; \bar{z})$ into a single variable $\underline{x}$. We then wish to consider functions $f(\varphi, \underline{x})$, analytic on a domain $D_f = \{(\varphi, \underline{x}) \| \operatorname{Im} \varphi_i| < \sigma,\ i = 1, \ldots, N,\ |x_i| < r_i,\ i = 1, \ldots\}$. If we expand f in a Taylor–Fourier series as $f(\varphi, \underline{x}) = \sum_{n,m} \hat{f}(n,m) e^{in \cdot \varphi} \underline{x}^m$, then the norm of f becomes $\| f \| = \sum_{n,m} |\hat{f}(n,m)| e^{\sigma |n|} \underline{r}^m$, with $\underline{r}^m \equiv \prod_i r_i^{m_i}$. Suppose $D_i = \{(\varphi, \underline{x}) \| \operatorname{Im} \varphi_i| < \tilde{\sigma}$ and $|x_i| < \tilde{r}_i\}$, and let $(\tilde{\varphi}, \tilde{\underline{x}}) = (\varphi + \tilde{\Phi}(\varphi), \tilde{\underline{x}}(\varphi, x))$ be an analytic change of variables mapping $D_i \to D_f$ and satisfying $\| \tilde{x}_i(\varphi, \underline{x}) - x_i \|_{D_i} \leqq (r_i - \tilde{r}_i)$ and $\| \tilde{\Phi}(\varphi) \|_{D_i} < (\sigma - \tilde{\sigma})$ (These are just the translation of the hypotheses of the lemma into the $(\varphi, \underline{x})$ variables.) Let $g(\varphi, \underline{x}) = f(\varphi + \tilde{\Phi}(\varphi), \tilde{\underline{x}}(\varphi, \underline{x}))$. Since $f, g, \tilde{\Phi}$, and $\tilde{\underline{x}}$ are all analytic functions we can expand them all in Taylor–Fourier series as

$$g(\varphi, \underline{x}) = \sum_{n,m} \hat{g}(n,m) e^{in\varphi} \underline{x}^m, \quad f(\varphi, \underline{x}) = \sum_{n,m} \hat{f}(n,m) e^{in\varphi} \underline{x}^m,$$

$$\tilde{\Phi}(\varphi) = \sum_n \hat{\psi}(n) e^{in \cdot \varphi}, \quad \text{and} \quad \tilde{\underline{x}}(\varphi, \underline{x}) = \sum_{n,m} \hat{\chi}(n,m) e^{in\varphi} \underline{x}^m,$$

with the usual multi-index notation. (Note that $\hat{\psi}(\varphi)$ and $\hat{\chi}(n,m)$ are vectors.) Since $g = f \circ (\tilde{\varphi}, \tilde{x})$ we have

$$\begin{aligned} g(\varphi, \underline{x}) &= \sum_{m,n} \hat{f}(n,m) e^{in(\varphi + \sum_{n'} \hat{\psi}(n') e^{in'\varphi})} \left(\sum_{n'',m''} \hat{\chi}(n'', m'') e^{in''\varphi} \underline{x}^{m''} \right)^m \\ &= \sum_{m,n} \hat{f}(n,m) e^{in\varphi} \left(\sum_k \frac{1}{k!} \left(i \sum_{n'} n \cdot \hat{\psi}(n') e^{in' \cdot \varphi} \right)^k \right) \left(\sum_{n''m''} \hat{\chi}(n'', m'') e^{in''\varphi} \underline{x}^{m''} \right)^m. \end{aligned}$$

Thus, we can express $\hat{g}(n,m)$ as a sum of $\hat{f}(n,m)$, $\hat{\psi}(n)$ and $\hat{\chi}(n,m)$ which we express symbolically as $\hat{g}(n,m) = \sum^{(n,m)} \hat{f} \hat{\psi} \hat{\chi}$. (We will note need the explicit form of this sum.) Thus,

$$\begin{aligned} \| f \circ (\tilde{\varphi}, \tilde{\underline{x}}) \|_{D_i} &= \| g \|_{D_i} = \sum_{n,m} |\hat{g}(n,m)| e^{\tilde{\sigma}|n|} \tilde{r}^m = \sum_{n,m} |(\textstyle\sum^{(m,n)} \hat{f} \hat{\psi} \hat{\chi})| e^{\tilde{\sigma}|n|} \tilde{r}^m \\ &\leqq \sum_{n,m} (\textstyle\sum^{(m,n)} |\hat{f}| |\hat{\psi}| |\hat{\chi}|) e^{\tilde{\sigma}|n|} \tilde{r}^m \\ &\leqq \sum_{n,m} |\hat{f}(n,m)| \left(\sum_k \frac{1}{k!} \left(\sum_{n'} |n| |\hat{\psi}(n')| e^{\tilde{\sigma}|n'|} \right)^k \right) e^{\tilde{\sigma}|n|} \\ &\quad \cdot \left(\sum_{n''m''} |\hat{\chi}(n'', m'')| e^{\tilde{\sigma}|n''|} \tilde{r}^{m''} \right)^m. \end{aligned}$$

The last of these inequalities here just "undid" the expansions of $(\sum_k (1/k!)(\sum_{n'} n \cdot \hat{\psi}(n') e^{in' \cdot \varphi})^k)^n$ and $(\sum_{n'',m''} \hat{\chi}(n'', m'') e^{in''\varphi} \underline{x}^{m''})^m$ that we made to express $\hat{g}$ in terms of $\hat{f}, \hat{\psi}$ and $\hat{\chi}$. We note that

$$\sum_k \frac{1}{k!} \left(\sum_{n'} |n| |\hat{\psi}(n')| e^{\tilde{\sigma}|n'|} \right)^k = \exp\left(\sum_{n'} |n| |\hat{\psi}(n')| e^{\tilde{\sigma}|n'|} \right) \leqq \exp |n| (\| \tilde{\Phi}(\varphi) \|_{D_i}) \leqq \exp |n| (\sigma - \tilde{\sigma}).$$

Similarly,

$$\left(\sum_{n'',m''}|\hat{\chi}(n'',m'')|e^{\tilde{\sigma}|n''|}\underline{\tilde{r}}^{m''}\right)^m=\prod_l\|\underline{\tilde{x}}_l(\varphi,\underline{x})\|_{D_i}^{m_i}\leqq r^m.$$

Thus,

$$\|f\circ(\tilde{\varphi},\underline{\tilde{x}})\|_{D_i}\leqq\sum_{n,m}|\hat{f}(n,m)|e^{\tilde{\sigma}|n|}\times e^{(\sigma-\tilde{\sigma})|n|}r^m=\|f\|_{D_f},$$

which completes the proof of the lemma.

We now return to Eq. (8.1), and further examine the transformed hamiltonian. We begin by writing

$$\begin{aligned}\{\chi\circ H^{(k)}\}\circ F^t&=\{\chi,Q^{(k)}+S^{(k)}\}\circ F^t+\{\chi,R^{(k)}\}\circ F^t\\&=\{\chi,Q^{(k)}+S^{(k)}\}+\int_0^t\{\chi,\{\chi,Q^{(k)}+S^{(k)}\}\}\circ F^s ds+\{\chi,R^{(k)}\}\circ F^t.\end{aligned}$$

We now recall that χ was chosen so that

$$\begin{aligned}\{\chi,Q^{(k)}+S^{(k)}\}+[R^{(k)}]^-=&-\{\chi_0,[R^{(k)}]\}^--\{\chi_1,[R_1^{(k)}]\}^-\\&-\{\chi_0,S_3^{(k)}+S_4^{(k)}\}^+-\{\chi_1,S_3^{(k)}\}^+-\{\chi_2,[R_0^{(k)}]\}^-\\&-\{\chi_1,[R_2^{(k)}]\}^--\{\chi_2,[R_1^{(k)}]\}^-+(Q^{(k+1)}-Q^{(k)})\\&-\{\chi_1,S^{\geqq}\}^--\{\chi_0,S^{\geqq}\}^--\{\chi_1,S_4\}^-.\end{aligned}$$

All of these terms with the exception of $(Q^{(k+1)}-Q^{(k)})$, can be bounded with the aid of Lemma 8.1, so we now look at this quadratic term in more detail. The terms contributing to $Q^{(k+1)}-Q^{(k)}$ arise in Eq. (6.5), from which we obtain an explicit formula for them:

$$\begin{aligned}&Q^{(k+1)}(I,J;z,\bar{z})-Q^{(k)}(I,J;z,\bar{z})\\&=\sum_{l=1}^N[\hat{R}^{(k)}(I,\delta_l,0;0,0)+\hat{v}^2(I,\delta_l,0;0,0)+\hat{v}^3(I,\delta_l,0;0,0)]J^l\\&\quad+\sum_{l=N+1}^\infty[\hat{R}^{(k)}(I,0,0;\delta_l,\delta_l)+\hat{v}^2(I,0,0;\delta_l,\delta_l)+\hat{v}^3(I,0,0;\delta_l,\delta_l)]z_l\bar{z}_l.\end{aligned}\tag{8.5}$$

The quantities in this expression are precisely those that appeared in the numerators of χ_2^J and χ_2^z. We bound them in the same way we did in Lemmas 6.3 and 6.4 and we find

Lemma 8.4. *If $\tilde{I}\in\mathscr{Q}^{(k)}$ and $\hat{\mathbb{L}}$ is as in the paragraph preceding 6.2 then*

$$\|Q^{(k+1)}-Q^{(k)}\|\leqq E_2^J(\varepsilon_k)+\hat{E}_2^z(\varepsilon_k),$$

on $D(\tilde{I},v^{(k)},\rho^{(k)}/4,\sigma^{(k)}-4\delta^{(k)},\underline{\tau}^{(k)}/(1+2^{-5}(k+1)^{-2})^2,\hat{\mathbb{L}},\tilde{\mathscr{T}}^{(k)}/(1+(k+1)^{-2})^2)$.

