A KAM Theorem for Some Degenerate Hamiltonian Systems

Randall David Paul
Department of Mathematics
University of Texas
Austin, Texas
tinker@math.utexas.edu

January 15, 1998

Abstract

We prove the existence of a large measure of invariant tori for a class of perturbed integrable systems in which the unperturbed Hamiltonian is degenerate (that is, its Hessian matrix does not have full rank). This class consists of systems for which–along with certain other restrictions–the unperturbed Hamiltonian added to the average of the perturbation is nondegenerate.

Our result is similar to a theorem of Arnold, whose proof is only sketched. In addition, we obtain explicit estimates on the measure of phase space not foliated by invariant tori.

We apply our result to a system designed by Weinberg to test the current theory of quantum mechanics. The system is a perturbed simple harmonic oscillator, to which KAM theorems with nondegeneracy conditions only on the unperturbed Hamiltonian do not apply.

Keywords: KAM theory, degenerate Hamiltonian systems, normal forms, non-linear quantum mechanics.

1 Introduction

The purpose of this paper is to prove a theorem extending the results of Kolmogorov, Arnold, and Moser to some cases involving degenerate Hamiltonian systems. We will deal with Hamiltonians of the following form:

$$H(p,q) = H_0(p) + H_1(p,q) \qquad p \in \mathbb{R}^n \quad q \in \mathbb{T}^n \tag{1.1}$$

where $|H_1| < \epsilon$.

We recall that the standard KAM theorem (see, for example, [1]) states that for H_1 sufficiently smooth, and H_0 satisfying the nondegeneracy condition,

$$\det \left[\frac{\partial^2 H_0}{\partial p_i \partial p_j} \right] \neq 0 \tag{1.2}$$

for small ϵ this system behaves to a large extent like the unperturbed system. That is, a large measure of the phase space is foliated by tori invariant under the flow given by

$$\dot{q} = \frac{\partial H}{\partial p} \qquad \dot{p} = -\frac{\partial H}{\partial q} \tag{1.3}$$

One of the problems with this theory has been that many systems of practical interest fail to satisfy (1.2). There are many other conditions that have been developed that are significantly weaker, for instance the "isoenergetic nondegeneracy" conditions mentioned by Arnold [2] and Neistadt [3]. Probably the most general nondegeneracy condition under which the KAM theorem has been proven is the Rüssman condition [4]. This requires that the range of the "frequency map" $\nabla_p H_0$, is not confined (even locally) to any hyperplane. Proofs of the KAM theorem under this or similar conditions appear in [5], [6], [7], and [8].

Our work will be concerned with systems that do not necessarily satisfy this condition. We follow an idea of Arnold and derive conditions that depend on the integrable system and the leading order term in the perturbation. As a matter of fact, one of the examples we will deal with involves perturbation from a harmonic oscillator, which certainly does not satisfy Rüssman's condition. Of course, we pay the price that the first order perturbation has to satisfy some other conditions.

In [9] (section I.8), Arnold states but only sketches the proof to a somewhat weaker version of our result. For comparison we give a precise statement of this theorem in section 3. In this same work Arnold proves a more elaborate–though quite specialized–theorem involving invariant tori near a closed, elliptic orbit. Though some of the issues involved in proving this theorem and ours are similar, the methods of proof are different.

Let us begin by informally discussing our result. A precise formulation and statement will be undertaken in section 3.

We weaken the nondegeneracy condition (1.2) by considering H_0 of the form:

$$H_0(p) = \sum_{i=1}^{l+m} \omega_i p_i + C(p_1, \dots p_l)$$

where $C(p) = O(\|p\|^2)$ and our system has n degrees of freedom. This function is clearly degenerate as m of the actions appear only linearly, and the remaining $n-l-m$ actions do not appear at all.

We will require that the average over the angles, q, of the entire Hamiltonian be "weakly nondegenerate" in the following sense. We require that for all actions p in some domain, the averaged Hamiltonian,

$$\overline{H}(p) = (2\pi)^{-n} \int_0^{2\pi} \cdots \int_0^{2\pi} H(p,q)dq$$

satisfies

$$\left\| \left[\frac{\partial^2 \overline{H}(p)}{\partial p_i \partial p_j} \right] v \right\| \geq \begin{cases} \mathtt{m}_0 \|v\| & \forall v \in \mathbb{R}^l \times \{0\}^{n-l} \\ \mathtt{m}_0 \epsilon \|v\| & \forall v \in \mathbb{R}^n \end{cases}$$

Intuitively, we are using the nondegeneracies of the averaged perturbation to compensate for difficulties introduced by a degenerate unperturbed Hamiltonian.

We also impose a smallness condition on the partial average of the perturbation, ϵH_1, over the so called "fast angles" of the system, $q_1 \dots q_{l+m}$. This smallness condition is necessary, and we will present a simple example demonstrating this in section 3.

The main result of this paper is that, for sufficiently small ϵ, the phase space of such a system will contain a set of nearly full measure (that is, the measure of its complement is bounded by a power of ϵ) composed of tori invariant under the flow induced by H. We will also provide estimates of the amount of phase space covered by the tori.

As an illustration of the applications of Theorem 1 we will consider in detail a system proposed by Weinberg [10].

Weinberg suggested that, given its paramount importance, the superposition principle of quantum mechanics should be experimentally tested. One way of doing this would be to consider sufficiently general non-linear theories that do not satisfy the superposition principle and establish experimental bounds for the non-linear effects. This, of course, requires one to be able to work out the detailed predictions of the non-linear theories.

In certain simple situations, such as particles of spin 3/2, symmetry and locality considerations restrict what terms can appear. The (linear) quantum mechanics of these systems can be described by a harmonic oscillator and the non-linear effects by a small term in a restricted class.

Using Theorem 1 we conclude that many of the orbits that will appear in the non-linear theory are qualitatively similar to the orbits in linear quantum mechanics. Hence, the two theories will be difficult to distinguish with experiments that do not select the initial conditions and that only measure qualitative properties such as frequencies.

We point out that, even if these non-linear theories of quantum mechanics have difficulties as fundamental models, they could serve as "effective models" incorporating some effects of quantum electrodynamics. Our theorem, then, suggests that the incorporation of these effects would only slightly alter most of the experimental predictions.

2 Preliminaries

This paper deals with a real analytic Hamiltonian function given in action/angle coordinates. The phase space for this system is $\mathbb{R}^n \times \mathbb{T}^n$, which we write as the product:

$$(\mathbb{R}^l \times \mathbb{T}^l) \times (\mathbb{R}^m \times \mathbb{T}^m) \times (\mathbb{R}^{n-l-m} \times \mathbb{T}^{n-l-m})$$

The three subspaces consist of conjugate pairs. These will be refered to as the "varying", "fast", and "slow" variables, respectively. The subspaces will be parametrized as follows.

$$
\begin{aligned}
(p_v, q_v) &= \{(p_i, q_i)\}_{i=1}^{l} \\
(p_f, q_f) &= \{(p_i, q_i)\}_{i=l+1}^{l+m} \\
(p_s, q_s) &= \{(p_i, q_i)\}_{i=l+m+1}^{n}
\end{aligned}
$$

The flow induced by a Hamiltonian function causes the different angular variables to cycle at different rates, and the subscripts are associated with these rates. It should be emphasized that the subscripts refer to the frequencies of the angular variables only. For instance, the "fast" actions, p_f, change very slowly in general; it is only their conjugate angles which change rapidly.

The frequencies of the "varying angles", q_v, are high, but vary throughout the domain of H. The "fast" frequencies are high, but relatively constant thoughout the domain of H, while the "slow" frequencies are low. A Hamiltonian which induces this sort of flow is

$$
H(p, q) = \omega_f \cdot p_f + H_0(p_v) + \epsilon H_1(p, q) \tag{2.1}
$$

where H_0 (as a function only of the "varying" actions) has nondegenerate Hessian matrix.

