Bounds on the Trajectories of a System of Weakly Coupled Rotators

C. Eugene Wayne[†]

Department of Mathematics, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

Abstract. In this paper we study a classical mechanical system of weakly coupled rotators on a one-dimensional lattice. Such systems are of interest in statistical mechanics. We prove that for any site in the system there is a "large" set of initial conditions for which there exists a canonical change of variables such that the trajectory of that site, in the transformed system, is essentially indistinguishable from that of an integrable system for a long (but finite) time. Alternatively, the trajectory of this site lies very close to a torus in the phase space of the system for times very long in comparison with the typical period of the unperturbed rotators. All the estimates in this theory are *independent* of the number of degrees of freedom in the system. We propose this mechanism as an explanation of certain numerical experiments.

1. Introduction and Main Results

In this paper we analyze the behavior of trajectories in the classical mechanical system with Hamiltonian

$$H(I, \phi) = \tfrac{1}{2} I \cdot I + \varepsilon \sum_{i=1}^{N} \cos(\phi_i - \phi_{i+1}), \tag{1.1}$$

where $I \in \mathbb{R}^N$ and $I \cdot I$ is the usual Euclidean inner product. More specifically given $j \in \{1, \ldots, N\}$ we will be interested in the behavior of $\{I_j(t), \phi_j(t)\}$, for long but finite times. In [8] Galgani and LoVecchio have studied a system of weakly coupled oscillators numerically. Looking at the trajectory of a single oscillator in the chain they find that as they increase the energy/mode of the system a transition occurs from a state in which the trajectory of the oscillator is confined to a relatively small region about the trajectory it would follow if uncoupled from the rest of the chain, to a state where the trajectory wanders over a large part of the phase plane. Furthermore they find that the energy/mode at which this transition occurs is the same whether the number of degrees of freedom, $N = 20$ or 100. Further numerical

studies seemed to confirm the fact that the point at which this transition occurs is independent of the number of degrees of freedom in the system [5,6]. One possible explanation of this phenomenon is that the transition represents the destruction of the invariant tori which the Kolmogorov–Arnol'd–Moser theory guarantees exist in weakly coupled systems. However, all estimates to date [7, 11] seem to show that the strength of the coupling for which the KAM theory applies must go to zero as N increases.

In a preprint of this paper we suggested that the mechanism described in the present work might explain these results. However, Benettin, Galgani and Giorgilli [2] have recently pointed out that the short range coupling in these models is destroyed when action-angle variables are introduced. Since the theory described below depends crucially on the short range nature of the interactions it will not apply to the above calculations. However, in [2] the authors also present the results of numerical experiments on the system with Hamiltonian (1.1). Their findings agree very well with our results. They observe that if the initial conditions of the system are such that some of the rotators are in resonance with their neighbors, the motion of those rotators has very little effect on the remainder of the system, i.e. the non-resonant regions undergo quasiperiodic behavior to a high degree of accuracy. This is precisely the prediction of the theory below. Namely, we show (Theorem 2.1) that given any set of initial conditions one can construct a canonical change of variables which results in a Hamiltonian consisting of three parts—an integrable part, a resonant part, and a non-resonant part. The effects of the resonant part of the interaction are localized in those regions where the initial conditions are resonant (i.e. give rise to small denominators in the iterative scheme used to construct the canonical transformation) while the non-resonant part is extremely small. Thus those sites in the non-resonant regions of the system behave essentially like an integrable system, and in Theorem 1.1 we show that given any site in the system "most" initial conditions are such that this site lies in the non-resonant region. Benettin, Galgani and Giorgilli have not yet studied the large N limit of (1.1), but they speculate that the behavior they observe should persist in that limit because of its essentially local character. Our theory predicts that this is indeed the case since all estimates presented below are *independent* of the number of degrees of freedom in the system. Thus even though the system is probably ergodic for large N, one would have to wait a very long time for the system to explore a significant portion of its phase space. We note that in [2], evidence is presented that inside the resonant regions of the system a kind of "local chaos" occurs. We have not been able to prove that such an effect is present, for as is often the case, the presence of chaos seems much harder to demonstrate in these systems than its absence. Because the Hamiltonian (1.1) is similar to those studied in statistical mechanics it is interesting to ask what the implications of this "weak chaos" are for the ergodic hypothesis, and even whether or not these phenomena could be observed in one dimensional physical systems.

Benettin, Galgani and Giorgilli interpret their results in light of classical perturbation theory and in [1] they show that the classical perturbation theory is well defined for the model (1.1) to any finite order in ε. The large N limit has not yet been studied, and it is not clear whether classical perturbation theory can handle the

more general short range interactions considered in [10], but their method may provide an alternate way to prove the results of the present paper for the system of rotators (1.1).

In order for a theory like that above to hold we expect that the "perturbing energy/mode," which in (1.1) is measured by constant ε, should be small in comparison with the "unperturbed energy/mode," E_0, which in this case is $\sum_{j=1}^{N} (\frac{1}{2}I_j^2)/N$. Thus we consider initial values of the action variables I in the N-dimensional (Euclidean) sphere, S_R, of radius $R = \sqrt{2E_0 N}$, for E_0 some fixed positive constant.

Theorem 1.1. *Suppose $\varepsilon < c$, for $0 < c$ a constant independent of N, (but which may depend on E_0). Then for all $j \in \{1,\ldots,N\}$ there is an open set $\mathcal{N}_j \subset S_R \times T^N$ such that if $(I,\phi) \in \mathcal{N}_j$, there is a canonical change of variables $C:(I',\phi') \to (I,\phi)$, such that the trajectories of the system with Hamiltonian $\tilde{H}(I'\phi') = H \circ C(I',\phi')$ and initial conditions $(I_0',\phi_0') = (C^{-1})(I_0,\phi_0)$ satisfy*

$$|I_j'(t) - I_{0j}'| \leq e^{-1/\varepsilon^a} \tag{1.2}$$

for some positive constant $a > 0$, independent of N. The inequality (1.2) holds for all $0 \leq t \leq T = 1/\varepsilon^{(1-b)}$, where the constant $b > 0$ may be chosen as small as one likes. Finally, $\mathcal{N}_j$ satisfies $\mathrm{vol}\, \mathcal{N}_j \geq (1 - \lambda(\varepsilon))\,\mathrm{vol}(S_R \times T^N)$, where $\lambda(\varepsilon) \sim \mathcal{O}(\varepsilon^\alpha)$ for α some small (N independent) constant.

Thus, the trajectories in the transformed system behave in a way, which for ε small is essentially indistinguishable from an integrable system. Note also that the time T for which (1.2) is valid is much longer than a typical period for the unperturbed system which is $\mathcal{O}(E_0^{-1/2})$. Given the form of the Hamiltonian (1.1) it is not surprising that we should be able to obtain some sort of estimates on the behavior of the trajectories for times of order $1/\varepsilon$. What is important here, is that we can control them to such a high degree of accuracy and that this control is independent of the number of degrees of freedom of the system. It seems likely that a result like Theorem 1.1 holds for times much longer than $1/\varepsilon$, in light of the numerical experiments of [2], and in the appendix we present some ideas as to why this might be, but at the moment these ideas fall far short of a proof. We can restate the theorem in terms of our original variables as follows. Let C be the canonical transformation constructed in Theorem 1.1, and let (I_0',ϕ_0') be as above. Consider the trajectories $I'(t) = I_0'$, $\phi'(t) = (\partial\tilde{H}/\partial I')|_{(I_0',\phi_0')}\, t + \phi_0'$ (where $\phi_i'(t)$ is evaluated $\mathrm{mod}(2\pi)$,) and set $(\tilde{I}(t), \tilde{\phi}(t)) = C(I'(t), \phi'(t))$. We then have

Theorem 1.1.' *If $\varepsilon < c$, and $0 \leq t \leq T$,*

$$|I_j(t) - \tilde{I}_j(t)| \leq e^{-1/\varepsilon^a}, |\phi_j(t) - \tilde{\phi}_j(t)| \leq e^{-1/\varepsilon^a}, \tag{1.3}$$

for some $a > 0$.

Here $(I(t), \phi(t))$ is the trajectory of the system with Hamiltonian (1.1), and initial conditions $(I_0, \phi_0) \in \mathcal{N}_j$. Thus the trajectories lie very close to the image of a torus for $0 \leq t \leq T$, a result consistent with the numerical experiments described earlier. We won't provide the details of the proof of Theorem 1.1'. It follows easily from the proof

of Theorem 1.1 and the bounds on the canonical transformation, C, which are implicit in [10].