Note that in addition we can read off from (8.5) expressions for the quantities $f_j^{(k)}(I)$ and $g_j^{(k)}(I)$ defined in Proposition 5.1. One has, for $k\geqq 0$,

$$\begin{aligned}&|f_l^{(k+1)}(I)-f_l^{(k)}(I)|\\&=|[\hat{R}^{(k)}(I,\delta_l,0;0,0)+\hat{v}^2(I,\delta_l,0;0,0)+\hat{v}^3(I,\delta_l,0;0,0)]|\\&\leqq 2\{\varepsilon_k+2NC_k\varepsilon_0E_0(\varepsilon_k)/\rho^{(k)}\delta^{(k)}+2^7C_k\varepsilon_0E_1(\varepsilon_k)[(N/\rho^{(k)}\delta^{(k)}+1)+1]\}/\rho^{(k)},\end{aligned}$$

and

$$
\begin{aligned}
&|g_l^{(k+1)}(I) - g_l^{(k)}(I)| \\
&= |[\hat{R}^{(k)}(I,0,0;\delta_l,\delta_l) + \hat{v}^2(I,0,0;\delta_l,\delta_l) + \hat{v}^3(I,0,0;\delta_l,\delta_l)| \\
&\leqq 2^6(1+k)^2(1+N_{k-M})^2\{\varepsilon_k + 2NC_k\varepsilon_0 E_0(\varepsilon_k)/\rho^{(k)}\delta^{(k)} \\
&\quad + 2^7 C_K\varepsilon_0 E_1(\varepsilon_k)[(N/\rho^{(k)}\delta^{(k)}) + 1]\}/(\mathcal{T}_l^{(k)})^2.
\end{aligned}
$$

We have bounded these expressions by once again noting that they appeared in the numerators of χ_2^J and χ_2^z and using the estimates derived in Lemmas 6.3 and 6.4. Using the definitions of the inductive constants from Sect. 5 it is easy to verify that if ε_0 is sufficiently small (this smallness condition does not depend on k), the quantities on the right-hand side of these inequalities are bounded by $\varepsilon_k \cdot \varepsilon^{-1/3}$, and $\varepsilon_k \cdot \varepsilon^{-1/3}/l^{2\zeta_k}$ respectively, when $k \geqq 1$. Next note that the quantities $\tilde{f}_j^{(1)}(I)$ and $\tilde{g}_j^{(1)}(I)$ defined in Eqs. (4.10) and (4.12) are $\hat{R}^{(0)}(I,\delta_j,0;0,0)$ and $\hat{R}^{(0)}(I,0,0;\delta_j,\delta_j)$ respectively. Thus,

$$
|f_j^{(1)}(I) - \tilde{f}_j^{(1)}(I)| \leqq |\hat{v}^2(I,\delta_j,0;0,0)| + |\hat{v}^3(I,\delta_j,0;0,0)| \leqq \varepsilon^{7/6},
$$

using the estimates derived in Sect. 6 for v^2 and v^3. One bounds $|g_j^{(1)}(I) - \tilde{g}_j^{(1)}(I)|$ in a similar fashion, and this completes the verification of induction hypothesis $(k+1.1)$.

In order to check $(k+1.2)$ and $(k+1.3)$ we give explicit expressions for $R^{(k+1)}$ and $S^{(k+1)}$. Given an analytic function G, we find it convenient to define G^R to be the sum of those terms in the power series for G of order two or less, and $G^S = G - G^R$. Note that one has immediately that $\|G^R\| \leqq \|G\|$ and $\|G^S\| \leqq \|G\|$. Then we have

$$
\begin{aligned}
S^{(k+1)} = S^{(k)} &+ \left[\int_0^1 \{\chi, R^{(k)}\} \circ F^t dt\right]^S + \left[\int_0^1 \left[\int_0^t \{\chi, \{\chi, Q^{(k)} + S^{(k)}\}\} \circ F^t ds\right] dt\right]^S \\
&- \{\chi_1, S^{\geqq}\}^- - \{\chi_0, S^{\geqq}\}^- - \{\chi_1, S^4\}^-. \qquad (8.6)
\end{aligned}
$$

While

$$
\begin{aligned}
R^{(k+1)} = [R^{(k)}]^+ &+ \left[\int_0^1 \{\chi, R^{(k)}\} \circ F^t dt\right]^R + \{\chi_0, S_3^{(k)} + S_4^{(k)}\}^+ + \{\chi_1, S_3^{(k)}\}^+ \\
&+ \left[\int_0^1 \left[\int_0^t \{\chi, \{\chi, Q^{(k)} + S^{(k)}\}\} \circ F^t ds\right]\right]^R - \{\chi_0, [R_2^{(k)}]\}^- \\
&+ \{\chi_1, [R_1^{(k)}]\}^- - \{\chi_2, [R_0^{(k)}]\}^- - [\chi_1, [R_2^{(k)}]\}^- - \{\chi_2, [R_1^{(k)}]\}^-. \qquad (8.7)
\end{aligned}
$$

Lemmas 8.1–8.4 now make it easy to verify the induction hypotheses. Let $D_i = D(\tilde{I}, v^{(k)}, \rho^{(k+1)}, \sigma^{(k+1)}, \underline{\tau}^{(k+1)}; \mathbb{L}^{(k+1)}, \underline{\mathcal{T}}^{k+1)})$ and $D_f = D(\tilde{I}, v^{(k)}, \rho^{(k)}, \sigma^{(k)}, \underline{\tau}^{(k)}, \mathbb{L}^{(k+1)}, \underline{\mathcal{T}}^{(k)})$.

Combining Lemmas 8.1 and 8.3, plus the induction hypothesis for $R^{(k)}$ we have

$$
\left\|\left[\int_0^1 \{\chi, R^{(k)}\} \circ F^t dt\right]\right\|_{D_i} \leqq \sup_{t \in [0,1]} \|\{\chi, R^{(k)}\} \circ F^t\|_{D_i} \leqq K(\varepsilon_k)\|R^{(k)}\|_{D_f} \leqq K(\varepsilon_k) \cdot \varepsilon_k. \qquad (8.8)
$$

Note further that

$$
\|\{\chi_0, [R_2^{(k)}]\}^- + \{\chi_1, [R_1^{(k)}]\}^- + \{\chi_2, [R_2^{(k)}]\}^- + \{\chi_2, [R_1^{(k)}]\}^-\| \leqq \|\{\chi, R^{(k)}\}\| \leqq K(\varepsilon_k)\varepsilon_k,
$$

and

$$\| \{\chi_0, S^{\geqq}\}^- + \{\chi_1, S^{\geqq}\}^- + \{\chi_1 S^4\}^- \| \leqq \| \{\chi, S\} \| \leqq K(\varepsilon_k) \cdot C_k \varepsilon_0,$$

using the induction hypothesis.

These two remarks, when combined with Lemma 8.4 imply

$$\| \{\chi, Q^{(k)} + S^{(k)}\} \| \leqq \varepsilon_k + (E_2^J(\varepsilon_k) + \hat{E}_2^z(\varepsilon_k)) + K(\varepsilon_k) \cdot \varepsilon_k + C_k \varepsilon_0 K(\varepsilon_k),$$

so that

$$\left\| \int_0^1 \int_0^t \{\chi, \{\chi, Q^{(k)} + S^{(k)}\}\} \circ F^s ds \right\|_{D_i}$$
$$\leqq [K(\varepsilon_k)][\varepsilon_k + (E_2^J(\varepsilon_k) + \hat{E}_2^z(\varepsilon_k)) + K(\varepsilon_k)\varepsilon_k + C_k \varepsilon_0 K(\varepsilon_k)].$$

Finally, we note that Lemma 8.2 implies

$$\| [R^{(k)}]^+ \|_{D_i} \leqq e^{-6\delta^{(k)} M_k} \quad \text{and} \quad \| [\{\chi, S\}^+] \|_{D_i} \leqq C_k \varepsilon_0 e^{-\delta^{(k)} M_k} K(\varepsilon_k).$$

Combining these remarks shows that

$$\| S^{(k+1)} \|_{D_i} \leqq C_k \varepsilon_0 + K(\varepsilon_k) \cdot \varepsilon_k + C_k \varepsilon_0 K(\varepsilon_k)$$
$$+ [K(\varepsilon_k)][\varepsilon_k + (E_2^J(\varepsilon_k) + \hat{E}_2^z(\varepsilon_k)) + K(\varepsilon_k)\varepsilon_k + C_k \varepsilon_0 K(\varepsilon_k)],$$

and

$$\| R^{(k+1)} \|_{D_i} \leqq e^{-6\delta^{(k)} M_k} \varepsilon_k + 2K(\varepsilon_k) \cdot \varepsilon_k + C_k \varepsilon_0 e^{-\delta^{(k)} M_k} K(\varepsilon_k)$$
$$+ [K(\varepsilon_k)][\varepsilon_k + (E_2^J(\varepsilon_k) + \hat{E}_2^z(\varepsilon_k)) + K(\varepsilon_k)\varepsilon_k + C_k \varepsilon_0 K(\varepsilon_k)].$$

Inserting the definitions of the various inductive constants it is elementary, if somewhat tedious, to verify that the induction hypotheses $(k+1.2)$ and $(k+1.3)$ of Proposition 5.1 are satisfied.