In this paper functions will be analytic in complex strips of the form:

$$
D_\rho = \{(p, q) \in \mathbb{C}^n \times [\mathbb{C}^n/\mathbb{Z}^n] : \forall i \quad |p_i| < \rho, |\mathrm{Im}(q_i)| < \rho\}
$$

The supremum norm on such a function will be denoted:

$$
\|f\|_\rho \equiv \sup_{(p,q) \in D_\rho} |f(p, q)|
$$

For some fixed action, Δp, the recentered complex domain is written

$$
\breve{D}_\rho = \{(p, q) \in \mathbb{C}^n \times [\mathbb{C}^n/\mathbb{Z}^n] : \forall i \quad |p_i - \Delta p_i| < \rho, |\mathrm{Im}(q_i)| < \rho\}
$$

and the norm on the intersection of two such domains is written

$$
\|f\|_\rho^\cap \equiv \sup_{(p,q) \in D_\rho \cap \breve{D}_\rho} |f(p, q)|
$$

The projection of D_ρ onto its action variables is denoted

$$\bar{D}_\rho = \{p \in \mathbb{C}^n : \forall i \quad |p_i| < \rho\}$$

Derivatives of these functions are estimated by Cauchy's Inequality, hence

$$\left\| \frac{\partial f}{\partial q} \right\|_{\rho-\delta} \leq \frac{\|f\|_\rho}{\delta}$$

The average over all the angles, and the average over just the fast and varying angles are denoted, respectively, by:

$$\begin{aligned}
\bar{f}(p) &= \frac{1}{(2\pi)^n} \int_0^{2\pi} \cdots \int_0^{2\pi} f(p,q)\, dq \\
\check{f}(p, q_s) &= \frac{1}{(2\pi)^{l+m}} \int_0^{2\pi} \cdots \int_0^{2\pi} f(p,q)\, dq_v dq_f - \bar{f}(p)
\end{aligned}$$

A function, f, can be uniquely written as $\tilde{f} + \check{f} + \bar{f}$, where $\tilde{f}$ is zero when averaged over the "fast" and "varying" angles.

We will impose Diophantine conditions on the frequencies of the "fast" or "slow" angles as well as the frequencies of the entire system. The sets of frequencies satisfying the three notions of Diophantine that we will use in this paper are denoted by:

$$\begin{aligned}
\Omega_f^c &= \left\{ \omega \in \mathbb{R}^m : \forall k \in \mathbb{Z}^m - \{0\} \quad |\omega \cdot k| \geq \frac{2c}{|k|^{m+1}} \right\} \\
\Omega_\epsilon^{c,\alpha} &= \left\{ \omega \in \mathbb{R}^{n-m-l} : \forall k \in \mathbb{Z}^{n-m-l} - \{0\} \quad |\omega \cdot k| \geq \frac{\epsilon^\alpha c}{|k|^s} \right\} \\
\Omega^c &= \{\omega \in \mathbb{R}^n : \forall k \in \mathbb{Z}^n - (\{0\}^{l+m} \times \mathbb{Z}^{n-l-m}) \quad |\omega \cdot k| \geq \frac{c}{|k|^s}\}
\end{aligned}$$

Above, $|k| = \sum_{i=1}^n |k_i|$ and $s = \max\{n+2, n-l+\beta(m+2)\}$. $\beta = \alpha(18-8\alpha)/(15-\gamma-4\alpha)$. It is well known that all these sets approach full measure as the constant, c, approaches zero.

To simplify the presentation we have chosen not to give explicit formulas for most constants that appear. Therefore, K will denote a number–possibly changing from line to line–which depends only on the fixed constants: ρ_0,

ρ_*, $\mathtt{m}_0$, $\mathtt{m}_*$, and the various dimensions, n, l, and m. These constants and the parameters α and γ, satisfy the following restrictions.

$$\begin{array}{ll} \rho_0 > \rho_* > 0 & \qquad \mathtt{m}_0 > \mathtt{m}_* > 0 \\ 1 < \gamma \leq 2 & \qquad 1 \leq \alpha < (\gamma+3)/4 \end{array} \tag{2.2}$$

3 Statement of the Main Result

The following theorem is the main result of this paper.

Theorem 1 *Given a real analytic function, $H : D_{2\rho_0} \to \mathbb{R}$, $\|H\|_{2\rho_0} < 1/2$, of the form:*

$$H(p,q) = \omega_f \cdot p_f + H_0(p_v) + H_1(p,q)$$

where $\omega_f \in \Omega_f^{\bar{c}}$, $(0 < \bar{c} \leq 1)$ and such that

$$\forall p \in \mathrm{Re}(D_{\rho_0}) \qquad \left\| \left[\frac{\partial^2 \overline{H}(p)}{\partial p_i \partial p_j} \right] v \right\| \geq \left\{ \begin{array}{ll} \mathtt{m}_0 \|v\| & \forall v \in \mathbb{R}^l \times \{0\}^{n-l} \\ \mathtt{m}_0 \epsilon \|v\| & \forall v \in \mathbb{R}^n \end{array} \right. \tag{3.1}$$

Assume that

$$\|H_1\|_{2\rho_0} < \epsilon, \qquad\qquad \|\check{H}_1\|_{2\rho_0} < \epsilon^\gamma \tag{3.2}$$

Then, *for*

$$\epsilon < \left(\frac{\bar{c}}{K}\right)^{\frac{18-8\alpha}{\gamma+3-4\alpha}} \tag{3.3}$$

we have that

$$\mathrm{Re}(D_{\rho_0}) = F \cup f$$

where
i) F is foliated by tori invariant under the flow induced by H.

ii) The tori are parametrized by

$$\phi_{\bar{p}}(q) = (\bar{p}, q) + g(\bar{p}, q) \tag{3.4}$$

where g is analytic and $\|g\|_{\rho_0} < K\epsilon^{\frac{4}{5}}$.
iii) The flow on each torus is conjugate to a linear flow.
iv) The measure of the set, f, not foliated by invariant tori is bounded by

$$\text{meas}(f) \leq K\epsilon^{\frac{\gamma-1}{10}} \tag{3.5}$$

Furthermore, if $l = 0$, (i.e. there are no "varying" variables) then the measure of f is bounded by a function of ϵ which asymptotically approaches $\epsilon^{\frac{\gamma-1}{4}}$ as $\epsilon \to 0$. That is

$$\text{meas}(f) \leq O(\epsilon^{\frac{\gamma-1}{4}}) \tag{3.6}$$

For comparison we state Arnold's result [9] as well (using his original notation).

Theorem 2 *(Arnold) Let the Hamiltonian be analytic and of the form*

$$H(p, q) = H_0(p_0) + \epsilon H_1(p, q) + \ldots \tag{3.7}$$

on the domain F of the phase space p,q (where $p = p_0, p_1$, p_0 is a vector of n_0 dimensions, p_1 is a vector of $n - n_0$ dimensions; $q = q_0, q_1$ are, respectively, the rapid and slow angular variables (mod 2π)).

We suppose that

1) the secular part of the perturbation

$$\bar{H}_1(p) = (2\pi)^{-n_0} \int_0^{2\pi} \cdots \int_0^{2\pi} H_1(p, q) dq_0 \tag{3.8}$$

does not depend on the phases q_1 of the slow motion.