The results of the present work are related to the studies of Nekhoroshev [9]. Nekhoroshev was able to show that given a nearly integrable Hamiltonian system, and *any* initial condition, (I_0, ϕ_0), then for $0 \leqq t \leqq T = e^{1/\varepsilon^a}$, $\sum_j |I_j(t) - I_{0j}| \leqq \varepsilon^b$, for some $b > 0$. Note that while Nekhoroshev does not have such good control over the trajectory at a particular site as Theorem 1.1 provides (partly because he is not working with systems with short range interactions and partly because he is considering arbitrary initial conditions) his result holds for an enormously long time. In particular it tells us that any Arnol'd diffusion which occurs in such a system must happen at a very slow rate. (The diffusion constant—if it exists—is bounded by e^{-1/ε^a}. See [4] for a (non-rigorous) discussion of these matters.) The drawback of Nekhoroshev's result from our point of view is that the constant a, as well as the degree of non-integrability which is allowed in order to apply the theorem, all go to zero as the number of degrees of freedom becomes large, which is the situation in which we are interested. (For instance, the constant $a \sim N^{-2}$ in [9].) The implications of results like those of Nekhoroshev, for systems studied in statistical mechanics is discussed in [3].

2. The Proof of Theorem 1.1

We review the results and terminology of [10] that we need below. Given a set $V \subset \mathbb{R}^N$, an N-vector ρ, and $\xi > 0$, define

$$W(\rho, \xi; V) = \bigcup_{I' \in V} \{(I, \phi) \in \mathbb{C}^{2N} \mid |I_i - I'_i| < \rho_i, |\operatorname{Im} \phi_i| < \xi \text{ for } i = 1, \ldots, N\}.$$

Theorem 1.1 and Corollary 1.2 of [10] imply

Theorem 2.1. *Given the Hamiltonian* (1.1), *and initial conditions* (I_0, ϕ_0), *if* $0 < \varepsilon < c$, *for some universal constant* c, *there is a canonical transformation* C; *analytic and invertible on* $W(\tilde{\rho}, 1; \{I'_0\})$, *with* $(I'_0, \phi'_0) = (C^{-1})(I_0, \phi_0)$, *and* $\tilde{\rho}$ *defined below. We write*

$$\tilde{H}(I', \phi') = H \circ C(I', \phi') = \tilde{h}(I') + \tilde{f}^{\mathrm{resonant}}(I', \phi') + \tilde{f}^{\mathrm{nonresonant}}(I', \phi'), \qquad (2.1)$$

where

$$\sup \left\{ \left| \frac{\partial \tilde{f}^{\mathrm{nonresonant}}}{\partial I'_j}(I', \phi') \right| + \left| \frac{\partial \tilde{f}^{\mathrm{nonresonant}}}{\partial \phi'_j}(I', \phi') \right| \right\} \leqq \varepsilon^{(3/2)^{k_0}} \leqq e^{-1/\varepsilon^a}, \qquad (2.2)$$

for $k_0 = $ *integer part of* $(B_3^{-1} |\ln \varepsilon|)$ *for* B_3 *some large constant. We emphasize that the constants* c *and* a *are both independent of the number of freedom in the system. We discuss the definition of, and bounds on,* $\tilde{f}^{\mathrm{resonant}}$ *below.*

Theorem 2.1 was proved by an inductive procedure. Given the Hamiltonian (1.1), let $h^0(I) = \frac{1}{2} I \cdot I$ and define $\omega^0(I) = \partial h^0 / \partial I \ (= I)$. Let $L_k = 4(3/2)^k$ and $M_k = (3/2)^k |\ln \varepsilon| / \delta$, for δ some universal constant that could be determined in the course of the proof. We note that it is convenient not to use the constant ε in (1.1) to measure the size of the interaction term in the Hamiltonian, but rather ε_0, which we define to

be the supremum of the interaction term over $W(\rho, \xi, V)$. It is easy to check that $\varepsilon_0 \sim \mathcal{O}(\varepsilon e^{2\xi})$. Define $\mathbb{X}_k = \{v \in \mathbb{Z}^N | d(\text{supp }v) \leq L_k, \ 0 < |v| \leq M_k\}$, where $d(\text{supp }v) =$ diameter of the support of v, considered as an integer valued function on the lattice $\{1, 2, \ldots, N\}$, and $|v| = \sum_{i=1}^{N} |v_i|$. The *primary resonances of order zero* are

$$R_p^0 = \{v \in \mathbb{X}_0 || \langle \omega^0(I_0), v \rangle | \leq \varepsilon^\beta [B_1 \exp[(3/2)|v| + L_0]]^{-1}\}.$$

for some constants $0 < \beta < 1$ and $B_1 > 1$, (which could again be determined from the proof) while the *secondary resonances* of *order zero* are

$$R_s^0 = \{v \in \mathbb{X}_1 \backslash R_p^0 | \overline{\text{supp }v} \cap \overline{\text{supp }v'} \neq \varnothing \text{ for some } v' \in R_p^0\}.$$

Here $\overline{\text{supp }v} = \{i, i+1, \ldots, j-1, j\}$, for i and j respectively the leftmost and rightmost sites in supp v. We then defined canonical transformations C^0, and $\tilde{C}^0$ such that $H^1(I', \phi') = H^0 \circ C^0(I, \phi) = h^1(I') + f^{1,r}(I', \phi') + f^{1,n,r}(I', \phi')$, with $h^1(I') = h^0(I') + f_0^0(I')$. (By $f_v^0(I)$ we mean the vth Fourier coefficient of the interaction term in our original Hamiltonian $H^0(I, \phi)$. For the Hamiltonian (1.1), $f_0^0(I) = 0$.) If we define $f^1(I, \phi) = H^1(I, \phi) - h^1(I)$; then $f^{1,r}(I, \phi) = \sum_{v \in R_p^0 \cup R_s^0} f_v^1(I')e^{iv\cdot\phi}$, while $f^{1,n,r}(I', \phi') = \sum_{v \notin R_p^0 \cup R_s^0}$ $f_v^1(I)e^{iv\cdot\phi}$. The canonical transformations obeyed $\tilde{C}^0 \circ C^0 = C^0 \circ \tilde{C}^0$ on their common domain, and we also defined $(I_1, \phi_1) = \tilde{C}^0(I_0, \phi_0)$. The canonical change of variables C^0 was chosen so that while $f^{1,r}$ was still $\mathcal{O}(\varepsilon_0)$, $f^{1,n,r} \sim \mathcal{O}(\varepsilon_0^{3/2})$. Continuing inductively we defined a sequence of canonical transformations $C^1, C^2, \ldots, C^{k_0}$, and their inverses $\tilde{C}^1, \tilde{C}^2, \ldots, \tilde{C}^{k_0}$, where $k_0 = $ integer part of $(B_3^{-1}|\ln \varepsilon|)$ for B_3 some large constant. Corresponding Hamiltonians were defined by

$$H^k(I', \phi') = H^{k-1} \circ C^{k-1}(I', \phi') = h^k(I') + f^k(I', \phi') = h^k(I') + f^{k,r}(I', \phi') + f^{k,n,r}(I', \phi').$$

In each case $h^k(I) = h^{k-1}(I) + f_0^{k-1}(I)$, and $f^{k,r}(I', \phi') = \sum_{v \in \tilde{R}^k} f_v^k(I)e^{iv\cdot\phi}$, with $\tilde{R}^k = \bigcup_{m=0}^{k-1} R_p^m \cup R_s^m$. As is typical in these accelerated convergence methods $f^{k,n,r} \sim \mathcal{O}(\varepsilon_k) \equiv \varepsilon_0^{(3/2)^k}$. If $\omega^k(I) = \partial h^k/\partial I$ (I) and $(I_k, \phi_k) = \tilde{C}^{k-1}(I_{k-1}, \phi_{k-1})$, the *primary resonances of order k* were given by

$$R_p^k = \{v \in \mathbb{X}_k \backslash \tilde{R}^k || \langle \omega^k(I_k), v \rangle | \leq \varepsilon^\beta [B_1 \exp[(3/2)|v| + L_k]]^{-1}\},$$

while the *secondary resonances of order k* were given by

$$R_s^k = \{v \in \mathbb{X}_{k+1} \backslash (\tilde{R}^k \cup R_p^k) | \overline{\text{supp }v} \cap \overline{\text{supp }v'} \neq \varnothing\}.$$

(The resonance vectors actually depend on the initial conditions (I_0, ϕ_0) but we suppress this dependence in the notation.)