IX. Estimates on Small Denominators

In this section we derive the estimates necessary to bound the denominators of the generating function in Sect. 5. This procedure takes two steps. We first show that for a large set of potentials, v, the operator L_v has eigenvalues that satisfy "good" small denominator conditions. We then show that if one starts with a potential with "good" frequencies then at each stage in the iterative process we can adjust the perturbed frequencies by moving the vector I in such a way that we maintain control over the small denominators.

Recall that $E_0 = \left\{ v \in L^2[0,1] \,|\, v(x) = v(1-x), \int_0^1 v = 0. \right\}$. Given $v \in E_0$ let $\mu_1 > \mu_2 > \cdots$ be the spectrum of $L_v = (d^2/dx^2) - v$. We assume that $\mu_1 < 0$. Recall that μ_n has the asymptotic form

$$-\mu_n = (n\pi)^2 + l^2(n), \tag{9.1}$$

and that furthermore any decreasing sequence of the form (9.1) is the spectrum of some $v \in E_0$. (Here, $l^2(n)$ is the n^{th} component of an l^2-sequence.) It is $\omega_n = \sqrt{|\mu_n|}$, rather than μ_n itself which enters our estimates so it is useful to have asymptotic formulae for the ω_n's. Define

$$l_k^2 = \left\{ (x_1, x_2, \ldots) | \sum_{n \geqq 1} (n^k x_n)^2 < \infty \right\}. \tag{9.2}$$

Proposition 9.1. *An increasing sequence of positive numbers* $\{\omega_n\}$ *gives rise to a sequence* $-\mu_n = \omega_n^2$ *of the form (9.1) if and only if*

$$\omega_n = (n\pi)(1 + x_n), \tag{9.3}$$

with $\{x_n\} \in l_2^2$.

Proof. First suppose that $\omega_n = \sqrt{|\mu_n|}$, and μ_n satisfies (9.1). Then $x_n = \sqrt{1 + l^2(n)/(n\pi)^2} - 1$, and, $|x_n| < a_n/(n\pi)^2$, for some l^2 sequence $\{a_n\}$. Thus,

$$\sum_{n \geqq 1} (n^2 x_n)^2 \leqq \sum_{n \geqq 1} \frac{a_n^2}{\pi^4} < \infty.$$

Conversely, suppose that $\{x_n\} \in l_2^2$ and set $\omega_n = (n\pi)(1 + x_n)$ and $-\mu_n = (n\pi)^2(1 + x_n)^2 = (n\pi)^2 + 2(n\pi)^2 x_n + (n\pi)^2 x_n^2$. We need only show $\{2(n\pi)^2 x_n + (n\pi)^2 x_n^2\} \in l^2$. However, $\{(n\pi)^2 x_n\} \in l^2$ since $\{x_n\} \in l_2^2$, and $\{(n\pi)^2 x_n^2\} \in l^2$ as we see by applying the Cauchy–Schwartz inequality. This completes the proof of the proposition.

The space l_2^2 is a convenient one with which to work, because it has associated with it a variety of Gaussian probability measures. For instance, we can define a probability measure on sequences $x = (x_1, x_2, x_3, \ldots)$, by taking a product of the probability measures with gaussian densities

$$dP_n^\alpha(x_n) = \frac{n^\alpha}{\sqrt{\pi}} e^{-n^{2\alpha} x_n^2} dx_n$$

on each component x_n. If we denote the product measure by dP^α, then l_2^2 has full measure with respect to dP^α if $\alpha > 3$. (In fact, l_2^2 has full measure if $\alpha > 3/2$, but we want $\alpha > 3$ for later purposes.)

Now set $\omega_n = (n\pi)(1 + x_n)$, and define $\Omega = (\omega_1, \ldots, \omega_N)$. We will verify that for all $\{x_n\}$ except for a set of measure zero with respect to dP^α, these frequencies satisfy the non-resonance conditions (D.1)–(D.3) of Sect. 3 for some choice of the constants $D_0^{(1)}, D_0^{(2)}$, and τ.

Fix $n \in \mathbb{Z}^N$, $n \neq 0$ and $j \geqq 0$. Then

$$\operatorname{Prob}(|n \cdot \Omega \pm j\pi| < [D_0^{(1)}]^{-1}(|n| + j)^{-\tau} = \int_S \left(\prod_{l=1}^N \frac{l^\alpha}{\sqrt{\pi}} e^{-l^{2\alpha} x_l^2} dx_l \right),$$

where

$$S = \left\{ x \left| \left| \sum_{l=1}^N n_l (l\pi)(1 + x_l) \pm j\pi \right| < [D_0^{(1)}]^{-1}(|n| + j)^{-\tau} \right. \right\}.$$

It is convenient to change variables in this integral to $z_l = l^\alpha x_l$ so that

$$\operatorname{Prob}(|n \cdot \Omega \pm j\pi| < [D_0^{(1)}]^{-1}|n|^{-\tau}) = \int_{\tilde{S}} \left(\prod_{l=1}^N \frac{e^{-z_l^2}}{\sqrt{\pi}} dz_l \right)$$

with

$$\tilde{S}=\left\{z\left|\left|\sum_{l=1}^{N} n_l(l\pi)+\sum_{l=1}^{N}\left(\frac{n_l\pi}{l^{\alpha-1}}\right)z_l\pm j\pi\right|<[D_0^{(1)}]^{-1}(|n|+j)^{-\tau}\right.\right\}.$$

Note that $\tilde{S}$ is a strip of width $2[D_0^{(1)}]^{-1}(|n|+j)^{-\tau}$ about the hyperplane $c(n)+\tilde{n}\cdot z=\mp j\pi$, where $c(n)=\sum_{l=1}^{N} n_l(l\pi)$ and $\tilde{n}=\left(\frac{n_1\pi}{1},\frac{n_2\pi}{2^{\alpha-1}},\frac{n_3\pi}{3^{\alpha-1}},\ldots,\frac{n_N\pi}{N^{\alpha-1}}\right)$. The volume of such a strip is easy to estimate, since we now have a Gaussian integral and we find

$$\operatorname{Prob}(|n\cdot\Omega\pm j\pi|<[D_0^{(1)}]^{-1}(|n|+j)^{-\tau})\leqq 2[D_0^{(1)}]^{-1}(|n|+j)^{-\tau}.$$

This gives the probability that the non-resonance condition (D.1) fails for some particular $n\in\mathbb{Z}^N$ and $j\geqq 0$. We estimate the total probability that it fails by summing over n and j. Provided $\tau>N+2$ we have.

$$\sum_{\substack{n\in\mathbb{Z}^N\\ n\neq 0\\ j\geqq 0}} 2[D_0^{(1)}]^{-1}(|n|+j)^{-\tau}\leqq 2[D_0^{(1)}]^{-1}3^{N+1}\left(1+\frac{1}{\tau-N-1}\right).$$

Since we can make this arbitrarily small by choosing $D_0^{(1)}$ large, we see that for almost every Ω, there exists some $D_0^{(1)}$ such that estimate (D.1) is satisfied for all $n\in\mathbb{Z}^N$, $n\neq 0$ and $j\geqq 0$.

Estimate (D.2) is proved in the same fashion so we do not reproduce the details here. Estimate (D.3) is slightly more difficult.

We will consider in detail the estimate of

$$\operatorname{Prob}(|\underline{n}\cdot\underline{\Omega}+(\omega_j-\omega_l)|<[D_0^{(2)}]^{-1}[|n|+|j-l|]^{-4\tau}).$$

The other three choices of plus and minus signs in condition (D.3) follow in analogous fashion. It is easy to estimate for fixed n,j and l,

$$\operatorname{Prob}(|\underline{n}\cdot\underline{\Omega}+(\omega_j-\omega_l)|<[D_0^{(2)}]^{-1}[|n|+|j-l|]^{-4\tau})\leqq 2[D_0^{(2)}]^{-1}[|n|+|j-l|]^{-4\tau}, \tag{9.4}$$

using the same methods as above. The problem is that if we now sum over $\underline{n}$, j and l, the sum diverges.