2) the determinants of the frequencies with respect to the impulses of order n_0 and n_1 do not vanish identically,

$$\left|\frac{\partial \omega_0}{\partial p_0}\right| = \left|\frac{\partial^2 H_0}{\partial p_0^2}\right| \neq 0 \qquad \left|\frac{\partial \xi_1}{\partial p_1}\right| = \left|\frac{\partial^2 \bar{H}_1}{\partial p_1^2}\right| \neq 0 \tag{3.9}$$

Then for sufficiently small $|\epsilon|$ and the majority (in the sense of Lebesgue measure) of initial conditions from F, the motion defined by the canonical *equations with Hamiltonian (3.7) is conditionally periodic. With these initial conditions the motion for all $-\infty < t < \infty$ is close to the conditionally periodic motion*

$$\dot{p}_0 = 0, \quad \dot{q}_0 = \omega_0(p_0) = \frac{\partial H_0}{\partial p_0}, \quad \dot{p}_1 = 0, \quad \dot{q}_1 = \epsilon \xi_1(p) = \epsilon \frac{\partial \bar{H}_1}{\partial p_1} \tag{3.10}$$

(with suitable conditions for the system (3.10).)

In [9] the proof of Theorem 2 was only sketched, however a complete proof of a much more elaborate theorem also appears in this paper. This theorem is proven by making a sequence of changes of variables defined on most of the phase space. The measure of the sets on which the changes of variables are not defined–the resonant regions–is estimated at each step. We, instead, follow Kolmogorov's original strategy (see [11]) of constructing a single invariant torus via an iterative scheme (that, in general, depends on the torus considered) and then showing a large measure of phase space can be filled with tori so constructed. This method provides information about the behavior of the system in a neighborhood of the torus. We believe that this method is also more flexible and leads to better quantitative estimates.

In both theorems stated above one requires that the averaged Hamiltonian be as nondegenerate as it can be given that some of the actions only appear to first order in ϵ. This weaker nondegeneracy must be compensated for by additional requirements on the average of the Hamiltonian over the rapidly changing angles.

For Theorem 1 we require that the difference between the fast angle average and the total average over all the angles be proportional to ϵ^γ. This corresponds to the requirement in Theorem 2 that the Hamiltonian averaged over the fast angles, q_0, depend on the slow angles, q_1, only on the order of the perturbation squared.

This type of condition is necessary. To see this consider the following system.

Example 1 *Let $G(p,q)$ be a Hamiltonian having two degrees of freedom and nondegenerate average, but only a small measure of invariant tori.*

Let

$$H(p,q) = p_1 + p_1^2 + \epsilon G(p_2, p_3, q_2, q_3) \tag{3.11}$$

Then for any $\epsilon > 0$ H has a small measure of invariant tori as well.

H satisfies all of the hypotheses of Theorem 1 excep the partial average condition. The fast angle is q_1 and the average over q_1 of the perturbation is the entire perturbation, ϵG. However for any nonzero ϵ the flow induced by H is just the Cartesian product of a circular flow (a constantly varying angle q_1 with a fixed action p_1) and a flow induced by ϵG. Clearly H will have no more invariant tori than G, which, by assumption, has very few.

The estimates in Theorem 1 on the measure of invariant tori lost to the perturbation are almost certainly not optimal. One reason for this is that our iterative scheme does not maintain separate estimates for the error terms which are independent of the actions p, and those which are linear in p. This simplier iterative scheme in general leads to poorer measure estimates. For example a naive application of this strategy to a smooth, nondegenerate, nearly integrable Hamiltonian system would give a bound for the measure of invariant tori lost of $\mathrm{O}(\epsilon^{1/4})$. More careful arguments (see [1] or [3]) show that the measure of this set is actually $\mathrm{O}(\epsilon^{1/2})$. It therefore does not seem unreasonable to hope that the measure estimate for the $l = 0$ case can also be improved to $\mathrm{O}(\epsilon^{1/2})$.

It is interesting that the measure estimate for the $l > 0$ case is so much worse than that for the $l = 0$ case. A priori one would think that the presence of frequencies which vary sharply would only tighten the measure estimates. We point out, though, that these frequencies (associated with the "varying" angles) have wide resonant regions. Moreover these high order terms in p_v produce large error terms which impede the convergence of a normal form. Our calculations seem to indicate that these difficulties, combined with the two scales at which resonances occur, more than make up for the positive effects of the partial nondegeneracy that exists in this case.

4 Proof of Theorem 1

Theorem 1 follows from the next theorem.

Theorem 3 Given $H : D_{\rho_0} \to \mathbb{R}$ *real analytic,* $\|H\|_{\rho_0} < 1/2$*, of the form:*

$$H(p,q) = \omega_v \cdot p_v + \omega_f \cdot p_f + \omega_s \cdot p_s + C_v(p_v) + A(q) + B(q) \cdot p + R(p,q)$$

where, for some $0 < c \leq 1$*,*

$$\omega_v \oplus \omega_f \oplus \omega_s \in \Omega^c, \quad \omega_s \in \Omega_\epsilon^{c,\alpha} \tag{4.1}$$

$$\left\| \left[\frac{\partial^2 \bar{H}}{\partial p_i \partial p_j}(0) \right] v \right\| \geq \left\{ \begin{array}{ll} \mathtt{m}_0 \|v\| & \forall v \in \mathbb{R}^l \times \{0\}^{n-l} \\ \mathtt{m}_0 \epsilon \|v\| & \forall v \in \mathbb{R}^n \end{array} \right. \tag{4.2}$$

$$\bar{B} = 0, \quad C_v = O(\|p\|^2), \quad R = O(\|p\|^2) \tag{4.3}$$

$$\|A + B \cdot p\|_{\rho_0} < \epsilon, \quad \|R\|_{\rho_0} < \epsilon \tag{4.4}$$

$$\|\breve{A} + \breve{B} \cdot p\|_{\rho_0} < \epsilon^\gamma \tag{4.5}$$

Then, *for*

$$\epsilon \leq \left(\frac{c}{K} \right)^{\frac{18-8\alpha}{\gamma+3-4\alpha}} \tag{4.6}$$

there is a canonical change of coordinates, $\psi : D_{\rho_*} \to D_{\rho_0}$*, so that:*

$$H \circ \psi(p,q) = \omega_v \cdot p_v + \omega_f \cdot p_f + \omega_s \cdot p_s + C_v(p_v) + R^+(p,q)$$

where,

$$\psi(p,q) = (p,q) + g(p,q), \quad \|g\|_{\rho_*} < \tfrac{K\epsilon}{c^2} \tag{4.7}$$

and

$$R^+ = O(\|p\|^2) \tag{4.8}$$

Moreover *let* $\breve{H}(p,q) = H(p - \Delta p, q)$ *and* $\breve{D}_{\rho_*}$ *be the domain* D_{ρ_*} *recentered at* Δp*. If* $\breve{H}$ *satisfies conditions (4.1)-(4.4) then the resulting map,* $\breve{\psi}$*, satisfies*

$$\|\psi - \breve{\psi}\|^{\cap}_{\rho_*} < \frac{K\epsilon}{c^2} \|\Delta p\| \tag{4.9}$$

We will prove Theorem 1 by showing that Theorem 3 can be applied to $H'(p,q) = H(p-\bar{p},q)$ for a large measure of actions, $\bar{p}$, in $\bar{D}_{\rho_0}$. The hypotheses of Theorem 1 imply that for each $\bar{p} \in \bar{D}_{\rho_0}$ the hypotheses of Theorem 3 hold with the exception of the Diophantine condition, (4.1) (We will choose c so that (4.6) is satisfied).