The *k^{th} order sites*, S^k, are all sites j such that:

(i) $j \in \text{supp }v$ for some $v \in R_p^k \cup R_s^{k-1}$ (for $k = 0$ take $R_s^{-1} \equiv \varnothing$.)
(ii) $j \notin S^m$, $m = 0, \ldots, k-1$.

Define

$$(b^m)_i = \begin{cases} c_1 \rho_{m+1}/k_0 & \text{if } i \in S^m, \quad m = 0, \ldots, k_0 - 1 \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

Here $0 < c_1 < 1$ is some universal constant, and ρ_{m+1} is defined inductively by $\rho_0 = 1, \rho_{m+1} = \rho_m \varepsilon^\beta [B_2 e^{2M_k + 2L_k}]^{-1}$, for $B_2 > 1$ some universal constant. Finally set

$$(\rho^k)_i = \begin{cases} \rho_k & \text{if } i \notin \bigcup_{m=0}^{k-2} S^m \\ (\rho^{k-1})_i - 8(b^m)_i & \text{if } i \in S^m, \quad m = 0, \ldots, k - 2. \end{cases} \tag{2.4}$$

The canonical transformation C in Theorem 2.1 is the composition $C^{k_0 - 1} \circ \cdots \circ C^0$, while $\tilde{\rho} = \rho^{k_0}$. The following proposition is immediate given Corollary 2.2 of [10].

Proposition 2.2. *If $\tilde{H}(I, \phi)$ is the Hamiltonian constructed in Theorem 2.1 then*

$$\sup \left\{ \left| \frac{\partial \tilde{\mathcal{J}}^{\text{resonant}}}{\partial I_j} \right| + \left| \frac{\partial \tilde{\mathcal{J}}^{\text{resonant}}}{\partial \phi_j} \right| \right\}$$

$$\leq \begin{cases} \varepsilon^{(1/4)(3/2)k_0} \leq e^{-1/\varepsilon^a} & \text{for some constant } a > 0 \\ & \text{if } j \notin \bigcup_{m=0}^{k_0 - 1} S^m \\ \varepsilon^{(1/4)(3/2)^m} & \text{if } j \in S^m, m = 0, \ldots, k_0 - 1, \end{cases} \tag{2.5}$$

with the supremum taken over $W(\tilde{\rho}, 1; \{\tilde{I}\})$. We remark that the factor of $(1/4)$ in the exponent on the right-hand side of (2.5) could be replaced by any constant $c < 1$ without difficulty.

Proposition 2.3. *If $T = 1/\varepsilon^b$, $0 < b < 1$, then for all $0 \leq t < T$, $(I'(t), \phi'(t)) \in W(\tilde{\rho}, 1; \{I'_0\})$, where $(I'(t), \phi'(t))$ is the solution of Hamilton's equations, with Hamiltonian $\tilde{H}(I', \phi')$ and initial conditions (I'_0, ϕ'_0). (Recall $(I'_0, \phi'_0) = (C^{-1})(I_0, \phi_0) = (I_{k_0}, \phi_{k_0}).)*

Proof. By Hamilton's equations

$$|I'_j(t) - I'_{0j}| \leq t \cdot \sup \left\{ \left| \frac{\partial \tilde{\mathcal{J}}^{\text{resonant}}}{\partial \phi'_j} \right| + \left| \frac{\partial \tilde{\mathcal{J}}^{\text{nonresonant}}}{\partial \phi'_j} \right| \right\}. \tag{2.6}$$

By (2.2) and (2.5) we can bound the right-hand side of (2.6) by $2t e^{(1/4)(3/2)k_0}$ if $j \notin \bigcup_{m=0}^{k_0 - 1} S^m$, and $2t \varepsilon^{(1/4)(3/2)^n}$, if $j \in S^n$ for some n, as long as $I'(t)$ is in $W(\tilde{\rho}, 1; \{I'_0\})$. On the other hand, $I'(t)$ will be in $W(\rho, 1; \{I'_0\})$ so long as $|I'_j(t) - I'_{0j}| < \tilde{\rho}_j$. ($\phi'(t) \in T^N$ and hence always in $W(\tilde{\rho}, 1; \{I'_0\})$. The definition of $\tilde{\rho}$ implies $\tilde{\rho}_j \geq (1/2)\rho_{m+1}$ if $j \in S^m$, and $\tilde{\rho}_j = \rho_{k_0}$ otherwise. Using the definition of ρ_k this in turn implies that we must have

$$2t \varepsilon^{(1/4)(3/2)^m} \leq \varepsilon^{\beta m} B_2^{-m} e^{-B_4(3/2)^m} \varepsilon^{B_5(3/2)^m/\delta} \quad \text{for all } m = 0, 1, \ldots, k_0.$$ Provided ε is sufficiently small (but independent of m) we see that this inequality can be satisfied for all $0 \leq t < 1/\varepsilon^b$, for all $0 < b < 1/4$, by picking δ large and β small. Using our remark above we can replace the factor of $(1/4)$ in the exponent on the left-hand side of this expression by any constant c, with $0 < c < 1$, and we see that this allows us to pick $0 \leq t < 1/\varepsilon^b$ for $0 < b < 1$.

Remark. Note that if there are no resonances of order m_0 or lower in the system $\left(\text{i.e. } \bigcup_{m=0}^{m_0} [R_p^m \cup R_s^m] = \varnothing \right)$, the trajectory $(I'(t), \phi'(t))$ will remain in $W(\tilde{\rho}, 1; \{I_0'\})$ for all time $0 \leq t < 1/\varepsilon^{b(3/2)^{m_0}}$, and Theorem 1.1 will apply for a correspondingly longer time. Thus associated to the sites of order m, S^m, is a characteristic time scale.

Remark. It is quite likely that Theorem 1.1 actually applies for a time much longer than $1/\varepsilon^b$, since when the trajectory leaves the region $W(\tilde{\rho}, 1; \{I_0'\})$, we can reapply Theorem 2.1 to construct a new canonical transformation, with new initial conditions chosen to be the image (with respect to C) of the point at which we left $W(\tilde{\rho}, 1; \{I_k'\})$, and then reapply Proposition 2.3.

One must control how the resonance vectors R_p^k change during this process, and so far I have not been able to do that. However, in the appendix I suggest a way to control the zeroth order resonances which would extend the applicability of the theory to time of $\mathcal{O}(1/\varepsilon^2)$.

Let $\mathcal{N}_j \subset S_R \times T^N$ be the set of initial conditions (I_0, ϕ_0) for which $j \notin \bigcup_{m=0}^{k_0} S^m$. From the conditions defining the resonances we see that $\mathcal{N}_j$ is open. Furthermore, by Hamilton's equations and Proposition 2.3 we see that for $(I_0, \phi_0) \in \mathcal{N}_j$, and

$$0 \leq t < T, \; |I_j'(t) - I_{0j}'| \leq 1/\varepsilon^b e^{-1/\varepsilon^a} \leq e^{-1/\varepsilon^{a'}},$$

so (1.2) is satisfied. Thus, the proof of Theorem 1.1 will be complete if we can show that $\mathrm{vol}\, \mathcal{N}_j \geq (1 - \varepsilon^\alpha) \mathrm{vol}\, S_R$. (We assume that the volume of T^N is normalized to one.) Let

$$\Gamma_j = \left\{ (I, \phi) \in S_R \times T^N \,\Big|\, \text{if } v \in \bigcup_{m=0}^{k_0} R_p^k \text{ for } (I, \phi), \text{ then } \mathrm{dist}(j, \mathrm{supp}\, v) \geq 3L_{k_0} \right\}.$$

Proposition 2.4. $\Gamma_j \subseteq \mathcal{N}_j$.

Proof. If $j \in \overline{\bigcup_{m=0}^{k_0} S^m}$, but $j \notin \overline{\mathrm{supp}\, v}$ for some $v \in \bigcup_{m=0}^{k_0} R_p^m$, (for (I, ϕ)) we must have $j \in \overline{\mathrm{supp}\, v}$, for some $v \in \bigcup_{m=0}^{k_0-1} R_s^m$. But every point in $\overline{\mathrm{supp}\, v}$ for $v \in R_s^0$ must be within a distance L_1 of some point in $\overline{\mathrm{supp}\, v'}$, for some $v' \in \bigcup_{m=0}^{k_0} R_p^m$ (in fact for $v' \in R_p^0$). Similarly, every point in $\overline{\mathrm{supp}\, v}$, for $v \in R_s^1$ must lie within a distance $L_1 + L_2$ of some point in $\overline{\mathrm{supp}\, v'}$ for $v' \in R_p^0 \cup R_p^1 \subset \bigcup_{m=0}^{k_0} R_p^m$. Continuing in this fashion, we find that any point in $\overline{\mathrm{supp}\, v}$, for $v \in R_s^{k_0-1}$, must lie within a distance $\sum_{j=1}^{k_0} L_j \leq 3L_{k_0}$ of some point in $\overline{\mathrm{supp}\, v'}$, for $v' \in \bigcup_{m=0}^{k_0} R_p^m$. But the definition of Γ_j forbids this, so $j \notin \bigcup_{m=0}^{k_0} S^m$, and $\Gamma_j \subseteq \mathcal{N}_j$.