We avoid this problem by rewriting

$$\begin{aligned}|n\cdot\Omega-(\omega_j-\omega_l)|&=|n\cdot\Omega-\pi(j-l)+\pi(j-l)-(\omega_j-\omega_l)|\\&\geqq|n\cdot\Omega-\pi(j-l)|-|\pi(j-l)-(\omega_j-\omega_l)|,\end{aligned}$$

and estimating these two pieces separately. We begin by noting that

$$\begin{aligned}\operatorname{Prob}(|\pi j-\omega_j|\geqq\delta)&=\int_{|\pi j-x|\geqq\delta} j^\alpha e^{-j^{2\alpha}x^2}\frac{dx}{\sqrt{\pi}}\\&\leqq\int_{|z|\geqq\delta j^{\alpha-1}/\pi} e^{-z^2}\frac{dz}{\sqrt{\pi}}\leqq 2\exp\left(-\frac{1}{2}\left(\frac{\delta j^{\alpha-1}}{\pi}\right)^2\right).\end{aligned} \tag{9.5}$$

We now estimate the set of frequencies for which (D.3) fails as follows:

First note that (D.1) implies that there exists a constant $\gamma(D_0^{(1)})$, with $\gamma(D_0^{(1)})\searrow 0$

as $D_0^{(1)} \nearrow \infty$ such that

$$\mathrm{Prob}(|\tilde{n}\cdot\Omega - \tilde{m}\pi| < [D_0^{(1)}]^{-1}[|\tilde{n}| + \tilde{m}]^{-\tau} \text{ for some } \tilde{n}\in\mathbb{Z}^N, \tilde{m} > 0) < \gamma(D_0^{(1)}).$$

For simplicity, denote this event by $\overline{\mathrm{D.1}}$, and its complement by D.1. Then

$$\sum_{\substack{n\in\mathbb{Z}^N\\ j\geqq N\\ m>0}} \mathrm{Prob}(|n\cdot\Omega - (\omega_{j+m} - \omega_j)| < [D_0^{(2)}]^{-1}[|n| + m]^{-4\tau}) \leqq \gamma(D_0^{(1)})$$

$$+ \sum_{\substack{n\in\mathbb{Z}^N\\ m>0}} \left\{ \sum_{N\leqq j\leqq [D_0^{(1)}][|n|+m]^\tau} \mathrm{Prob}(|n\cdot\Omega - (\omega_{j+m} - \omega_j)| < [D_0^{(2)}]^{-1}[|n|+m]^{-4\tau} \textit{ and } \mathrm{D.1}) \right\}$$

$$+ \sum_{\substack{n\in\mathbb{Z}^N\\ m>0}} \left\{ \sum_{j>[D_0^{(1)}][|n|+m]^\tau} \mathrm{Prob}(|n\cdot\Omega - (\omega_{j+m} - \omega_j)| < [D_0^{(2)}]^{-1}[|n|+m]^{-4\tau} \textit{ and } \mathrm{D.1}) \right\}.$$

We now estimate separately the two sums over j. To estimate the first one, note that (9.4) implies it is less than or equal to

$$\sum_{\substack{n\in\mathbb{Z}^N\\ m>0}} \{2[D_0^{(1)}][D_0^{(2)}]^{-1}[|n| + m]^{-3\tau}\} \leqq C(N,\tau)\{[D_0^{(1)}]/[D_0^{(2)}]\},$$

provided $\tau > N + 2$.

To estimate the second sum over j we note that

$$|n\cdot\Omega - (\omega_{j+m} - \omega_j)| \geqq |n\cdot\Omega - \pi m| - (|\pi j - \omega_j| + |\pi(j+m) - \omega_{j+m}|),$$

so that, using the fact that (D.1) holds

$$|n\cdot\Omega - (\omega_{j+m} - \omega_j)| \geqq [D_0^{(2)}]^{-1}[|n| + m]^{-4\tau},$$

provided $D_0^{(2)} > 2D_0^{(1)}$ and $(|\pi j - \omega_j| + |\pi(j+m) - \omega_{j+m}|) < \frac{1}{2}[D_0^{(1)}]^{-1}[|n| + m]^{-\tau}$. On the other hand, (9.5) implies

$$\mathrm{Prob}((|\pi j - \omega_j| + |\pi(j+m) - \omega_{j+m}|) > \tfrac{1}{2}[D_0^{(1)}]^{-1}[|n| + m]^{-\tau})$$

$$\leqq 4\exp\left(-\frac{1}{2\pi^2}[D_0^{(1)}]^{-2}(|n| + m)^{-2\tau} j^{2(\alpha-1)}\right).$$

Thus,

$$\sum_{\substack{n\in\mathbb{Z}^N\\ m>0}} \left\{ \sum_{j>[D_0^{(1)}][|n|+m]^\tau} \mathrm{Prob}(|n\cdot\Omega - (\omega_{j+m} - \omega_j)| < [D_0^{(2)}]^{-1}[|n| + m]^{-4\tau} \textit{ and } \mathrm{D.1}) \right\}$$

$$\leqq \sum_{\substack{n\in\mathbb{Z}^N\\ m>0}} \sum_{j>[D_0^{(1)}][|n|+m]^\tau} 4\exp\left(-\frac{1}{2\pi^2}[D_0^{(1)}]^{-2}(|n| + m)^{-2\tau} j^{2(\alpha-1)}\right).$$

$$\leqq \sum_{\substack{n\in\mathbb{Z}^N\\ m>0}} \left| 8[D_0^{(1)}]^2[|n| + m]^{2\tau} \exp\left\{-\left(\frac{1}{2\pi^2}[(D_0^{(1)})(|n| + m)^\tau]^{2(\alpha-2)}\right)\right\}\right|$$

$$\leqq C(N,\tau)[D_0^{(1)}]^2 \exp\left\{-\frac{1}{2\pi^2}[D_0^{(1)}]^{2(\alpha-3)}\right\}.$$

(Recall that $\alpha > 3$, so all the sums converge.)

Thus, the set of points which violate (D.3) has measure bounded by

$$C(N,\tau)\left[[D_0^{(1)}]/[D_0^{(2)}]+[D_0^{(1)}]^2\exp\left\{-\frac{1}{2\pi^2}[D_0^{(1)}]^{2(\alpha-3)}\right\}\right]+\gamma(D_0).$$

Thus, if we choose $D_0^{(2)}=[D_0^{(1)}]^2$, the probability that (D.3) is violated goes to zero as $D_0^{(1)}\to\infty$, so (D.1)–(D.3) are almost surely satisfied.

Note that this defines a measure on the set of potentials in E_0 whose spectrum is purely negative if we set the measure of a subset of E_0 to be equal to the measure of the set of frequencies of those potentials with respect to dP^α. Thus, this argument proves that the sets $\mathscr{E}_0(j)$ and $\mathscr{F}_0(N)$ in Theorems 2.1 and 2.2 do indeed have full measure if we restrict ourselves to potentials whose eigenvalues are all negative.

We now complete the argument by showing that if we start with a potential whose frequencies satisfy (D.1)–(D.3) we can maintain control of these frequencies throughout the iterative process. Here, we follow the ideas of [E] closely.

We construct the set $\mathscr{Q}^{(k)}$ in Proposition 5.1 by successively eliminating points I from $B(v^0,I^0)$ which violate the small denominator conditions at each stage of the iteration. At the initial stage of the iteration the frequencies are independent of I, so we take $\mathscr{Q}^{(0)}=B(v^0,I^0)$.

At subsequent stages of the iteration the frequencies will depend on I, but we are able to prove:

Lemma 9.2. *Suppose we have constructed a set* $\mathscr{Q}^{(k)}\subset B(v^0,I^0)$, *such that for* $I\in\mathbb{C}^N$ *and* $\operatorname{dist}(I,\mathscr{Q}^{(k)})<v^{(k)}$ *one has*

(Dk.1) $|n\cdot\Omega^{(k)}(I)\pm j\pi|\geqq[D_k^{(1)}]^{-1}(|n|+j)^{-\tau},\quad n\in\mathbb{Z}^N,\quad 0<|n|\leqq M_k,\quad j\geqq 0.$

(Dk.2) $|n\cdot\Omega^{(k)}(I)\pm\omega_j^{(k)}(I)|\geqq[D_k^{(1)}]^{-1}[|n|+j]^{-\tau},\quad n\in\mathbb{Z}^N,\quad |n|\leqq M_k,\quad j\geqq N+1,$

(Dk.3) $|n\cdot\Omega^{(k)}(I)\pm(\omega_j^{(k)}(I)\pm\omega_l^{(k)}(I))|\geqq[D_k^{(2)}]^{-1}[|n|+|j-l|]^{-4\tau}$
$n\in\mathbb{Z}^N,\quad |n|\leqq M_k,\quad j,l\geqq N+1.$

Then if ε_0 *is sufficiently small, there exists* $\mathscr{Q}^{(k+1)}\subset\mathscr{Q}^{(k)}$ *with* $\operatorname{meas}(\mathscr{Q}^{(k)}\backslash\mathscr{Q}^{(k+1)})\leqq\mathscr{O}(1/k^2|\log\varepsilon_0|)\operatorname{meas}(\mathscr{Q}^{(0)})$ *such that if* $I\in\mathbb{C}^N$ *and* $\operatorname{dist}(I,\mathscr{Q}^{(k+1)})<v^{(k+1)}$, *then* $\Omega^{(k+1)}$, *and* $\{\omega_j^{(k+1)}\}$ satisfy (D(k + 1).1) – (D(k + 1).3).