The gradient with respect to the actions of the averaged Hamiltonian, $\nabla_p \bar{H}$, provides a map from the space of actions to the space of frequencies. This map is invertible because of the nondegeneracy condition, (3.1). Let

$$S = \left[\nabla_p \bar{H}\right]^{-1} (\Omega^c \cap \Omega_\epsilon^{c,\alpha}) \cap \bar{D}_{\rho_0} \tag{4.10}$$

S is the set of actions which correspond to suitably nonresonant frequencies. We claim that this is a set of large measure for small ϵ and c.

Lemma 1 *For* $c \leq \bar{c}$

$$\mathrm{meas}(\bar{D}_{\rho_0} - S) \leq \begin{cases} \frac{K}{\mathtt{m}_0^n}(c + \left(\frac{\epsilon}{\bar{c}}\right)^{\beta-1} + c\epsilon^{\alpha-1}) & \text{if} \quad l \neq 0 \\ \frac{K}{\mathtt{m}_0^n}(\left(\frac{\epsilon}{\bar{c}}\right)^{\beta-1} + c\epsilon^{\alpha-1}) & \text{if} \quad l = 0 \end{cases}$$

To prove Lemma 1 first note that the image of $\bar{D}_{\rho_0}$ in the space of frequencies is contained in a rectangle, W.

$$\nabla_p \bar{H}(\bar{D}_{\rho_0}) \subset [-K,K]^l \times [\omega_f - K\epsilon, \omega_f + K\epsilon]^m \times [-K\epsilon, K\epsilon]^{n-l-m} \equiv W$$

Let

$$P_k^{c,s} = \{x \in \mathbb{R}^n : |x \cdot k| < \frac{c}{|k|^s}\}$$

where

$$k = k_v \oplus k_f \oplus k_s \in (\mathbb{Z}^l \oplus \mathbb{Z}^m \oplus \mathbb{Z}^{n-l-m}) - \{0\}^n$$

Consider first the area of the surface of intersection

$$\{x \in W : x \cdot k = 0\}$$

If the $n-1$ dimensional plane, $x \cdot k = 0$, intersects the l dimensional plane, $\mathbb{R}^l \times \{0\}^{n-l}$, transversally then the area will be proportional to ϵ^{n-l}. If not the area will be larger, proportional to ϵ^{n-l-1}. More precisely, the area is bounded by $K\|k\|\epsilon^{n-l}$ if $k_v \neq 0$, and $K\epsilon^{n-l-1}$ if $k_v = 0$.

Then

$$\begin{aligned}
\operatorname{meas}(P_k^{c,s} \cap W) &= \operatorname{area}(\{x \in W : x \cdot k = 0\}) \left(\frac{2c}{|k|^s}\right) \\
&\leq \begin{cases} \frac{Kc\epsilon^{n-l}}{|k|^{s-1}} & \text{if} \quad k_v \neq 0 \\ \frac{Kc\epsilon^{n-l-1}}{|k|^s} & \text{if} \quad k_v = 0 \end{cases}
\end{aligned}$$

Similarly,

$$\operatorname{meas}(P_{k_s}^{\epsilon^\alpha c,s} \cap W) \leq \frac{Kc\epsilon^{n-l-1+\alpha}}{|k|^s}$$

We estimate the measure of $W - (\Omega^c \cup \Omega_\epsilon^{c,\alpha})$ by summing over these planar regions. These sums converge, but of particular interest is the case where $k_v = 0$, but $k_f \neq 0$. Let $B_{K\epsilon}(\omega_f) \subset \mathbb{R}^{n-l}$ be the $K\epsilon$ cube about the frequency $\omega_f \oplus 0$. First note that for $\omega_f \in \Omega_f^{\bar{c}}$ and $y \in \mathbb{R}^l \times B_{K\epsilon}(\omega_f)$, if

$$|k| < \left(\frac{\bar{c}}{K\epsilon}\right)^{\frac{1}{m+2}} \tag{4.11}$$

then

$$\begin{aligned}
|y \cdot k| &\geq |\omega_f \cdot k| - |(y - \omega_f) \cdot k| \\
&\geq \frac{2\bar{c}}{|k_f|^{m+1}} - K\epsilon \|k\| \\
&\geq \frac{\bar{c}\|k\|}{|k_f|^{m+2}}
\end{aligned}$$

Therefore, since $y \in P_k^{c,s}$ implies

$$|y \cdot k| < \frac{\bar{c}\|k\|}{|k|^s}$$

we have

$$(\mathbb{R}^l \times B_{K\epsilon}(\omega_f)) \cap P_k^{c,s} = \emptyset$$

for all k with $k_v = 0$, $k_f \neq 0$, and satisfying (4.11).

We have that

$$\begin{aligned}
\operatorname{meas}\left(\bigcup_{k_v=0,k_f\neq 0} P_k^{c,s}\cap W\right) &= \operatorname{meas}\left(\bigcup_{k_v=0,|k_f|\geq\left(\frac{\bar{c}}{K\epsilon}\right)^{\frac{1}{m+2}}} P_k^{c,s}\cap W\right)\\
&\leq K\bar{c}\epsilon^{n-l-1}\sum_{k_v=0,|k_f|\geq\left(\frac{K\bar{c}}{\epsilon}\right)^{\frac{1}{m+2}}} |k|^{-s}\\
&\leq Ka\bar{c}\epsilon^{n-l-1}\left[\frac{\bar{c}}{K\epsilon}\right]^{\frac{n-l-s}{m+2}}\\
&\leq Ka\bar{c}^{1-\beta}\epsilon^{n-l-1+\beta} \qquad (4.12)
\end{aligned}$$

Here a is a constant arising from the sum, and depends only on the dimensions, n and l.

The estimate on the full set of resonant frequencies is

$$\begin{aligned}
&\operatorname{meas}(W-(\Omega^c\cup\Omega_\epsilon^{c,\alpha}))\\
&\leq \sum_{k_v\neq 0}\operatorname{meas}(P_k^{c,s}\cap W)+\sum_{\substack{k_f\neq 0\\ k_v=0}}\operatorname{meas}(P_k^{c,s}\cap W)\\
&\quad+\sum_{\substack{k_v=k_f=0\\ k_s\neq 0}}\operatorname{meas}(P_k^{\epsilon^\alpha c,s}\cap W)\\
&\leq Kc\epsilon^{n-l}\sum_{k_v\neq 0}|k|^{-s}+K\bar{c}^{1-\beta}\epsilon^{n-l-1+\beta}\\
&\quad+Kc\epsilon^{n-l-1+\alpha}\sum_{k_s\neq 0}|k|^{-s}\\
&\leq K\epsilon^{n-l}(c+(\epsilon/\bar{c})^{\beta-1}+c\epsilon^{\alpha-1}) \qquad (4.13)
\end{aligned}$$

Therefore,

$$\begin{aligned}
\operatorname{meas}(\bar{D}_{\rho_0}-S) &= \operatorname{meas}(\left[\nabla_p\bar{H}\right]^{-1}(W-\Omega^c\cap\Omega_\epsilon^{c,\alpha}))\\
&\leq \frac{1}{\mathtt{m}_0^n\epsilon^{n-l}}\operatorname{meas}(W-\Omega^c\cap\Omega_\epsilon^{c,\alpha}) \qquad (4.14)
\end{aligned}$$

which implies the $l\neq 0$ case in Lemma 1.

If $l=0$, the first sum in (4.13) is empty, and only the terms proportional to a power of ϵ contribute to the measure. □

Proceeding with the proof of Theorem 1, define a map,

$$\Theta : S \times T^n \to \mathbb{R}^n \times T^n$$
$$\Theta(\bar{p}, q) = \psi_{\bar{p}}(\bar{p}, q)$$

where $\psi_{\bar{p}}$ is the map provided by Theorem 3 for $H' = H(p - \bar{p}, q)$. The hypotheses of Theorem 1 and the fact that $\bar{p} \in S$ guarentee that Θ is well-defined. It will be shown that Θ has a Lipshitz inverse with Lipshitz constant close to one. From this we conclude that the measure of the image of Θ–which is the same as the measure of the set F–is roughly proportional to the measure of S as $\epsilon \to 0$.