Proposition 2.5. $\mathrm{vol}\, \Gamma_j \geq (1 - \varepsilon^\alpha) \mathrm{vol}\, S_R$, *for some small constant α.*

By the remarks above, $\mathrm{vol}\, \mathcal{N}_j \geq \mathrm{vol}\, \Gamma_j \geq (1 - \varepsilon^\alpha) \mathrm{vol}\, S_R$, so the proof of Theorem 1.1 is complete.

The idea behind the proof of Proposition 2.5 is a simple geometrical argument. Let Γ_j^0 be the set with no zeroth order primary resonances within a distance $3L_{k_0}$ of j. For the Hamiltonian (1.1), $\omega^0(I) = I$, so (assuming vol $T^N = 1$)

$$\text{vol}[(S_R \times T^N)\setminus\Gamma_j^0] \leq \sum_{\substack{v \in \aleph_0 \\ \text{dist}(j,\text{supp } v) \leq 3L_{k_0}}} \int_{|\langle I,v\rangle| \leq \varepsilon^\beta[B_1 e^{(3/2)|v|+L_0}]^{-1}} dI.$$

The integral on the right-hand side of (2.7) is a slice of thickness $\varepsilon^\beta[B_1 e^{(3/2)|v|+L_0}]^{-1}$ out of an N dimensional sphere, and hence can be bounded by $\varepsilon^\beta BE_0^{-1/2}$ $[B_1 e^{(3/2)|v|+L_0}]^{-1}$ vol S_R, for some constant B. Since there are at most $2^{2M}2^L$ vectors v with $|v| = M$, $d(\text{supp } v) = L$, and leftmost point in supp v fixed, and since there are at most $(3L_{k_0} + L_0 + 1)$ choices for the leftmost point in supp v, vol$[(S_R \times T^N)\setminus\Gamma_j^0] \leq (3L_{k_0} + L_0 + 1)\varepsilon^\beta E_0^{-1/2}(2/e)^{L_0}$, if B_1 is sufficiently large. Since $\varepsilon^\beta L_{k_0} \leq \varepsilon^\alpha$, provided the constant B_3 in the definition of L_{k_0} is sufficiently large we obtain vol $\Gamma_j^0 \geq (1 - \varepsilon^\alpha)$ vol S_R. One continues to estimate the contributions of higher order resonances in the same fashion, but the procedure is somewhat more complicated due to the fact that $\langle \omega^k(I), v\rangle$ is required to be small rather than just $\langle I, v\rangle$, so we leave the details for the next section.

3. The Proof of Proposition 2.5

Throughout this section and the next we will assume that $\varepsilon < c$, the constant in Theorem 1.1, and won't continue to restate this hypothesis. Let

$$R_{j,\ell}^k = \left\{(I, \phi) \in S_R \times T^N | \exists v \in \bigcup_{m=0}^k R_p^m, \text{ with dist}(j, \text{supp } v) < \ell\right\}.$$

(Recall that R_p^m depends on (I_0, ϕ_0), but we have suppressed this dependence in the notation.)

Lemma 3.1. *There exist* $\alpha > 0$, $B > 1$ *such that*

$$\text{vol } R_{j,\ell}^k \leq \left\{3 \sum_{m=0}^k BL_m(1 + (2/e^c)^{L_0} + \cdots + (2/e^c)^{L_m})\right.$$
$$\left. + B\ell(1 + (2/e^c)^{L_0} + \cdots + (2/e^c)^{L_k})\right\}\varepsilon^\alpha \text{ vol } S_R, \qquad (3.1)$$

for $c = 1 - \varepsilon^\gamma$, *and* γ *some small positive constant.*

Note that $R_{j,3L_{k_0}}^{k_0} = (S_R \times T^N)\setminus\Gamma_j$. Thus, since $L_{k_0}\varepsilon^\alpha < \varepsilon^{\alpha'}$, provided the constant B_3 in the definition of k_0 is large enough, Proposition 2.5 follows.

Proof. The proof is by induction on k. Note that the $k = 0$ case follows easily from the discussion following Proposition 2.5.

Now consider the case $k = 1$. Recall that $\omega^1(I) = \omega^0(I) + (\partial f_0^0/\partial I)(I)$, and $(I_1, \phi_1) = \tilde{C}^0(I_0, \phi_0) = (I_0 + \psi^0(I_0, \phi_0), \phi_0 + \pi^0(I_0, \phi_0))$. The last equality coming from (3.14) of [10]. Hence $\omega^1(I_1) = \omega^1(I_0 + \psi^0(I_0, \phi_0)) \equiv \tilde{\omega}_{\phi_0}^1(I_0)$. Thus we have a function defined on $S_R \times T^N$ with the property that $v \in R_p^1$ (of (I, ϕ)) if and only if $v \notin \tilde{R}^0$, $v \in \aleph_1$,

and $|\langle \tilde{\omega}_\phi^1(I), v \rangle| \leqq \varepsilon^\beta [B_1 e^{(3/2)|v|+L_1}]^{-1}$. Note that since there are only a finite number of choices for the sets of zeroth order resonances there will be only a finite number of different choices for the generating function of the canonical transformations $\tilde{C}^0$ as our initial conditions range over $S_R \times T^N$. Furthermore, the estimates of [10] show that if the generating function is fixed, $(I', \phi') = \tilde{C}^0(I, \phi)$ depends smoothly on (I, ϕ). The boundary between regions where different canonical transformations (i.e. those with different generating functions) are defined is determined by the conditions $|\langle \omega^0(I), v \rangle| = \varepsilon^\beta [B_1 e^{(3/2)|v|+L_0}]^{-1}$, for $v \in \mathbb{X}_0$. (To keep our notation uniform, define $\tilde{\omega}_\phi^0(I) = \omega^0(I)$.) Thus, $\omega_\phi^1(I)$ may be discontinuous across these hyperplanes. However, by deleting a set of arbitrarily small volume about these hyperplanes we obtain an open subset $\mathscr{D}^1 \subset S_R \times T^N$ on which $\tilde{\omega}_\phi^1$ is smooth, and where the estimates of [10] can be used to bound its derivatives.

Now consider the case of arbitrary k. Suppose we have constructed (for each $(I_0, \phi_0) \in S_R \times T^N$) the sequence of hamiltonians $H^0(I, \phi), H^1(I, \phi), \ldots, H^k(I, \phi)$, and the sequence of points $(I_j, \phi_j) = \tilde{C}^{j-1} \circ \cdots \circ \tilde{C}^0(I_0, \phi_0)$, for $j = 1, \ldots, k$. (To save writing let $\tilde{C}_j = \tilde{C}^j \circ \cdots \circ \tilde{C}^0$.) As before there are finitely many choices for $\tilde{C}_j$ corresponding to all possible different choices of sets of resonance vectors. Write $\tilde{C}_j(I, \phi) = (I + \psi^j(I, \phi), \phi + \pi^j(I, \phi))$ and define $\omega_\phi^j(I) = \omega^j(I + \psi^{j-1}(I, \phi))$. Once again the discontinuities in $\tilde{\omega}_\phi^k$ occur when $|\langle \tilde{\omega}_\phi^j(I), v \rangle| = \varepsilon^\beta [B_1 e^{(3/2)|v|+L_j}]^{-1}$, for some $v \in \mathbb{X}_j$, $v \notin \tilde{R}^j, j = 0, \ldots, k-1$. It is easy to show (using the estimates of [10]) that the set of (I, ϕ) for which one of these equalities is satisfied is just the union of a finite number of hypersurfaces in $S_R \times T^N$. (One shows that $\langle \omega_\phi(I), v \rangle$ is a small perturbation of $\langle I, v \rangle$ and then uses an inverse function theorem argument.) Thus by excluding a set of arbitrarily small measure it is possible to construct an open subset $\mathscr{D}^k \subset S_R \times T^N$ on which $\tilde{\omega}_\phi^k(I)$ is smooth, and where one has:

Lemma 3.2.

$$\sup_{\mathscr{D}^k} \left| \frac{\partial \tilde{\omega}_{\phi i}^k}{\partial I_j}(I) - \delta_{ij} \right| \leqq \min \begin{cases} \varepsilon^{(3/8)} \\ \varepsilon^{(1/8)|i-j|} \end{cases}. \tag{3.2}$$

The proof of this lemma is presented in the next section.