Proof. Note that (D0.1)–(D0.4) hold for $I\in\mathscr{Q}^0$. (In fact, since $\Omega^{(0)}$ and $\omega_j^{(0)}$ are independent of I, they hold for all I.) Furthermore, if ε_0 is sufficiently small, (D1.1)–(D1.3) and (D2.1)–(D2.3) are automatically satisfied for $I\in\mathscr{Q}^{(1)}=\mathscr{Q}^{(2)}=B(v^0,I^0)$. In order for (D1.1) to fail, we must have $|j|\leqq\mathscr{O}(|n|)\approx\mathscr{O}(|M_1|)$. Thus,

$$|n\cdot\Omega^{(1)}(I)\pm j\pi|\geqq|n\cdot\Omega^{(0)}(I)\pm j\pi|-|n\cdot\Omega^{(1)}(I)-n\cdot\Omega^{(0)}|$$
$$\geqq[D_0^{(0)}](|n|+j)^{-\tau}-|n|\varepsilon^{7/6}\geqq[D_0^{(1)}](|n|+j)^{-\tau},$$

if $|n|$ and j are $\leqq\mathscr{O}(M_1)$ and ε_0 is sufficiently small. Here, the next to last inequality used (D0.1) and the estimate of (k.1) of Proposition 5.1. The proofs that (D1.2), (D1.3) and (D2.1)–(D2.3) are satisfied are similar so we omit them. In fact, using this idea, one could show that (Dk.1)–(Dk.3) are automatically satisfied for any finite k, by choosing ε_0 sufficiently small. Choosing $k=2$ is sufficient to make the induction argument work, but, we make no claim that is the optimal choice.

We now show inductively that we can construct sets, $\mathscr{Q}^{(k+1)}$, for $k \geqq 24$, on which (Dk + 1.1)–(Dk + 1.3) hold, and then finally show that these estimates hold on complex neighborhoods of $\mathscr{Q}^{(k+1)}$ of size $\nu^{(k+1)}$.

We will first estimate how many frequencies must be excluded in order to insure that for $I \in \mathscr{Q}^{(k+1)}$,

$$|n \cdot \Omega^{(k+1)}(I) - \pi j| \geqq 2[D_{k+1}^{(1)}]^{-1}[|n| + j]^{-\tau}.$$

This estimate is rather standard. From the estimates on $\partial\Omega/\partial I$ at the end of Sect. 4, we see that the map $I \to \Omega^{(k+1)}(I)$ is a diffeomorphism from $\mathscr{Q}^{(k)} \to \Omega^{(k+1)}(\mathscr{Q}^{(k)})$. In fact, it is easy to see that one can, by interpolation, extend $\Omega^{(k+1)}(I)$ to all of $B(\nu^{(0)}, I^0)$ in such a way that the extended function is unchanged on $\mathscr{Q}^{(k)}$, but satisfies $\det(\partial\Omega^{(k+1)}/\partial I) = C\varepsilon^N(1 + \mathcal{O}(\varepsilon^{1/2}))$ on all of $B(\nu^{(0)}, I^0)$. (And hence is a diffeomorphism on the whole ball.) Since $\Omega^{(k+1)}$ is a diffeomorphism we may treat the frequencies Ω as independent variables, and then at the end of the computation, pull back our estimates on the set of frequencies that must be excluded via $(\Omega^{(k+1)})^{-1}$ to give estimates on the size of the sets $\mathscr{Q}^{(k+1)}$.

The estimates on $\Omega^{(k+1)}(I)$ imply that there are positive constants $\underline{c}$ and $\bar{c}$, independent of k, such that

$$B(\underline{c}\varepsilon\nu^{(0)}, \Omega^{(k+1)}(I^0)) \subset \Omega^{(k+1)}(\mathscr{Q}^{(k)}) \subset B(\bar{c}\varepsilon\nu^{(0)}, \Omega^{(k+1)}(I^0)).$$

Thus we have,

$$\operatorname{meas}(\{\Omega \mid |n \cdot \Omega - \pi j| < [D_{k+1}^{(1)}]^{-1}[|n| + j]^{-\tau}\}) = \int\limits_{\substack{\Omega^{(k+1)}(\mathscr{Q}^{(k)}) \cap \\ \{\Omega \mid |n \cdot \Omega - \pi j| < [D_{k+1}^{(1)}]^{-1}[|n|+j]^{-\tau}\}}} d\Omega \leqq 4[D_{k+1}^{(1)}]^{-1}[|n| + j]^{-\tau} \cdot c(N)(\bar{c}\nu^{(0)}\varepsilon)^{N-1}.$$

This last inequality results from the fact that the set over which we are integrating is contained in a strip of width $4[D_{k+1}^{(1)}]^{-1}[|n| + j]^{-\tau}$ about a hyperplane through $B(\bar{c}\varepsilon\nu^{(0)}, \Omega^{(k+1)}(I^0))$.

The estimates on $\partial\Omega^{(k+1)}/\partial I$ imply $(\nu^{(0)}\varepsilon)^N \leqq c(N) \operatorname{meas}(\Omega^{(k+1)}(B(\nu^{(0)}, I^0)))$, so

$$\begin{aligned} &\operatorname{meas}(\{\Omega \mid |n \cdot \Omega - \pi j| < 2[D_{k+1}^{(1)}][|n| + j]^{-\tau}\}) \\ &\quad \leqq c(N)[D_{k+1}^{(1)}]^{-1}(\nu^{(0)}\varepsilon)^{-1} \operatorname{meas}(\Omega^{(k+1)}(B(\nu^{(0)}, I^0)))[|n| + j]^{-\tau}. \end{aligned}$$

Since $[D_{k+1}^{(1)}]^{-1}(\nu^{(0)}\varepsilon)^{-1} \ll 1$, for $k \geqq 2$, the measure of the set of excluded points is very small. If we now sum over n and j, we see that the measure of the set of points where (D(k + 1).1) fails is bounded by

$$C(N, \tau)[D_{k+1}^{(1)}]^{-1}(\nu^{(0)}\varepsilon)^{-1} \operatorname{meas}(\Omega^{(k+1)}(B(\nu^{(0)}, I^{(p)}))),$$

provided $\tau > N + 2$.

We now consider the set of points where hypothesis (D(k + 1).2) fails. Note first of all that (D(k + 1).2) is automatically satisfied for $k > 0$ unless $|n| + j$ is rather large. To see this note that

$$|n \cdot \Omega^{(k+1)} - \omega_j^{(k+1)}| = |n \cdot \Omega^{(0)} - \omega_j^{(0)} + n \cdot (\Omega^{(k+1)} - \Omega^{(0)}) + (\omega_j^{(k+1)} - \omega_j^{(0)})|.$$

It is easy to estimate $(\Omega^{(k+1)} - \Omega^{(0)})$ and $\omega_j^{(k+1)} - \omega_j^{(0)}$ using the bounds of induction hypothesis $(k+1.1)$ of Proposition 5.1, and we find $|\Omega^{(k+1)}-\Omega^{(0)}| \leqq c\varepsilon^{5/4}$, while $|\omega_j^{(k+1)}-\omega_j^{(0)}| \leqq c\varepsilon^{5/4}/j^{2\zeta_k}$. Since $\Omega^{(0)}$, and $\omega_j^{(0)}$ satisfy (D0.2) we have

$$|n\cdot\Omega^{(k+1)}-\omega_j^{(k+1)}| \geqq D_0^{(1)}[|n|+j]^{-\tau} - c\varepsilon^{5/4}(|n|+1),$$

which is automatically bounded below by $2D_{k+1}^{(1)}[|n|+j]^{-\tau}$ unless $|n|+j$ is $\mathcal{O}(\varepsilon^{-5/4(\tau+1)})$. Once again we treat Ω as the independent variable. Thus, we must estimate

$$\operatorname{meas}(\{\Omega|\,|n\cdot\Omega-\omega_j|<2[D_{k+1}^{(1)}]^{-1}[|n|+j]^{-\tau}\}) = \int\limits_{\substack{(\Omega^{(k+1)}(\mathcal{D}^{(k)}))\cap \\ \{\Omega|\,|n\cdot\Omega-\omega_j|<2[D_{k+1}^{(1)}]^{-1}[|n|+j]^{-\tau}\})}} d\Omega. \tag{9.6}$$

If ω_j were independent of Ω, this would be identical to the previous estimate. In that case, if $|n\cdot\Omega-\pi j|<\delta$, and we moved Ω by an amount δ in the direction of $\underline{n}$, the inequality would reverse. Here, if $|n\cdot\Omega-\omega_j^{(k+1)}(\Omega)|<\delta$, and we move Ω in the direction of $n, \omega_j(\Omega)$ may also move in such a way that the inequality still holds. However, if we differentiate this expression with respect to Ω, we have

$$\frac{\partial}{\partial\Omega_m}(n\cdot\Omega-\omega_j^{(k+1)}(\Omega)) = n_m - \sum_l \frac{\partial\omega_j^{(k+1)}}{\partial I_l}\cdot\left(\frac{\partial\Omega^{(k+1)}}{\partial I}\right)^{-1}_{lm}.$$

Note that if

$$\sum_m \left| n_m - \sum_l \left(\frac{\partial\omega_l^{(k+1)}}{\partial I_l}\right)\left(\frac{\partial\Omega^{(k+1)}}{\partial I}\right)^{-1}_{lm}\right| > c > 0, \tag{9.7}$$

we can, by moving Ω a distance δ/c in the direction of n, insure that $|n\cdot\Omega-\omega_j^{(k+1)}(\Omega)| \geqq \delta$.