By (4.7),

$$\begin{aligned}
\sup_{(p,q)\in Re(D_{\rho_0})} \|\Theta(p,q) - (p,q)\| &= \sup_{(p,q)\in Re(D_{\rho_0})} \|\psi_p(p,q) - (p,q)\| \\
&= \sup_{(p,q)\in Re(D_{\rho_0})} \|g_p(p,q)\| \\
&< \frac{K\epsilon}{c^2} \qquad (4.15)
\end{aligned}$$

Likewise,

$$\begin{aligned}
&\|\Theta(p,q) - (p,q) - \Theta(\bar{p},\bar{q}) - (\bar{p},\bar{q})\| \\
&\quad \le \|g_p(p,q) - g_{\bar{p}}(\bar{p},\bar{q})\| \\
&\quad \le \|g_p(p,q) - g_p(\bar{p},\bar{q})\| + \|g_p(\bar{p},\bar{q}) - g_{\bar{p}}(\bar{p},\bar{q})\| \\
&\quad \le \sup_{(p,q)\in Re(D_{\rho_0})} \|\nabla g_p\| \|(p,q) - (\bar{p},\bar{q})\| + \|\psi_p - \psi_{\bar{p}}\|_{\rho *} \\
&\quad \le \frac{K\epsilon}{c^2\rho_0}\|(p,q) - (\bar{p},\bar{q})\| + \frac{K\epsilon}{c^2}\|p - \bar{p}\| \\
&\quad \le \frac{K\epsilon}{c^2}\|(p,q) - (\bar{p},\bar{q})\| \qquad (4.16)
\end{aligned}$$

Therefore $(\Theta - \mathrm{Id})$ has Lipshitz constant bounded by $K\epsilon/c^2$, so by the Lipshitz inverse function theorem there exists $\Theta^{-1} : \mathrm{Image}(\Theta) \to S \times \mathbb{T}^n$ with Lipshitz constant bounded by $(1 - K\epsilon/c^2)^{-1}$. This implies that

$$\mathrm{meas}(\mathrm{Image}(\Theta)) \ge (1 - K\epsilon/c^2)\mathrm{meas}(S \times \mathbb{T}^n)$$

So if $l \neq 0$ (we will see that $\epsilon < c^2/2K$) then

$$\begin{aligned}
\mathrm{meas}(D_{\rho_0} - \mathrm{Image}(\Theta)) &\le K\mathrm{meas}(\bar{D}_{\rho_0} - S) \\
&\le K\left(c + (\epsilon/c)^{\beta-1} + c\epsilon^{\alpha-1}\right) \qquad (4.17)
\end{aligned}$$

We wish to apply Theorem 3 setting

$$c = K\epsilon^{\sigma}$$

where $\sigma = (\gamma + 3 - 4\alpha)/(18 - 8\alpha)$.

Note that this means (4.6) is satisfied, and (3.3) implies $c \leq \bar{c}$ (so Lemma 1 may be applied).

Substituting into (4.17),

$$\begin{aligned} \text{meas}(D_{\rho_0} - &\text{Image}(\Theta)) \\ &\leq K\left(\epsilon^{\sigma} + \epsilon^{(1-\sigma)(\beta-1)} + \epsilon^{\sigma+\alpha-1}\right) \end{aligned} \tag{4.18}$$

By our choice of β in Section 2, we have that $(1-\sigma)(\beta-1) = \sigma + \alpha - 1$, so the dominant terms above are the first and the third. (4.18) is minimized (as $\epsilon \to 0$) by setting $\alpha = 1$. This means that $\sigma = (\gamma - 1)/10$, implying (3.5).

If $l = 0$, we have

$$\text{meas}(D_{\rho_0} - \text{Image}(\Theta)) \leq K\epsilon^{\sigma+\alpha-1} \tag{4.19}$$

(4.19) is minimized by letting α tend toward $(\gamma+3)/4$. This forces σ to tend toward zero, constraining the size of ϵ by (3.3). This implies (3.6).

Let $F = \text{Image}(\Theta)$ and f equal its complement in $Re(D_{\rho_0})$. For each fixed $p \in S$, $\Theta(p, q) = \psi_p(p, q)$ parametrizes a torus in F, so F is foliated as claimed. These tori are invariant since, by construction,

$$H \circ \phi_{\bar{p}}(p, q) = \nabla_p \bar{H}(\bar{p}) \cdot (p - \bar{p}) + O(\|p - \bar{p}\|^2)$$

So Hamilton's equations at $p = \bar{p}$ are

$$\dot{q} = \nabla_p \bar{H}(\bar{p}) \qquad\qquad \dot{p} = 0$$

$\square$.

5 Proof of Theorem 3

The theorem will be proven with an iterative argument that uses quadratic convergence, following Kolmogorov. In our case, however, the first step of the iteration is fundamentally different. The averaged Hamiltonian, $\bar{H}$, is nondegenerate only to order ϵ in the "slow" variables. This factor is not troublesome in later steps when the perturbation is on the order of ϵ^γ or higher, but leads to some difficulty when the perturbation is of an order commensurate with the nondegeneracy. As was shown earlier (example 1), one must impose a restriction on the size of the most problematic term, $\check{H}_1$, in order for the theorem to be true at all.

Because of these difficulties, the first change of coordinates is treated separately in Lemma 2. The estimates for subsequent iterative steps will be developed in Lemma 3. We now state the lemmae and show that they imply Theorem 3, but postpone their proofs until section 6.

Let $\tau = n + 2s + 1$, and $\delta_0 = \min\{(\rho_0 - \rho_*)/2, (\mathtt{m}_0 - \mathtt{m}_*)/2\}$. Though δ_0 is a fixed constant, and therefore could be absorbed into the generic factor K, it is left separate in the following lemma for clarity. In Lemma 3, the corresponding δ's are not fixed beforehand, and therefore must be kept track of separately from the other constants.

Lemma 2 **Given** *H satisfying the hypotheses of Theorem 3 (with $\epsilon_0 = \epsilon$)* **then**, *for*

$$\epsilon_0 < \frac{c^4 \delta_0^{4\tau+3}}{K} \tag{5.1}$$

there is an analytic, canonical change of coordinates, $\phi : D_{\rho_0 - \delta_0} \to D_{\rho_0}$, of the form:

$$\phi(p,q) = (p,q) + g(p,q) \tag{5.2}$$

so that

$$H \circ \phi = \omega_v \cdot p_v + \omega_f \cdot p_f + \omega_s \cdot p_s + C_v(p_v) + [A^+(q) + B^+(q) \cdot p] + R^+(p,q)$$

where,

$$\|A^+ + B^+ \cdot p\|_{\rho_0 - \delta_0} < \frac{K \epsilon_0^\gamma}{c^4 \delta_0^{4\tau+2}} \tag{5.3}$$

$$\left\| \left[\frac{\partial^2 \overline{H \circ \phi}}{\partial p_i \partial p_j}(0) \right] v \right\| \geq \begin{cases} \mathtt{m}_+ \|v\| & \forall v \in \mathbb{R}^l \times \{0\}^{n-l} \\ \mathtt{m}_+ \epsilon_0 \|v\| & \forall v \in \mathbb{R}^n \end{cases} \tag{5.4}$$

with $\mathtt{m}_+ \geq \mathtt{m}_0 - \delta_0$

$$R^+ = O(\|p\|^2) \qquad \|R^+\|_{\rho_0-\delta_0} < \frac{K\epsilon_0}{c^2\delta_0^{2\tau+2}} \tag{5.5}$$