Given a set $A \subset S_R \times T^N$, let $\chi(A)$ denote its characteristic function, and let $R_{j,\ell}^{k;c}$ denote the complement of $R_{j,\ell}^k$. Then

$$\text{vol } R_{j,\ell}^k = \int\limits_{T^N} \int\limits_{S_R} \chi(R_{j,\ell}^k) dI d\phi$$

$$\leqq \int\limits_{T^N} \int\limits_{S_R} \chi(R_{j,\ell+L_K}^{k-1}) dI d\phi + \int\limits_{T^N} \int\limits_{S_R} \chi(R_{j,\ell+L_K}^{k-1;c}) \chi(R_{j,\ell}^k) dI d\phi, \tag{3.3}$$

where $dI = \prod\limits_{i=1}^N dI_i$ and $d\phi = \prod\limits_{i=1}^N (d\phi_i/2\pi)$. By the induction hypothesis, (3.1) bounds the first integral on the right-hand side of (3.3) by

$$\{(1 + (2/e^c)^{L_0} + \cdots + (2/e^c)^{L_{k-1}})(\ell + L_k)B$$

$$+ 3 \sum_{m=0}^{k-1} BL_m(1 + (2/e^c)^{L_0} + \cdots + (2/e^c)^{L_m})\} \varepsilon^\alpha \text{ vol } S_R. \tag{3.4}$$

The second term on the right-hand side of (3.3) is controlled by

$$\int_{T^N}\int_{S_R} \chi(R_{j,\ell+L_K}^{k-1;c})\chi(R_{j,\ell}^k)dId\phi$$

$$\leq \delta + \sum_{\substack{v\in\mathbb{X}_k \\ \text{dist}(j,\text{supp}\,v)\leq\ell}} \int_{\substack{|\langle\bar\omega_\phi^k(I),v\rangle|\leq\varepsilon^\beta[B_1 e^{(3/2)|v|}+L_k]^{-1} \\ (I,\phi)\in\mathscr{D}^k}} d\phi\int dI\chi(R_{j,\ell+L_K}^{k-1;c}), \qquad (3.5)$$

where δ bounds the volume removed from $S_R\times T^N$ in constructing $\mathscr{D}^k$, and may, as we remarked, be chosen arbitrarily small. We note that the presence of the factor of $\chi(R_{j,\ell+L_k}^{k-1;c})$ in the integrand on the right-hand side of (3.5) allows us to ignore the contributions of all resonances of order less than k, in the sum over v. The right-hand side of (3.5) is bounded with the help of the next lemma which we prove in Sect. 4.

Lemma 3.3. *If* dist $(j,\text{supp}\,v)\leq\ell$, *and* $v\in\mathbb{X}_k$,

$$\int_{\substack{|\langle\bar\omega_\phi^k(I),v\rangle|\leq\varepsilon^\beta[B_1 e^{(3/2)|v|}+L_k]^{-1} \\ (I,\phi)\in\mathscr{D}^k}} d\phi\int dI\chi(R_{j,\ell+L_k}^{k-1;c}) \leq \varepsilon^\beta e^{-(3/2)|v|}e^{-cL_k}(\text{vol}\,S_R)/E_0^{1/2}. \qquad (3.6)$$

where c is the same constant as in Lemma 3.1.

Since there are at most $2^L 2^{2M}$ vectors v with $|v| = M$, $d(\text{supp}\,v) = L$, and the leftmost point in supp v fixed, and also at most $(L_k + \ell + 1)$ choices for the leftmost point in supp v given that dist$(j,\text{supp}\,v)\leq\ell$, we see that (3.6) bounds the right-hand side of (3.5) by

$$\delta + \sum_{M=1}^\infty\sum_{L=0}^{L_k}(L_k + \ell + 1)2^L 2^{2M}\varepsilon^\beta e^{-(3/2)|M|}e^{-cL_k}(\text{vol}\,S_R)/E_0^{1/2}$$

$$\leq \delta + B(L_k + \ell + 1)(2/e^c)^{L_k}\varepsilon^\beta(\text{vol}\,S_R)/E_0^{1/2}, \qquad (3.7)$$

which when combined with (3.4) yields the general case of Lemma 3.1. (We use the fact that $\varepsilon^\beta/E_0^{1/2}$ can be taken less than ε^α, for some $\alpha < \beta$.)

4. The Last Lemmas

We finish by proving Lemmas 3.2 and 3.3. We first need some additional results from [10]. Let $\xi_0 = \frac{1}{4}|\ln\varepsilon|$, $\xi_k = \xi_{k-1} - 3(\delta + 2)$, for δ some constant that could be determined from the proof of [10]. Note that if the constant B_3 in the definition of k_0 is large enough $\xi_{k_0}\geq 1$. The transformations C^k and $\tilde C^k$ whose construction was discussed in Sect. 2 are analytic on $W(\rho^{k+1}, \xi_{k+1}; \{I_{k+1}\})$ and map this set into $W(\rho^k, \xi_k; \{I_k\})$. Define

$$S(\rho^k;(I,\phi)) = \{(I',\phi')\in\mathbb{C}^{2N}\,|\,|I_j' - I_j| < (\rho^k)_j \text{ and } |\phi_j' - \phi_j| < (\rho^k)_j\}.$$

Recall that $C^k(I_{k+1}, \phi_{k+1}) = (I_k, \phi_k)$ (or equivalently $(I_{k+1}, \phi_{k+1}) = \tilde C^k(I_k, \phi_k)$.) This fact, combined with the bounds on C^k in [10] imply that if $a\in(0,1]$, the image of $S(a\rho^{k+1};(I_{k+1}, \phi_{k+1}))$ with respect to C^k actually contains $S(a\rho^{k+1}/(1 + \varepsilon_k^{1/2});(I_k, \phi_k))$ (with $\varepsilon_k = \varepsilon^{(1/2)(3/2)^k}$), which implies in particular that $\tilde C^k$ maps this set (which is contained in its domain of analyticity) into $S(a\rho^{k+1};(I_{k+1}, \phi_{k+1}))$. Define x and x', elements of $\mathbb{C}^{2N}$ by $x_i = \{I_i$ if $1\leq i\leq N$, ϕ_{i-N} if $N < i\leq 2N\}$ and set $x_i' = \{I_i'$ if $1\leq i\leq N$, ϕ_{i-N}' if $N < i\leq 2N\}$. Then on the appropriate domains we have $x = x(x') = C^k(x')$ and $x' = x'(x) = \tilde C(x)$. Lemma 4.2 of [10] gives

Lemma 4.1. *Let* $\mathscr{D} = W(\rho^{k+1}, \xi_{k+1}; \{I_{k+1}\})$. *Then*

$$\sup_{\mathscr{D}} \left| \frac{\partial x_i}{\partial x_j'}(x') - \delta_{ij} \right| \leq \min\left(\varepsilon^{-2\alpha} k_0 \rho_{k+1}^{-1} \varepsilon_k B^{Lk}, e^{-m(1-\eta_k)(1-3\beta_k)\delta(i,j)}\right), \tag{4.1}$$

for some small positive α, $m = \frac{1}{2}|\ln \varepsilon|$, $\eta_k = kc_1/k_0 + c_2 \sum_{j=0}^{k} (3/2)^{-j}$, *for* c_1 *and* c_2 *universal constants chosen so that* $\eta_{k_0} < 1/8$, $8\beta_k = c_1/k_0 + c_2(3/2)^{-k}$, *with* c_1 *and* c_2 *the same constants as in the definition of* η_k, *and* $\delta(i,j) = |i(\mathrm{mod}\ N) - j(\mathrm{mod}\ N)|$.

By the analytic inverse function theorem, the matrix of derivatives $(\partial x_i'/\partial x_j)(x)$ is the inverse of the matrix $(\partial x_i/\partial x_j')(x')$. Using your favorite method for inverting matrices, for example the expansion of Appendix B of [10] we find.