By the estimates of Sect. 4 (Lemma 4.7 and Eq. (4.12)) combined with the estimates of induction hypothesis $(k+1.1)$ of Proposition 5.1, we have $\|(\partial\Omega^{(k+1)}/\partial I)^{-1}\| \leqq C(N)\varepsilon^{-1}$ and $|\partial\omega_j^{(k+1)}/\partial I_l| \leqq c\varepsilon/j^{2\zeta_k}$. Thus,

$$\sum_l \left|\left(\frac{\partial\omega_j^{(k+1)}}{\partial I_l}\right)\left(\frac{\partial\Omega^{(k+1)}}{\partial I}\right)^{-1}_{lm}\right| \leqq C(N)/j^{2\zeta_k}.$$

By the remarks of the previous paragraph we see that we need only consider those cases in which $|n|+j \geqq \mathcal{O}(\varepsilon^{-5/4(\tau+1)})$. In this case, the estimates in the preceding sentence imply

$$\left| n_m - \sum_l \left|\left(\frac{\partial\omega_j^{(k+1)}}{\partial I_l}\right)\left(\frac{\partial\Omega^{(k+1)}}{\partial I}\right)^{-1}_{lm}\right|\right| \geqq c_0 > 0,$$

for some m, so we see that the set $\{\Omega|\,|n\cdot\Omega-\omega_j^{(k+1)}(\Omega)|<2[D_{k+1}^{(1)}]^{-1}[|n|+j]^{-\tau}\}$ is contained in a strip of width at most $(4/c_0)[D_{k+1}^{(1)}]^{-1}[|n|+j]^{-\tau}$ inside $B(\bar{c}v^{(0)}\varepsilon, \Omega^{(k+1)}(I^0))$, allowing us to bound (9.6) by $(4C(N)/c_0)[D_{k+1}^{(1)}]^{-1}[|n|+j]^{-\tau}$ $(\bar{c}v^{(0)}\varepsilon)^{N-1}$. Summing over n and j as before we find that (D(k + 1).2) fails for a set of points Ω whose measure is at most $C(N,\tau)[D_{k+1}^{(1)}]^{-1}(v^{(0)}\varepsilon)^{-1}\operatorname{meas}(\Omega^{(k+1)}$ $(B(v^{(0)}, I^0)))$, provided $\tau > N+2$.

Finally, we consider condition (D(k + 1).3). Just as we did in the case of (D.3) above, we will verify the case $|n\cdot\Omega^{(k+1)}-(\omega_j^{(k+1)}-\omega_l^{(k+1)})|$, the other three combinations of plus and minus signs being handled in like fashion.

We begin by showing that (D(k + 1).3) fails to hold, only under very special circumstances. Note that Proposition 9.1 implies that there exists $c>0$ such that $|\omega_j^{(0)}-j\pi|\leqq c/j$ for all $j\geqq N+1$. If we combine this with the estimates of the induction hypothesis $(k+1.1)$ of Proposition 5.1 we can readily establish that $|(\omega_j^{(k+1)}(I)-\omega_{j+m}^{(k+1)}(I))|\geqq\tilde{c}m$, for all I such that $\operatorname{dist}(I,\mathscr{Q}^{(k)})<v^{(k)}$, provided ε is sufficiently small. (This smallness condition does not depend on k.)

If we now define $\Omega^s=\sup_{I,k,l}|\Omega_l^{(k)}(I)|$, we see

$$|n\cdot\Omega^{(k+1)}(I)-(\omega_j^{(k+1)}(I)-\omega_{j+m}^{(k+1)}(I))|\geqq\tilde{c}/2\geqq 2[D_{k+1}^{(2)}]^{-1}[|n|+m]^{-4\tau},$$

for ε sufficiently small, unless $|n|\Omega^s>\tilde{c}m/2$. On the other hand we see that since (D0.3) holds we have

$$\begin{aligned}&|n\cdot\Omega^{(k+1)}(I)-(\omega_j^{(k+1)}(I)-\omega_{j+m}^{(k+1)}(I))|\\ &\quad\geqq|n\cdot\Omega^{(0)}-(\omega_j^{(0)}-\omega_{j+m}^{(0)})|-|n\cdot(\Omega^{(k+1)}(I)-\Omega^{(0)})+(\omega_{j+m}^{(k+1)}(I)-\omega_{j+m}^{(0)})\\ &\qquad-(\omega_j^{(k+1)}(I)-\omega_j^{(0)})|\geqq[D_0^{(2)}]^{-1}[|n|+m]^{-4\tau}-c\varepsilon^{(5/4)}[|n|+1]^{\tau}.\end{aligned}$$

Thus we see that (D(k + 1).3) is automatically satisfied unless $|n|+m\approx\mathcal{O}(\varepsilon^{-5/4(\tau+1)})$, which since $|n|>\tilde{c}m/2\Omega^s$, implies $|n|\approx\mathcal{O}(\varepsilon^{-5/4(\tau+1)})$.

If we recall that (D(k + 1).1) holds we obtain a second condition under which (D(k + 1).3) is automatically satisfied. Note that

$$\begin{aligned}&|n\cdot\Omega^{(k+1)}(I)-(\omega_j^{(k+1)}(I)-\omega_{j+m}^{(k+1)}(I))|\\ &\quad\geqq|n\cdot\Omega^{(k+1)}(I)-m\pi|-|(\omega_j^{(k+1)}(I)-j\pi)-(\omega_{j+m}^{(k+1)}(I)-(j+m)\pi)|\\ &\quad>2[D_{k+1}^{(1)}]^{-1}(|n|+m)^{-\tau}-(c/j^{4/5}+c/(j+m)^{4/5})\\ &\quad>2[D_{k+1}^{(1)}]^{-1}(|n|+m)^{-\tau}-2c/j^{4/5}.\end{aligned}$$

The second of these inequalities follows from the fact that the asymptotic estimates of Proposition 9.1, combined with the inductive estimate $(k+1.1)$ of Proposition 5.1 imply that there is a constant $c>0$ such that $|\omega_j^{(k+1)}(I)-j\pi|<c/j^{4/5}$ for all j.

Since $D_{k+1}^{(2)}>4D_{k+1}^{(1)}$, this last expression is bigger than $2(D_{k+1}^{(2)})^{-1}(|n|+m)^{-4\tau}$, and (D(k + 1).3) is automatically satisfied, when $j<\{(4c)[D_{k+1}^{(1)}](|n|+m)^{\tau}\}^{5/4}$.

Having now examined the situations in which (D(k + 1).3) is automatically satisfied let us examine cases where it can fail. Using the methods described in (D(k + 1).2) we can readily establish that

$$\begin{aligned}&\operatorname{meas}(\{\Omega|\,|n\cdot\Omega-(\omega_j^{(k+1)}-\omega_l^{(k+1)})|<2[D_{k+1}^{(2)}]^{-1}[|n|+|j-l|]^{-4\tau}\})\\ &\quad\leqq c[D_{k+1}^{(2)}]^{-1}[|n|+|j-l|]^{-4\tau}C(N,\tau)(cv^{(0)}\varepsilon)^{N-1}.\end{aligned}$$

In deriving this estimate, one is forced to show that

$$\sum_m\left|n_m-\sum_l\left[\frac{\partial\omega_j^{(k+1)}}{\partial I_l}-\frac{\partial\omega_{j+m}^{(k+1)}}{\partial I_l}\right]\left(\frac{\partial\Omega}{\partial I}\right)^{-1}_{lm}\right|\geqq c_0>0.$$