$$\|g\|_{\rho_0-\delta_0} < \frac{K\epsilon_0}{c^2\delta_0^{2\tau}} \tag{5.6}$$

Moreover *given* $\breve{H}$ *as in Theorem 3, the resulting transformed Hamiltonian,* $\breve{H} \circ \breve{\phi}$*, satisfies*

$$\|A^+ - \breve{A}^+\|^{\cap}_{\rho_0-\delta_0} < \frac{K\epsilon_0^{\gamma}}{c^4\delta_0^{4\tau+2}}\|\Delta p\|, \quad \|B^+ - \breve{B}^+\|^{\cap}_{\rho_0-\delta_0} < \frac{K\epsilon_0^{\gamma}}{c^4\delta_0^{4\tau+2}}\|\Delta p\| \tag{5.7}$$

$$\|H \circ \phi - \breve{H} \circ \breve{\phi}\|^{\cap}_{\rho_0-\delta_0} \leq \frac{K(1+\delta_0)\|\Delta p\|}{\delta_0} \tag{5.8}$$

and

$$\|\phi - \breve{\phi}\|^{\cap}_{\rho_0-\delta_0} < \frac{K\epsilon_0\|\Delta p\|}{c^2\delta_0^{2\tau}} \tag{5.9}$$

Lemma 3 Given *positive constants,* $\rho > \rho_*$*,* $M < 2K/\rho_0$*,* $N < 1 + K\delta_0^{-2\tau-2}$*, and a real analytic function,* $H : D_\rho \to \mathbb{R}$*,* $\|H\|_\rho < 1$ *of the form:*

$$H(p,q) = \omega_v \cdot p_v + \omega_f \cdot p_f + \omega_s \cdot p_s + [A(q) + B(q) \cdot p] + C_v(p_v) + R(p,q)$$

where,

$$\omega_v \oplus \omega_f \oplus \omega_s \in \Omega^c, \quad \omega_s \in \Omega^{c,\alpha}_{\epsilon} \tag{5.10}$$

$$C_v = O(\|p\|^2), \quad R = O(\|p\|^2), \quad \|R\|_\rho < N\epsilon \tag{5.11}$$

$$\|A + B \cdot p\|_\rho < \eta \tag{5.12}$$

$$\left\| \left[\frac{\partial^2 \bar{H}}{\partial p_i \partial p_j}(0) \right] v \right\| \geq \begin{cases} \mathtt{m}\|v\| & \forall v \in \mathbb{R}^l \times \{0\}^{n-l} \\ c^2 \mathtt{m}\epsilon \|v\| & \forall v \in \mathbb{R}^n \end{cases} \tag{5.13}$$

Then *if*

$$\eta < \frac{c^6 \delta^{2\tau+3} \epsilon^{2\alpha-1}}{K} \tag{5.14}$$

there exists an analytic, canonical change of coordinates, $\phi : D_\rho \to D_{\rho-\delta}$ *of the form*

$$\phi(p, q) = (p, q) + g(p, q) \tag{5.15}$$

so that

$$H \circ \phi = \omega_v \cdot p_v + \omega_f \cdot p_f + \omega_s \cdot p_s + A^+(q) + B^+(q) \cdot p + C_v(p_v) + R^+(p, q)$$

where,

$$\|A^+ + B^+ \cdot p\|_{\rho-\delta} < \frac{K\eta^2}{c^8 \epsilon^{4\alpha-3} \delta^{4\tau+2}} \tag{5.16}$$

$$\|R^+\|_{\rho-\delta} < (N + \delta)\epsilon \tag{5.17}$$

$$\left\| \left[\frac{\partial^2 \overline{H \circ \phi}}{\partial p_i \partial p_j}(0) \right] v \right\| \geq \begin{cases} \mathtt{m}_+\|v\| & \forall v \in \mathbb{R}^l \times \{0\}^{n-l} \\ c^2 \mathtt{m}_+\epsilon \|v\| & \forall v \in \mathbb{R}^n \end{cases} \tag{5.18}$$

with $\mathtt{m}_+ \geq \mathtt{m} - \delta$

$$\|g\|_{\rho-\delta} < \frac{K\eta}{c^2 \epsilon^{2\alpha-1} \delta^{2\tau}} \tag{5.19}$$

Moreover if $\breve{H} : \breve{D}_{\rho-\delta} \to \breve{D}_\rho$ *is another Hamiltonian satisfying the hypotheses of Lemma 3 and*

$$\|A - \breve{A}\|_\rho^\cap < \eta\|\Delta p\| \qquad \|B - \breve{B}\|_\rho^\cap < \eta\|\Delta p\| \tag{5.20}$$

$$\|H - \breve{H}\|_\rho^\cap \leq M\|\Delta p\| \tag{5.21}$$

then

$$\|A^+ - \breve{A}^+\|_{\rho-\delta}^\cap < \frac{K\eta^2\|\Delta p\|}{c^8\epsilon^{4\alpha-3}\delta^{4\tau+2}} \qquad \|B^+ - \breve{B}^+\|_{\rho-\delta}^\cap < \frac{K\eta^2\|\Delta p\|}{c^8\epsilon^{4\alpha-3}\delta^{4\tau+2}} \tag{5.22}$$

$$\|H \circ \phi - \breve{H} \circ \breve{\phi}\|_{\rho-\delta}^\cap \leq (M + \delta)\|\Delta p\| \tag{5.23}$$

and lastly,

$$\|\phi - \breve{\phi}\|_{\rho-\delta}^\cap < \epsilon\delta^2\|\Delta p\| \tag{5.24}$$

Using these lemmae the proof of Theorem 3 proceeds. After the application of Lemma 2, one tries to repeatedly apply Lemma 3. We may apply Lemma 2 since for α in the range given in section 2

$$\frac{\gamma + 3 - 4\alpha}{18 - 8\alpha} < \frac{1}{4}$$

so (4.6) implies (5.1).

For $i > 0$ we have the following recurrence relation between the error in the i-th and $i+1$-st steps.

$$\eta_{i+1} = \frac{K\eta_i^2}{c^8\epsilon^{4\alpha-3}\delta_i^{4\tau+2}} \tag{5.25}$$

We must choose the sequence, $\{\delta_i\}$, so that for a sufficiently small η_1 the sequence, $\{\eta_i\}$, converges exponentially to zero. Following [12] we solve for δ_i in (5.25).

$$\delta_i^{4\tau+2} = \frac{K\eta_i^2}{c^8\epsilon^{4\alpha-3}\eta_{i+1}}$$

We want to choose δ_i so that

$$\eta_i = 2^{-(4\tau+2)(i-1)}\eta_1$$

Substituting,

$$\begin{aligned}
\delta_i^{4\tau+2} &= \frac{K2^{-2(4\tau+2)(i-1)}\eta_1^2}{c^8\epsilon^{4\alpha-3}2^{-(4\tau+2)i}\eta_1} \\
&= \frac{2^{8\tau+4}K2^{-(4\tau+2)i}\eta_1}{c^8\epsilon^{4\alpha-3}} \qquad (5.26)
\end{aligned}$$

Therefore we choose

$$\delta_i = 2^{-i}\left[\frac{K\eta_1}{c^8\epsilon^{4\alpha-3}}\right]^{\frac{1}{4\tau+2}}$$

Lemma 2 gives a bound for η_1 in terms of ϵ_0.