Lemma 4.2. *Let* $\mathscr{D} = S(\rho^{k+1}/(1 + \varepsilon_k^{1/2}); (I_k, \phi_k))$. *Then*

$$\sup_{\mathscr{D}} \left| \frac{\partial x_i'}{\partial x_j}(x) - \delta_{ij} \right| \leq \min\left(\varepsilon^{-2\alpha} k_0 \rho_{k+1}^{-1} \varepsilon_k B^{Lk}, B^{\delta(i,j)} e^{-m(1-\eta_k)(1-3\beta_k)\delta(i,j)}\right), \tag{4.2}$$

where the constant B *in (4.2) may not be the same as that in (4.1).*

Note that given the bounds of Lemma 4.2 on $x'(x) = \tilde{C}^0(x)$, and $x'(x) = \tilde{C}^1(x)$, we can apply Lemma 4.1 of [10] to bound derivatives of $x'(x) = \tilde{C}^1 \circ \tilde{C}^0(x)$, for x in $S(\rho^2/2; (I_0, \phi_0))$. Continuing inductively we obtain

Lemma 4.3. *Let* $(I'(I, \phi), \phi'(I, \phi)) = \tilde{C}_k(I, \phi) = \tilde{C}^k \circ \tilde{C}^{k-1} \circ \cdots \circ \tilde{C}^0(I, \phi)$. *Then if* x *is as above and* $x'(x) = \tilde{C}_k(x)$, *we have*

$$\sup_{\mathscr{D}} \left| \frac{\partial x_j'}{\partial x_i}(x) - \delta_{ij} \right| \geq \min\left(\varepsilon_0^{(1-\eta_k)/2}, e^{-m(1-\eta_{k+1})\delta(i,j)}\right), \tag{4.3}$$

for $\mathscr{D} = S(\rho^{k+1}/2; (I_0, \phi_0))$, *and* $k = 0, \ldots, k_0 - 1$.

Next recall that by (2.9) of [10] on $W(\rho^k, \xi_k; \{I_k\})$,

$$\left| \frac{\partial \omega_i^k}{\partial I_j}(I) - \delta_{ij} \right| \leq \begin{cases} e^{-(1-\eta_k)m|i-j|} & \text{if} \quad i \neq j \\ Bk_0 \sum_{j=0}^{k-1} \varepsilon_j \rho_{j+1}^{-1} & \text{if} \quad i = j. \end{cases} \tag{4.4}$$

Also, since $\omega^0(I) = I$ for the present example, and $\omega^k(I) = \omega^{k-1}(I) + (\partial f^k/\partial I)(I)$, the bound (2.5) of [10] implies $|\omega_i^k(I) - I_i| \leq 2\varepsilon_0$ for $i = 1, \ldots, N$, and I in $W(\rho^k, \xi_k; \{I_k\})$. Since $\tilde{C}_k: S(\rho^{k+1}/2; (I_0, \phi_0)) \to W(\rho^{k+1}, \xi_{k+1}; \{I_{k+1}\})$ we can bound derivatives of $\omega^k(I'(I, \phi))$ on this domain, using Lemma 4.3, (4.4) and Lemma 4.1 of [10], obtaining

Lemma 4.4. *On* $S(\rho^{k+1}/2; (I_0, \phi_0))$

$$\left| \frac{\partial \omega_i^k}{\partial I_j}(I'(I, \phi)) - \delta_{ij} \right| \leq \begin{cases} \varepsilon^{3/8} & \text{if} \quad i = j \\ B^{(\delta(i,j)+1)} e^{-m(1-\eta_{k+1})\delta(i,j)} & \text{if} \quad i \neq j. \end{cases} \tag{4.5}$$

On the other hand, if $(I_0, \phi_0) \in \mathscr{D}^k$, $\tilde{\omega}_{\phi_0}^k(I_0) = \omega^k(I'(I_0, \phi_0))$, so (4.5) proves Lemma 3.2. ($I'(I, \phi)$ is as in Lemma 4.3.). We remind the reader that in [10], the constants c_1 and c_2 in the definition of η_k were chosen so that $\eta_k < (1/8)$ for $k = 0, 1, \ldots, k_0$. Note also that our remarks above, and the bound on $|(I_{k+1})_i - (I_k)_i|$ that comes from (3.15) of [10] imply that on $\mathscr{D}^k$, $|\tilde{\omega}_{\phi i}^k(I) - I_i| \leq 2\varepsilon^{1/2}$, an observation we will find useful later.

We now turn to Lemma 3.3. We will prove that for any $\phi \in T^N$, $\zeta > 0$, $v \in \mathbb{X}_k$, and $\mathrm{dist}(j, \mathrm{supp}\, v) \leq \ell$,

$$\int_{\substack{|\langle \tilde{\omega}_\phi^k, v \rangle| \leq \zeta \\ (I,\phi) \in \mathscr{D}^k \cap R_{j,\ell}^{k,-1,c} + L_k}} dI \leq (2^3 \zeta e^{\varepsilon^\gamma L_k}) \operatorname{vol} S_R / E_0^{1/2}, \tag{4.6}$$

for some $\gamma > 0$, from which (3.6) follows. We assume, to simplify some of the subsequent notation that if $j \in \mathrm{supp}\, v$, then $j \in \{1, \ldots, L_k\}$. This results in no loss of generality. Let $\mathcal{O}$ be an orthogonal transformation (from $\mathbb{R}^N \to \mathbb{R}^N$) such that $\mathcal{O}v = \hat{e}_1 |v|$, where $(\hat{e}_1)_j = \delta_{ij}$. Then if $y = \mathcal{O}I$, and $\Omega(y) = \mathcal{O}\tilde{\omega}_\phi^k(\mathcal{O}^{-1}y)$,

$$\int_{\substack{|\langle \omega_\phi^k(I), v \rangle| < \zeta \\ (I,\phi) \in \mathscr{D} \cap R}} dI = \int_{\substack{|\Omega_1(y)| \leq \zeta/|v| \\ y \in \mathcal{O}(\mathscr{D} \cap R)}} dy, \tag{4.7}$$

where we have suppressed the superscripts and subscripts on $\mathscr{D}$ and R to save space and the notation $y \in \mathcal{O}(\mathscr{D} \cap R)$ means $(\mathcal{O}^{-1}y, \phi) \in \mathscr{D} \cap R$. Now let $z_i = \Omega_i(y)$ if $i = 1, \ldots, L_k$, and $z_i = y_i$ if $i = L_k + 1, \ldots, N$. We will show below that $T : y \to z$ is one-to-one on the set of y such that $(\mathcal{O}^{-1}y, \phi) \in \mathscr{D} \cap R$, and that $|\det(\partial T/\partial y)^{-1}| \leq e^{\varepsilon^\sigma L_k}$ on that set, for some $\sigma > 0$. Thus,

$$\int_{\substack{|\Omega_1(y)| \leq \zeta/|v| \\ y \in \mathcal{O}(\mathscr{D} \cap R)}} dy \leq e^{\varepsilon^\sigma L_k} \int_{\substack{|z_1| \leq \zeta/|v| \\ z \in T(\mathcal{O}(\mathscr{D} \cap R))}} dz, \tag{4.8}$$

where $z \in T(\mathcal{O}(\mathscr{D} \cap R))$ means $T^{-1}(z) \in \mathcal{O}(\mathscr{D} \cap R)$. Let $\hat{z} = (z_2, \ldots, z_N)$. Let $(T_{z_1}^{-1}(\hat{z}))_j = (T^{-1}(z_1, \hat{z}))_j$ for $j = 2, \ldots, N$ and set $\hat{y} = T_{z_1}^{-1}(\hat{z})$. Note that $T_{z_1}^{-1}$ is one-to-one for all $(z_1, \hat{z}) \in T(\mathcal{O}(\mathscr{D} \cap R))$. We show below that $\sup_{z_1, \hat{z}} |\det(\partial T_{z_1}^{-1}/\partial \hat{z})| \leq e^{\varepsilon^{\sigma'} L_k}$, so that the right-hand side of (4.8) is bounded by

$$2(\zeta/|v|) e^{(\varepsilon^\sigma + \varepsilon^{\sigma'}) L_k} \sup_{y_1} \int_{(y_1, \hat{y}) \in \mathcal{O}(\mathscr{D} \cap R)} d\hat{y}.$$

But since $\mathcal{O}$ is an orthogonal transformation, and $\mathscr{D} \subset S_R \times T^N$, $(y_1, \hat{y}) \in \mathcal{O}(\mathscr{D} \cap R)$ implies $(y_1, \hat{y}) \in S_R$, and hence that $\|\hat{y}\| \leq \sqrt{2E_0 N}$, where $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^{N-1}$. Thus, $\int_{(y_1, \hat{y}) \in \mathcal{O}(\mathscr{D} \cap R)} d\hat{y} \leq$ volume of an $N - 1$ dimensional sphere of radius $\sqrt{2E_0 N}$ and (4.6) follows easily.