This follows easily, since the sum over l is bounded by $C(N)$, while $|n|\approx\mathcal{O}(\varepsilon^{-5/4(\tau+1)})$ by our remarks above. Thus the total measure of the set of points which fail to satisfy (D(k + 1).3) is

$$\operatorname{meas}(\{\Omega \mid |n\cdot\Omega-(\omega_j^{(k+1)}-\omega_{j+m}^{(k+1)})|<2[D_{k+1}^{(2)}]^{-1}[|n|+m]^{-4\tau}, \text{for some } n, j \text{ and } l\})$$

$$\leqq \sum_{\substack{n\in\mathbb{Z}^N; m\in\mathbb{Z}, m\geqq 0\\ |n|+m>c\varepsilon^{-5/4(\tau+1)}}} \left(\sum_{\substack{N<j\\ \leqq(4c[D_{k+1}^{(1)}][|n|+m]^{\tau})^{5/4}}} C[D_{k+1}^{(2)}]^{-1}[|n|+m]^{-4\tau}\cdot C(N,\tau)(cv^{(0)}\varepsilon)^{N-1}\right)$$

$$\leqq \sum_{\substack{n\in\mathbb{Z}^N; m\in\mathbb{Z}, m\geqq 0\\ |n|+m>c\varepsilon^{-5/4(\tau+1)}}} \tilde{C}[D_{k+1}^{(1)}]^{(5/4)}[D_{k+1}^{(2)}]^{(-1)}[|n|+m]^{(-11\tau/4)}(cv^{(0)}\varepsilon)^{N-1}$$

$$\leqq C(N,\tau)\frac{[D_{k+1}^{(1)}]^{5/4}}{[D_{k+1}^{(2)}]}\cdot(\varepsilon)^2\cdot(cv^{(0)}\varepsilon)^{N-1}$$

$$\leqq C(N,\tau)\frac{[D_{k+1}^{(1)}]^{5/4}}{[D_{k+1}^{(2)}]}\cdot(\varepsilon/v^{(0)})\cdot\operatorname{meas}(\Omega^{(k+1)})(B(v^{(0)},I^0)).$$

In this sequence of inequalities we have assumed that $\tau>\max(2N,4)$ in order to insure that the sum converges. Note further that the set of excluded points is very small since $D_{k+1}^{(2)} \gg [D_{k+1}^{(1)}]^{5/4}(\varepsilon/v^{(0)})$.

Combining this estimate, with our estimates on the size of the set of points where (D(k + 1).1) and (D(k + 1).2) fail, we see that by removing from $\Omega^{(k+1)}(\mathscr{Q}^{(k)})$, a set of measure less than or equal to

$$\left(C(N,\tau)[D_{k+1}^{(1)}]^{-1}(v^{(0)}\varepsilon)^{-1}+C(N,\tau)\frac{[D_{k+1}^{(1)}]^{5/4}}{[D_{k+1}^{(2)}]}(\varepsilon/v^{(0)})\right)\times\operatorname{meas}(\Omega^{(k+1)}B(v^{(0)},I^0)),$$

we can insure that $(D(k+1).1)$–(D(k + 1).4) are all satisfied. Inserting the definitions of the inductive constants we see that this set, which is by definition $\Omega^{(k+1)}(\mathscr{Q}^{(k+1)})$, removes a set of points whose measure is at most $\mathscr{O}(1/(k+1)^2|\log\varepsilon_0|)$ $\operatorname{meas}(\Omega^{(k+1)}(B(v^0,I^0))$.

We must now show that the estimates actually hold on a complex neighborhood of $\mathscr{Q}^{(k+1)}$.

Consider condition (D(k + 1).1). Let I be a point whose distance from $\mathscr{Q}^{(k+1)}$ is less than $v^{(k+1)}$. Then there is a point $\tilde{I}\in\mathscr{Q}^{(k+1)}$ such that $|I-\tilde{I}|<v^{(k+1)}$. Thus

$$|n\cdot\Omega^{(k+1)}(I)-\pi j|\geqq|n\cdot\Omega^{(k+1)}(\tilde{I})-\pi j|-|n\cdot(\Omega^{(k+1)}(I)-\Omega^{(k+1)}(\tilde{I}))|.$$

From the construction of $\mathscr{Q}^{(k+1)}$ above we have bounds on the first term on the right-hand side of this expression. The second term is bounded by the aid of Taylor's theorem and our estimates on the derivative $\partial\Omega^{(k+1)}/\partial I$. This allows us to bound the right-hand side of the previous inequality from below by

$$2[D_{k+1}^{(1)}]^{-1}[|n|+j]^{-\tau}-c\cdot|n|\varepsilon v^{(k+1)}.$$

If we now note that our remarks above show that condition (D(k + 1).1) is automatically satisfied unless $j<c|n|$ and the fact that we only consider n which satisfy $|n|\leqq M_{k+1}$ the definitions of the various inductive constants show that this expression is bounded below by

$$[D_{k+1}^{(1)}]^{-1}[|n|+j]^{-\tau},$$

completing the verification of (D(k + 1).1).

The verification of (D(k + 1).2) is very similar and so we conclude by checking (D(k + 1).3). We choose a point $\tilde{I}$ just as in the previous paragraph and write

$$\begin{aligned}&|n\cdot\Omega^{(k+1)}(I)-(\omega_j^{(k+1)}(I)-\omega_{j+m}^{(k+1)})(I)|\\&\quad\geqq|n\cdot\Omega^{(k+1)}(\tilde{I})-(\omega_j^{(k+1)}(\tilde{I})-\omega_{j+m}^{(k+1)}(\tilde{I}))|-|n\cdot(\Omega^{(k+1)}(\tilde{I})-\Omega^{(k+1)}(I))\\&\qquad-(\omega_j^{(k+1)}(\tilde{I})-\omega_{j+m}^{(k+1)}(\tilde{I}))-(\omega_j^{(k+1)}(I)-\omega_{j+m}^{(k+1)}(I))|.\end{aligned}$$

Once again we bound this using the bounds derived for $\tilde{I}\in\mathscr{Q}^{(k+1)}$ above and Taylor's theorem and we find that this expression is bounded from below by

$$2[D_{k+1}^{(2)}]^{-1}[|n|+m]^{-4\tau}-c\cdot(|n|+2)\varepsilon\nu^{(k+1)}.$$

We now recall that we showed that (D(k + 1).3) was satisfied automatically unless $|n|>\tilde{c}m/2\Omega^s$, and use the fact that $|n|<M_{k+1}$ and we find that this expression is bounded below by

$$[D_{k+1}^{(2)}]^{-1}[|n|+m]^{4\tau}.$$

This completes the verification that the small denominator estimates hold on a complex neighborhood of $\mathscr{Q}^{(k+1)}$.

Thus far we have derived estimates on sets in $\Omega^{(k)}(B(\nu^{(0)},I^0))$ for which the small denominator estimates fail, but this immediately yields estimates on the size of the sets in $B(\nu^{(0)},I^0)$ for which they fail. To see this note that the estimates at the end of Sect. 4 when combined with the estimates of hypothesis (k.1) of Proposition 5.1, and pair of dimensional estimates then imply that

$$\det\left(\frac{\partial\Omega^{(k)}}{\partial I}\right)=\varepsilon^N\cdot(C_1+\mathscr{O}(\sqrt{\varepsilon})).$$

Thus, we immediately conclude that

$$\begin{aligned}&\text{meas}\frac{(\{I\in\mathscr{Q}^{(k)}|\,|n\cdot\Omega^{(k+1)}(I)|<[D_{k+1}^{(1)}]^{-1}|n|^{-\tau}\})}{\text{meas}\,(B(\nu^0,I^0))}\\&\quad\leqq\frac{(1+\mathscr{O}(\sqrt{\varepsilon}))\,\text{meas}\,(\{\Omega\in\Omega^{(k+1)}(\mathscr{Q}^{(k)})|\,|n\cdot\Omega|<[D_{k+1}^{(1)}]^{-1}|n|^{-\tau}\})}{\text{meas}\,(\Omega^{(k+1)}(B(\nu^0,I^0)))},\end{aligned}$$

and similarly for conditions (Dk.2)–(Dk.3).

Combining this with our previous estimate for the right-hand side of this inequality we see that

$$\text{meas}\,(\mathscr{Q}^{(k)}\backslash\mathscr{Q}^{(k+1)})\leqq\mathscr{O}(1/k^2|\log\varepsilon_0|)\,\text{meas}\,(B(\nu^0,I^0)).\tag{9.8}$$

Since $\mathscr{Q}^{(0)}=B(\nu^0,I^0)$ we see immediately that $\text{meas}\,(\mathscr{Q}^{(k)})\geqq(1-\mathscr{O}(1/|\log\varepsilon_0|))$ $\text{meas}\,(B(\nu^0,I^0))$ for all k, and combining this observation with (9.8) we obtain $\text{meas}\,(\mathscr{Q}^{(k+1)})\geqq(1-\mathscr{O}(1/k^2|\log\varepsilon_0|))\,\text{meas}\,(\mathscr{Q}^{(k)})$. This verifies the last unproven statement of Proposition 5.1.