$$\eta_1 < \frac{K\epsilon_0^\gamma}{c^4}$$

Noting that the parameter "ϵ" in Lemma 3 actually corresponds to ϵ_0/c^2 in Lemma 2, and substituting

$$\delta_i = 2^{-i}\left[\frac{K\epsilon_0^\gamma}{\epsilon^{4\alpha-3}c^{12}}\right]^{\frac{1}{4\tau+2}} = 2^{-i}\left[\frac{K\epsilon_0^{\gamma-4\alpha+3}}{c^{18-8\alpha}}\right]^{\frac{1}{4\tau+2}} \leq \frac{2^{-i}}{K}$$

By (2.2) we have that $\gamma - 4\alpha + 3 > 0$. This and (4.6) tell us that the series $\sum \delta_i$, may be made small by choosing the constant K, in (4.6) (and, in turn, (3.3)) large. Hence the sum can be chosen small enough that the following conditions, necessary for repeated application of Lemma 3, are met.

$$\rho_i = \rho_1 - \sum_{j=1}^{i} \delta_j > \rho_*$$

$$\mathtt{m}_i = \mathtt{m}_1 - \sum_{j=1}^{i} \delta_j > \mathtt{m}_*$$

$$N_i = \frac{K}{\delta_0^{2\tau+2}} + \sum_{j=1}^{i} \delta_j < 1 + \frac{K}{\delta_0^{2\tau+2}}$$

$$M_i = \frac{K(1+\delta_0)}{\delta_0} + \sum_{j=1}^{i} K\delta_j < \frac{2K}{\delta_0}$$

Checking (5.14), we see by (5.26)

$$\delta_i^{4\tau+2} = \frac{2^{8\tau+4}K2^{-(4\tau+2)i}\eta_1}{c^8\epsilon^{4\alpha-3}} = \frac{2^{4\tau+2}K\eta_i}{c^8\epsilon^{4\alpha-3}}$$

Solving back for η_i,

$$\eta_i = \frac{c^8\epsilon^{4\alpha-3}\delta_i^{4\tau+2}}{2^{4\tau+2}K} < \frac{c^6\epsilon^{2\alpha-1}\delta_i^{2\tau+3}}{4K}$$

Once assured that Lemma 3 can be applied iteratively, we can construct the sequence of canonical mappings, $\{\psi_i\}$, where

$$\psi_i = \phi_1 \circ \phi_2 \circ \ldots \circ \phi_i$$

This is a sequence of bounded, analytic, canonical maps, so by Montel's theorem there is a convergent subsequence. (Actually, the sequence can be shown to be Cauchy, but this is unnecessary for the proof.) The limiting map,

$$\Psi_\infty : D_{\rho_*} \to D_{\rho_0}$$

is clearly the canonical change of coordinates sought.

To show (4.9), it is sufficient to prove

$$\|\psi_i - \breve{\psi}_i\|^{\cap}_{\rho_i} < \frac{K\epsilon_0}{c^2}\|\Delta p\|$$

independently of i.

$$\begin{aligned}
\|\psi_i - \breve{\psi}_i\|^{\cap}_{\rho_i} &\leq \|(\phi_1 \circ \ldots \circ \phi_{i-1} \circ \phi_i) - (\phi_1 \circ \ldots \phi_{i-1} \circ \breve{\phi}_i)\|^{\cap}_{\rho_i} \\
&\quad + \|(\phi_1 \ldots \circ \phi_{i-1} \circ \breve{\phi}_i) - (\breve{\phi}_1 \circ \ldots \breve{\phi}_{i-1} \circ \breve{\phi}_i)\|^{\cap}_{\rho_i} \\
&\leq \|\nabla \psi_{i-1}\|_{\rho_i - 3\delta_i/4} \|\phi_i - \breve{\phi}_i\|^{\cap}_{\rho_i} + \|\psi_{i-1} - \breve{\psi}_{i-1}\|^{\cap}_{\rho_{i-1}} \\
&\leq K\epsilon\delta_i \|\Delta p\| + \|\psi_{i-1} - \breve{\psi}_{i-1}\|^{\cap}_{\rho_{i-1}}
\end{aligned}$$

Now, by Lemma 2,

$$\|\psi_1 - \breve{\psi}_1\|^{\cap}_{\rho_1} \quad = \|\phi_1 - \breve{\phi}_1\|^{\cap}_{\rho_1} \quad \leq \frac{K\epsilon_0}{c^2} \|\Delta p\|$$

By induction,

$$\begin{aligned}
\|\psi_i - \breve{\psi}_i\|^{\cap}_{\rho_i} &\leq \frac{K\epsilon_0}{c^2} \|\Delta p\| + \sum_{j=1}^{i} K\delta_j \epsilon \|\Delta p\| \\
&\leq \frac{K\epsilon_0}{c^2} \|\Delta p\|
\end{aligned}$$

So the theorem is proven. □

6 Proofs of Lemmae 2 and 3

To prove these lemmae we will need to control solutions to equations of the form

$$\sum_{i=1}^{n} \lambda_i \frac{\partial f}{\partial q_i} = g \tag{6.1}$$

The following lemma is obtained by a slight modification of a well known argument (see, for instance, [12] (pg 221) whose argument we follow closely).

Lemma 4 Given $\omega, \tilde{\omega} \in \mathbb{R}^n$ *so that* $\forall k \notin \mathcal{K} \subset \mathbb{Z}^n$

$$|k \cdot \omega| \geq \frac{c}{|k|^s} \qquad |k \cdot \tilde{\omega}| \geq \frac{c}{|k|^s} \tag{6.2}$$

and given a real analytic function, $g : D_\rho \to \mathbb{C}$ satisfying $\forall k \in \mathcal{K}$

$$\int_0^{2\pi} \cdots \int_0^{2\pi} g(p,q) e^{-ik\cdot q} dq = 0 \tag{6.3}$$

then

i) There exist functions, $f, \tilde{f}$, unique up to the addition of a function of p, which satisfy:

$$\sum_{i=1}^{n} \omega_i \frac{\partial f}{\partial q_i} = g \qquad \sum_{i=1}^{n} \tilde{\omega}_i \frac{\partial \tilde{f}}{\partial q_i} = g \tag{6.4}$$

$$\|f\|_{\rho-\delta} \leq \frac{K\|g\|_\rho}{c\delta^{n+s}} \qquad \|\tilde{f}\|_{\rho-\delta} \leq \frac{K\|g\|_\rho}{c\delta^{n+s}} \tag{6.5}$$

$$\left\|\frac{\partial f}{\partial q_i}\right\|_{\rho-\delta} \leq \frac{K\|g\|_\rho}{c\delta^{n+s+1}} \qquad \left\|\frac{\partial \tilde{f}}{\partial q_i}\right\|_{\rho-\delta} \leq \frac{K\|g\|_\rho}{c\delta^{n+s+1}} \tag{6.6}$$

ii) The difference of these functions is proportional to the difference of the frequencies

$$\left\|\frac{\partial f}{\partial q_i} - \frac{\partial \tilde{f}}{\partial q_i}\right\|_{\rho-\delta} \leq \frac{K\|g\|_\rho \|\omega - \tilde{\omega}\|}{c\delta^{n+2s+2}} \tag{6.7}$$

The proof relies on three well known facts.

i) If f is analytic in D_ρ and $\hat{f}_k$ is its k-th Fourier coefficient,

$$|\hat{f}_k| \leq \|f\|_\rho \exp(-|k|\rho) \tag{6.8}$$

ii) If $|\hat{f}_k| \leq C \exp(-|k|\rho)$ for all $k \in \mathbb{Z}^n$ then

$$\|f\|_{\rho-\delta} \leq C \left(\frac{4}{\delta}\right)^n \tag{6.9}$$

iii) For $C, s, \delta > 0$ we have

$$C^s \leq \left(\frac{s}{e\delta}\right)^s \exp(C\delta) \tag{6.10}$$

(6.4) is equivalent the equation of Fourier coefficients, $(\omega \cdot k)\hat{f}_k = \hat{g}_k$. Thus the solution, f, will have Fourier coefficients

$$\hat{f}_k = \begin{cases} \frac{\hat{g}_k}{\omega\cdot k} & k \notin \mathcal{K} \\ 0 & k \in \mathcal{K} \end{cases}$$

since $\hat{g}_k = 0$ for $k \in \mathcal{K}$. Likewise for $\tilde{f}$. So too we have the uniqueness result since only the average, $\hat{f}_0$, is not determined by the formula above.