We now bound the two Jacobian determinants and show that $T(y)$ is single valued. First note that $\mathcal{O}$ may be chosen so that $\mathcal{O}_{ij} = \delta_{ij}$ if either i or $j > L_k$. Also, $(\partial T_i/\partial y_j)(y) = \delta_{ij}$ if i or $j > L_k$. Thus, $\det(\partial T/\partial y)$ is equal to the determinant of the $L_k \times L_k$ block in the upper left-hand corner of this matrix. An easy calculation shows that $(\partial T/\partial y)(y))' = \mathcal{O}'((\partial \omega_\phi^k/\partial I)(\mathcal{O}^{-1}y))' \mathcal{O}^{-1'}$. Here, if M is an $N \times N$ matrix M' means the restriction to elements M_{ij} with $i, j \leq L_k$. Thus, $\det(\partial T/\partial y) = \det(\partial \tilde{\omega}_\phi^k/\partial I)' = \exp[\operatorname{tr}(\ln(\partial \tilde{\omega}_\phi^k/\partial I)')]$. If we write $(\partial \tilde{\omega}_\phi^k/\partial I)' = \mathbb{1} + ((\partial \tilde{\omega}_\phi/\partial I) - \mathbb{1})'$, then Lemma 3.2 bounds $|(\partial \tilde{\omega}_\phi^k/\partial I) - \mathbb{1})'_{ij}|$ by $\varepsilon^{3/8}$ if $i = j$ and $\varepsilon^{(1/8)|i-j|}$ if $i \neq j$. This information makes it easy to estimate the power series for $\ln[\mathbb{1} + ((\partial \tilde{\omega}_\phi/\partial I)(I) - \mathbb{1})']$ inductively, and we find $|\operatorname{tr} \ln \Delta(\partial \tilde{\omega}_\phi^k/\partial I)']| \leq \varepsilon^\sigma L_k$ for some $\sigma > 0$. (In fact $\sigma = (1/10)$ suffices.) Hence $e^{-\varepsilon^\sigma L_k} \leq |\det(\partial T/\partial y)| \leq e^{\varepsilon^\sigma L_k}$, verifying the first of our bounds. (Note that this is not

enough to ensure that T is one-to-one, because the domain on which T is defined is not connected.)

Now note that $((\partial T_{z_1}^{-1}/\partial \hat{z})(\hat{z}))^{-1}$ is equal to the matrix $((\partial T/\partial y)(y))'$ with its first row and first column struck out. Since $\det(\partial T/\partial y)'$ is the amount by which the transformation $(\partial T/\partial y)'$ changes the volume of an L_k dimensional cube in $\mathbb{R}^{L_k}$ we estimate $|\det(\partial T_{z_1}^{-1}/\partial \hat{z})^{-1})|$ by bounding the amount by which $(\partial T/\partial y)'$ can increase the volume of an arbitrary $L_k - 1$ dimensional cube (in $\mathbb{R}^{L_k}$). A little calculation (using Lemma 3.2) shows that if x is any unit vector in $\mathbb{R}^{L_k}$, $\|(\partial T/\partial y)'x\| \leq (1 + \varepsilon^\sigma)^{1/2}$, for some $\sigma > 0$, (we used the fact that $\varepsilon^{\sigma''}L_k < \varepsilon^\sigma$ for $k = 0, \ldots, k_0$ in this calculation) and hence, $(\partial T/\partial y)'$ cannot increase the volume of an $L_k - 1$ dimensional cube by more than a factor of $(1 + \varepsilon^\sigma)^{(L_k - 1)/2} \leq e^{\varepsilon^{\sigma'} L_k}$, which implies $|\det(\partial T_{z_1}^{-1}/\partial \hat{z})^{-1}| \leq e^{\varepsilon^{\sigma'} L_k}$ as claimed.

Finally, we show that on $\mathcal{O}(\mathcal{D} \cap R)$, $T(y)$ is single valued. Suppose $T(y^1) = T(y^2)$. Note first that this implies $y_j^1 = y_j^2$ for $j > L_k$, and $(\mathcal{O}\tilde{\omega}_\phi^k (\mathcal{O}^{-1}y^1))_j = (\mathcal{O}\tilde{\omega}_\phi^k (\mathcal{O}^{-1}y^2))_j$ if $1 \leq j \leq L_k$. This last fact in turn implies $\tilde{\omega}_{\phi j}^k(\mathcal{O}^{-1}y^1) = \tilde{\omega}_{\phi j}^k(\mathcal{O}^{-1}y^2)$, for $j = 1, \ldots, L_k$. If $I^1 = \mathcal{O}^{-1}y^1$, $I^2 = \mathcal{O}^{-1}y^2$, then $I_j^1 = I_j^2$ for $j > L_k$. This fact, coupled with the fact that I^1 and I^2 are both in $R_{j,\ell + L_k^-}^{k-1,c}$ implies that any $v \in \bigcup_{m=0}^{k-1} R_p^m$ (for either (I^1, ϕ) or (I^2, ϕ)) must satisfy supp $v \cap \{1, 2, \ldots, L_k\} = \varnothing$. Since $I_j^1 = I_j^2$ for $j > L_k$, we see that $\bigcup_{m=0}^{k-1} R_p^m$ is the same for both (I^1, ϕ) and (I^2, ϕ).

Since (I^1, ϕ) and (I^2, ϕ) have the same sets of resonance vectors, the generating functions for the canonical changes of variables $\tilde{C}^m$, $m = 1, \ldots, k - 1$ will be the same for (I^1, ϕ) and (I^2, ϕ). (See Sect. 3 of [10] for an explanation of the construction of these generating functions.) Next note our observation following Lemma 4.4 implies $|\tilde{\omega}_{\phi m}^k(I^1) - \tilde{\omega}_{\phi m}^k(I^2)| \geq |I_m^1 - I_m^2| - 4\varepsilon^{1/2}$, so that we can assume $\sup_j |I_j^1 - I_j^2| \leq 4\varepsilon^{1/2}$.

Since $4\varepsilon^{1/2} < \rho_1$ if the constant c in Theorem 1.1 is sufficiently small, we have $\sup|I_j^1 - I_j^2| < \rho_1$. Let ℓ be such that $\rho_\ell \leq \sup_j |I_j^1 - I_j^2| < \rho_{\ell-1}$. (If $\sup_j |I_j^1 - I_j^2| < \rho_k$, the necessary modification of the argument below is very easy.) Suppose $\sup_j |I_j^1 - I_j^2| = |I_m^1 - I_m^2|$. Write

$$\tilde{\omega}_{\phi m}^k(I^1) - \tilde{\omega}_{\phi m}^k(I^2) = \sum_{j=\ell}^k [\tilde{\omega}_{\phi m}^j(I^1) - \tilde{\omega}_{\phi m}^{j-1}(I^1) - \tilde{\omega}_{\phi m}^j(I^2) + \tilde{\omega}_{\phi m}^j(I^2)] + \tilde{\omega}_{\phi m}^{\ell-1}(I^1) - \tilde{\omega}_{\phi m}^{\ell-1}(I^2).$$

Recall that $\tilde{\omega}_{\phi m}^j(I^1) = \omega_m^j(I_j(I^1, \phi))$, (where $(I_j, \phi_j) = \tilde{C}_{j-1}(I^1, \phi)$) so that

$$\tilde{\omega}_{\phi m}^j(I^1) - \tilde{\omega}_{\phi m}^{j-1}(I^1) = \tilde{\omega}_m^{j-1}(I_j(I^1, \phi)) - \omega_m^{j-1}(I_{j-1}(I^1, \phi)) + \frac{\partial f_0^{j-1}}{\partial I_m}(I_j(I^1, \phi)).$$