Once one knows that the denominators obey condition (Dk.1)–(Dk.3) one can replace (Dk.2) and (Dk.3) by the forms which we used in Sect. 5 and which we denote (Dk.2)′ and (Dk.3)′. Recall that $\omega_j^{(0)}\sim(j\pi)$. Thus, there is some constant, c, such that $\omega_j^{(k)}(I)>cj$, for all $j>N$, $k=1,2,\ldots$, and $I\in\mathbb{C}^N$ satisfying $\text{dist}\,(I,\mathscr{Q}^{(k)})<\nu^{(k)}$.

Let $\Omega^s = \sup_{I,l,k} |\Omega_l^{(k)}(I)|$. Then if $j \geqq 2|n|\Omega^s/c$,

$$|n \cdot \Omega^{(k)}(I) \pm \omega_j^{(k)}(I)|^{-1} \leqq \frac{2c}{j},$$

regardless of whether (Dk.2) holds or not. On the other hand, if (Dk.2) holds and $j < 2|n|\Omega^s/c$, we have

$$|n \cdot \Omega^{(k)}(I) \pm \omega_j^{(k)}(I)|^{-1} \leqq [D_k^{(1)}][|n| + j]^\tau \leqq [D_k^{(1)}]\left(\frac{2\Omega^s}{c}\right)^2 \frac{|n|^{\tau+1}}{j}.$$

Thus, for j large, the ‘small" denominators actually help the convergence of the sum over j! Similarly, we find

$$|n \cdot \Omega^{(k)}(I) \pm (\omega_j^{(k)}(I) \pm \omega_l^{(k)}(I)|^{-1} \leqq [D_k^{(2)}]\left((c'\Omega^s)^2 \frac{|n|^{4\tau+2}}{(1 + |j-l|)}\right).$$

Thus we have

Corollary 9.4. *Under the hypotheses of Proposition 9.2 conditions* (Dk.2). *and* (Dk.3) *can be replaced by*

(Dk.2)′ $$|n \cdot \Omega^{(k)}(I) \pm \omega_j^{(k)}(I)|^{-1} \leqq [D_k^{(1)}](c'\Omega^s)^2 \frac{|n|^{\tau+1}}{j}$$

and

(Dk.3)′ $$|n \cdot \Omega^{(k)}(I) \pm (\omega_j^{(k)}(I) \pm \omega_l^{(k)}(I))|^{-1} \leqq [D_k^{(2)}](c'\Omega^s)^2 \frac{|n|^{4\tau+2}}{1 + |j-l|}.$$

We conclude by completing the proof of Theorem 2.3. We noted in Sect. 5 that if $I \in \mathscr{Q}^{(\infty)} = \bigcap_{k \geqq 0} \mathscr{Q}^{(k)}$, one can repeat the iterative step infinitely often, and construct a quasi-periodic orbit corresponding to this value of I. The estimates of the induction hypothesis (k.1) of Proposition 5.1 imply that $\lim_{k \to \infty} \Omega^{(k)}(I) = \Omega^{(\infty)}(I)$ exists which gives the frequencies of the quasi-periodic orbit. We prove Theorem 2.3 by showing that the set of frequencies $\Omega^{(\infty)}(\mathscr{Q}^{(\infty)})$ has large measure. In fact, using the estimates of (k.1), and the Whitney embedding theorem as was done in [CG] or [P3] one can construct a diffeomorphism $\tilde{\Omega}^{(\infty)}$ on $B(v^{(0)}, I^0)$ such that $\det(\partial\tilde{\Omega}^{(\infty)}/\partial I) = C \cdot \varepsilon^N(1 + \mathcal{O}(\sqrt{\varepsilon}))$, and if $I \in \mathscr{Q}^{(\infty)}, \Omega^{(\infty)}(I) = \tilde{\Omega}^{(\infty)}(I)$. This allows us to estimate

$$\frac{\text{meas}\,(\tilde{\Omega}^{(\infty)}(\mathscr{Q}^{(\infty)}))}{\text{meas}\,(\tilde{\Omega}^{(\infty)}(B(v^{(0)}, I^0)))} \geqq (1 - \mathcal{O}(\sqrt{\varepsilon}))\frac{\text{meas}\,(\mathscr{Q}^{(\infty)})}{\text{meas}\,(B(v^{(0)}, I^0))} \geqq \left(1 - \mathcal{O}\left(\frac{1}{|\log \varepsilon_0|}\right)\right). \tag{9.10}$$

Since we get a quasi-periodic orbit with frequency corresponding to each point in $\tilde{\Omega}^{(\infty)}(\mathscr{Q}^{(\infty)})$, and since (9.10) implies that this set has large Lebesgue measure the proof of Theorem 2.3 is complete.

Acknowledgements. It is a pleasure to thank Claudio Albanese, Jean–Pierre Eckmann, Jürg Fröhlich, Rafael de la Llave, Serge Parmentier, and Robert Sachs for useful discussion about non-linear wave equations.

I wish also to thank Haken Eliasson and Jürgen Pöschel for explaining their work on low dimensional tori to me.

Finally, I wish to acknowledge the hospitality of the Zentrum for Theoretische Studien, ETH-Zürich, and the Dept. de Physique Théorique, Univ. de Genève where I was a visitor while much of this research was completed.

References

[AF] Albanese, C., Fröhlich, J.: Periodic solutions of some infinite-dimensional hamiltonian systems associated with non-linear partial difference equations, I. Commun. Math. Phys. **116**, 475–502 (1988)

[AFS] Albanese, C., Föhlich, J., Spencer, T.: Periodic solutions of some infinite dimensional hamiltonian systems associated with non-linear partial differential equations, II. Commun. Math. Phys. **119**, 677–699 (1988)

[B] Brézis, H.: Periodic Solutions of nonlinear vibrating strings and duality principles. Bull. AMS. **8**, 409 (1983)

[Bo] Borg, G.: Eine Umkehrung der Sturm–Liouvilleschen Eigenwertaufgabe. Acta. Math. **78**, 1–96 (1946)

[BN] Brézis, H., Nirenberg, L.: Forced vibrations for a nonlinear wave equation. Commun. Pure. Appl. Math. **31**, 1–30 (1978)

[CG] Chierchia, L., Gallavotti, G.: Smooth prime integrals for quasi-integrable Hamiltonian systems. Il Nuovo Cimento **67B**, 277 (1982)

[E] Eliasson, L. H.: Perturbations of stable invariant tori for hamiltonian systems. Ann. Scuola Norm. Su. Pisa **15**, 115–148 (1988)

[FSW] Fröhlich, J., Spencer, T., Wayne, C. E.: Localization in disordered nonlinear dynamical systems. J. Stat. Phys. **42**, 247 (1986)

[G] Graff, S. M.: On the continuation of hyperbolic invariant tori for Hamiltonian systems. J. Diff. Eqs. **15**, 1–70 (1974)

[GL] Gelfand, I., Levitan, B.: On the determination of a differential equation from its spectral function. Am. Math. Soc. Trans. **1**, 263–304 (1955)

[H] Hale, J. K.: Periodic solutions of a class of hyperbolic equations containing a small parameter. Arch. Rat. Mech. Anal. **23**, 380–398 (1967)

[N] Nikolenko, N. V.: The method of Poincaré normal forms in problems of integrability of equations of evolution type. Russ. Math. Surv. **41**, 63–114 (1986)

[P1] Pöschel, J.: Small denominators with spatial structure. Commun. Math. Phys. (to appear)

[P2] Pöschel, J.: On elliptic lower dimensional tori in hamiltonian systems. Preprint, Math. Z. (to appear)

[P3] Pöschel, J.: Integrability of Hamiltonian systems on Cantor sets. Commun. Pure. Appl. Math. **35**, 653 (1982)

[PT] Pöschel, J., Trubowitz, E.: Inverse spectral theorÿ. Boston: Academic Press 1987

[R] Rabinowitz, P.: Free vibrations for a semilinear wave equation. Commun. Pure. Appl. Math. **30**, 31–68 (1977)

[Ru] Rüssman, H.: Private communication

[V] Villot, M.: Theorie classique des perturbations et grand nombre de degres de liberte, Thèse de doctorat de l'université de Provence (1985)

[VB] Vittot, M., Bellissard, J.: Invariant tori for an infinite lattice of coupled classical rotators. Preprint, CPT-Marseille, (1985)

[Z] Zehnder, E.: Generalized implicit function theorems with application to some small divisor problems, II. Commun. Pure. Appl. Math. **29**, 41–111 (1976)

Communicated by A. Jaffe

Received February 24, 1989

Note added in proof. After I had submitted this manuscript, J. Pöschel called my attention to the related work of S. B. Kuksin [K1, K2]. Although technically Kuksin's method and the present paper are rather different, the spirit is quite similar as both use KAM theory to construct quasi-periodic solutions of non-linear wave equations like (1.1), and both use the asymptotics of the eigenvalues of the Sturm–Liouville operator to establish convergence of the canonical transformations in the theory.

[K1] Kuksin, S. B.: Hamiltonian perturbations of infinite dimensional linear systems with an imaginary spectrum. Funct. Anal. Appl. **21**, 22–37 (1987)
[K2] Kuksin, S. B.: Conservative perturbations of infinite-dimensional linear systems depending on a vector parameter **23**, 62–63 (1989)