Using (2.5) of [10] and the remark following (4.5) to bound $\partial f_0^{j-1}/\partial I_m$ and $|(I_j)_\ell - (I_{j-1})_\ell|$ respectively, we readily establish that $|\tilde{\omega}_{\phi m}^j(I^1) - \tilde{\omega}_{\phi m}^{j-1}(I^1)| \leq \varepsilon_{j-1}^{(3/4)}$. On the other hand, $\tilde{\omega}_{\phi m}^{\ell-1}(I^1)$ is an analytic function on $S(\rho^\ell/2; (I^1, \phi))$ and $\tilde{\omega}_{\phi m}^{\ell-1}(I^1)$ is analytic on $S(\rho^\ell/2; (I^2, \phi))$. Since the sets of resonance vectors for (I^1, ϕ) and (I^2, ϕ) are the same, these functions will agree on the (non-empty) intersection of these two

domains so we have an analytic function on $S(\rho', (I^1, \phi)) \cup S(\rho', (I^2, \phi))$. There is a path γ contained in this domain, joining I^1 to I^2, and made up of L_k pieces, γ_i, of length $|I_i^1 - I_i^2|$ along which only the component I_i varies. By the fundamental theorem of calculus

$$|\tilde{\omega}_{\phi m}^{\ell-1}(I^1) - \tilde{\omega}_{\phi m}^{\ell-1}(I^2)| = \left| \int_\gamma dI' \frac{\partial \tilde{\omega}_{\phi m}^\ell}{\partial I}(I') \right| \geqq |I_m^1 - I_m^2|(1 - B\varepsilon^{(1/10)}),$$

where we used (4.5) to control the derivatives in the integrand, and the fact that $|I_m^1 - I_m^2| \geqq |I_i^1 - I_i^2|$ for all i. But this implies $|\tilde{\omega}_{\phi m}^k(I^1) - \omega_{\phi m}^k(I^2)| \geqq |I_m^1 - I_m^2|$ $(1 - B\varepsilon^{(1/10)}) - 2^2\varepsilon_\ell^{1/2} \geqq \rho_\ell (1 - B\varepsilon^{(1/10)}) - 2^2\varepsilon_\ell^{1/2} > 0$. This contradicts the assumption that $\tilde{\omega}_{\phi m}^k(I^1) = \tilde{\omega}_{\phi m}^k(I^2)$ and establishes the fact that $T(y)$ is single valued.

Appendix

We outline a method we hope will extend Theorem 1.1 to all times $t < 1/\varepsilon^2$, as well as lead to other results such as a bound on the diffusion constant for Arnol'd diffusion.

Given an initial condition $(I_0, \phi_0) \in \mathcal{N}_j$, construct the canonical transformation C which eliminates non-resonance harmonics. Let the trajectories of the transformed system evolve until

$$|\langle I(t), v \rangle| = (3/4)B_v \varepsilon^\beta [e^{(3/2)|v| + L_0}]^{-1}, \tag{A.1}$$

for some $v = (0, \ldots, 0, \mp 1, \pm 1, 0, \ldots, 0)$, where $(I(t), \phi(t)) = C(I'(t), \phi'(t))$ and $B_v \equiv 1$. Note that if this doesn't occur for $0 \leqq t \sim \mathcal{O}(\varepsilon^{-2})$ we're done. Furthermore, that $(I'(t), \phi'(t))$ will not leave the region on which C is defined before (A.1) occurs is guaranteed by the observation following Lemma 3.1 of [10] that we could define C on the largest connected domain in $S_R \times T^N$ with a fixed set of resonance vectors.

Next note that because the canonical transformations, C and $\tilde{C}$, differ from the identity only by $\mathcal{O}(\varepsilon)$, and since Hamilton's equations imply that to $\mathcal{O}(1)$ $I'(t) = I_0' + \sum_{v_i : v_i \in R_p^0} c_i(t)v_i$, for $t \leqq (\varepsilon^{-2})$,

$$I(t) \cong I_0' + \sum_{v_i : v_i \in R_p^0} c_i(t)v_i. \tag{A.2}$$

We now define $B_{v'} = \frac{1}{8}$ if $v' = (0, \ldots, 0, \pm 1, \mp 1, 0, \ldots, 0)$ and $\langle v', v \rangle \neq 0$, where v is the vector in (A.1). If v' is of this form, and if $\langle v', v \rangle \neq 0$, we say v' is a zeroth order primary resonance for $\tilde{I}(t)$ if and only if $|\langle I, v' \rangle| \leqq B_v \varepsilon^\beta [e^{(3/2)|v| + L_0}]^{-1}$. If $\langle v', v \rangle = 0$, v' is a zeroth order primary resonance vector for $I(t)$ if and only if it was in R_p^0 for I_0. Such a definition allows us to begin the iterative procedure used to prove Theorem 2.1. Furthermore, (A.1), (A.2) and the fact that all $v_j \in R_p^0$ are of the form $(0, \ldots, 0, \pm 1, 0, \ldots, 0)$ imply that the set of zeroth order resonance vectors for $\tilde{I}(t)$ must be contained in that of I_0. Thus, the fact that the motion of the system is dominated by the zeroth order resonances for times $t < \mathcal{O}(\varepsilon^{-2})$ means that the number of zeroth order resonances cannot increase.

One then reapplies Theorem 2.1 to $(I(t), \phi(t))$. One will in general encounter *new* higher order resonances in this procedure, but they will be localized in the sense that

their support will be close to the support of that v which appears in (A.1). (One probably has to define the higher order resonances by a formula like (A.3), with $B_{v'}$ less than one if v' is close to v.) Having constructed the new canonical transformation for $(I(t), \phi(t))$, one repeats the entire procedure, and continues in this fashion until $t \sim \mathcal{O}(\varepsilon^{-2})$. A few remarks are in order:

1) The biggest stumbling block in the above program is the fact that if the same vector v appears in (A.1) many times in this process, $B_v \to 0$, and since the value of ε which allows us to apply Theorem 2.1 must be less than B_v, the value of ε to which Theorem 1.1 applies goes to zero. We note that this difficulty could be avoided if we could show that the set of zeroth order resonance vectors for $(I(t), \phi(t))$ was *strictly* contained in the set of resonance vectors for (I_0, ϕ_0).

2) Since we generate new higher order resonances at each stage in this process, we might worry that the site j, which originally was not an element of $\bigcup_0^{k_0} S^m$ might 'fall into" one of these sets and we would lose control of it. However, since the effects of the "renormalized" interactions in the theory decay as $\mathcal{O}(\varepsilon^{|i-j|})$, where i and j are the sites under consideration, and since the site j of interest is at least a distance L_{k_0} from any zeroth order resonance such effects should be $\mathcal{O}(\varepsilon^{L_{k_0}}) = \mathcal{O}(e^{-1/\varepsilon^a})$ and hence easily controlled.

3) Since the theorem on the elimination of non-resonant harmonics in [10] applies to any initial condition a result, like that of this appendix, which allows one to control the resonance vectors of the resulting trajectories for times longer than ε^{-1}, would yield a bound, independent of the number of degrees of freedom, on the rate of Arnol'd diffusion in the system, assuming it occurs.

References

1. Benettin, G., Galgani, L., Giorgilli, A.: Classical perturbation theory for systems of weakly coupled rotators. Preprint
2. Benettin, G., Galgani, L., Giorgilli, A.: Numerical investigations on a chain of weakly coupled rotators in light of classical perturbation theory. Preprint
3. Benettin, G., Galgani, L., Giorgilli, A.: Boltzmann's ultraviolet cutoff and Nekhoroshev's theorem on Arnold's diffusion. Preprint
4. Chirikov, B.: A universal instability of many-dimensional oscillator system, Phys. Rep. **52**, 263 (1979)
5. Casartelli, M., Diana, E., Galgani, L., Scotti, A.: Numerical computations on a stochastic parameter related to the Kolmogorov entropy. Phys. Rev. **A13**, 1921 (1976)
6. Diana, E., Galgani, L., Casartelli, M., Casati, G., Scotti, A.: Stochastic transition in a classical nonlinear dynamical system. A Lennard-Jones chain. Theor. Math. Phys. **29**, 1022 (1976)
7. Gallavotti, G.: Perturbation theory for classical Hamiltonian systems. In: Scaling and self-similarity in physics Fröhlich J. (ed.) Boston, MA: Birkhaüser Boston 1983
8. Galgani, L., LoVecchio, G.: Stochasticity thresholds for coupled oscillators. Il Nuovo Cimento **52B**, 1 (1979)
9. Nekhoroshev, N.: An exponential estimate on the time of stability of nearly integrable Hamiltonian systems. Russ. Math. Surv. **32** 1 (1977)

10. Wayne, C. E.: On the elimination of non-resonance harmonics. Commun Math. Phys (in press)
11. Wayne, C. E.: The KAM theory of systems with short range interactions, Parts I and II. Commun. Math. Phys. **96**, 311 and 330 (1984)