Using (6.8) and (6.10), letting $C = |k|$ and $\delta = \delta/2$,

$$\begin{aligned} |\hat{f}_k| = \frac{|\hat{g}_k|}{|\omega \cdot k|} &\leq \frac{|k|^s}{c} e^{-|k|\rho} \|g\|_\rho \\ &\leq \frac{\|g\|_\rho}{c} \left(\frac{2s}{e\delta}\right)^s e^{-|k|(\rho-\delta/2)} \end{aligned}$$

So, by (6.9),

$$\|f\|_{(\rho-\delta/2)-\delta/2} \leq \frac{\|g\|_\rho}{c} \left(\frac{2s}{e\delta}\right)^s \left(\frac{8}{\delta}\right)^n \leq \frac{K\|g\|_\rho}{c\delta^{n+s}}$$

The Fourier coefficients of $\frac{\partial f}{\partial q_i}$ are $k_i \hat{f}_k$, so

$$|k_i \hat{f}_k| \leq \frac{|k|^{s+1}}{c} e^{-|k|\rho} \|g\|_\rho$$

and again by (6.9) and (6.10),

$$\left\| \frac{\partial f}{\partial q_i} \right\|_{\rho-\delta} \leq \frac{\|g\|_\rho}{c} \left(\frac{2(s+1)}{e\delta}\right)^{s+1} \left(\frac{8}{\delta}\right)^n \leq \frac{K\|g\|_\rho}{c\delta^{n+s+1}}$$

The argument is identical for $\tilde{f}$.

The Fourier coefficients of $\frac{\partial}{\partial q_i}(f - \tilde{f})$ are bounded as follows

$$\begin{aligned} |k_i(\hat{f}_k - \hat{\tilde{f}}_k)| &\leq |k_i| \left| \frac{\hat{g}_k}{\omega \cdot k} - \frac{\hat{g}_k}{\tilde{\omega} \cdot k} \right| \\ &\leq \frac{|\hat{g}_k||k_i|}{|\omega \cdot k||\tilde{\omega} \cdot k|} |k \cdot (\tilde{\omega} - \omega)| \\ &\leq \frac{|\hat{g}_k||k|^{2s+2}}{c} \|\tilde{\omega} - \omega)\| \\ &\leq \frac{\|g\|_\rho e^{-|k|\rho} |k|^{2s+2}}{c} \|(\tilde{\omega} - \omega)\| \end{aligned}$$

So another application of (6.9) and (6.10) completes the proof. $\square$

In this and the following proof, we will construct canonical changes of coordinates near the identity which will make the torus $p = 0$ a closer approximation to an invariant torus. Canonical transformations will be generated using the Lie method as described in [12]. Using this method the canonical transformation is the time one flow, ϕ, along a Hamiltonian vector field. It is well-known that such flows preserve the symplectic form. The vector field will in turn be generated by a real analytic Hamiltonian, $\chi : D_\rho \to \mathbb{C}$, of the form

$$\chi(p, q) = X(q) + Y(q) \cdot p$$

The transformed Hamiltonian, H, according to Taylor's theorem then has the form

$$H \circ \phi = H + \{\chi, H\} + \frac{1}{2}\{\chi, \{\chi, H\}\} \circ \phi_t \tag{6.11}$$

where $\{f, g\} = \sum_{i=1}^{n} \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i}$ are the Poisson brackets, and ϕ_t is the flow along the same vector field for time $t \le 1$. Since it is our plan to try to eliminate the dependence on q in H up to order p^2, we now expand the Hamiltonian in powers of p, writing

$$\begin{aligned} C_v(p_v) &= \frac{1}{2}\sum_{i,j \le l} E^v_{ij} p_i p_j + F^v(p_v) \\ R(p, q) &= \frac{1}{2}\sum_{i,j \le n} E_{ij}(q) p_i p_j + F(p, q) \end{aligned}$$

where F^v and F are $O(\|p\|^3)$ and the average of A is assumed to be zero (this constant term does not effect the flow generated by H and so can be ignored). Also by (4.3) $\bar{B}_i = 0$. Substituting,

$$\begin{aligned} H \circ \phi = {} & \omega_v \cdot p_v + \omega_f \cdot p_f + \omega_s \cdot p_s \\ & + C_v + \frac{1}{2}Ep \cdot p + F + \{Y \cdot p, \frac{1}{2}(E^v + E)p \cdot p\} + \{\chi, F^v + F\} && (6.12) \\ & + \Big[\{X, \omega \cdot p\} + \tilde{A}\Big] + \Big[\{Y \cdot p, \omega \cdot p\} + \tilde{B} \cdot p + \{X, \frac{1}{2}E^v p \cdot p\}\Big] && (6.13) \\ & + \{\chi, A + B \cdot p\} + \{X, \frac{1}{2}Ep \cdot p\} + \check{A} + \check{B} \cdot p + \frac{1}{2}\{\chi, \{\chi, H\}\} \circ \phi_t && (6.14) \end{aligned}$$

The terms in (6.12) are terms of order $\|p\|^2$, and therefore merely contribute to R^+. We need only show that they change very little, and therefore do not adversely effect later iterations. This will be done later.

It is our object to choose X and Y so that (6.13) is zero. Then all dependence on q of order less than p^2 will be contained in terms on line (6.14) which we will show are small.

We choose a function X canceling $\tilde{A}$ by solving the equation

$$\sum_{i=1}^{n} \omega_i \frac{\partial X}{\partial q_i} = \tilde{A} \tag{6.15}$$

Since $\tilde{A}$ has average zero over the "fast" and "varying" angles Lemma 4 can be applied with $\mathcal{K} = \{0\}^{l+m} \times \mathbb{Z}^{n-l-m}$. Therefore there exists an X satisfying (6.15) with bounds (dropping the subscripts from ϵ_0, ρ_0, and δ_0)

$$\begin{aligned} \|X\|_{\rho-\delta/4} &\leq \frac{K\|\tilde{A}\|_{\rho}}{c(\delta/4)^{\tau-1}} \leq \frac{K\epsilon}{c\delta^{\tau-1}} \\ \left\|\frac{\partial X}{\partial q_i}\right\|_{\rho-\delta/4} &\leq \frac{K\epsilon}{c\delta^{\tau}} \end{aligned}$$

Likewise, there exists a vector valued function Y which solves the equation

$$\left[\sum_{i=1}^{n} \omega_i \frac{\partial Y}{\partial q_i}\right] \cdot p = \left[\tilde{B} - E^v \frac{\partial X}{\partial q}\right] \cdot p \tag{6.16}$$

Once again applying Lemma 4,

$$\begin{aligned} \|Y\|_{\rho-\delta/2} &\leq \frac{K}{c(\delta/4)^{\tau-1}} \left[\|B\|_{\rho-\delta/4} + K \left\|\frac{\partial X}{\partial q_i}\right\|_{\rho-\delta/4} \|E^v\|\right] \\ &\leq \frac{K\epsilon}{c^2\delta^{2\tau-1}} \\ \left\|\frac{\partial Y}{\partial q_i}\right\|_{\rho-\delta/2} &\leq \frac{K\epsilon}{c^2\delta^{2\tau}} \end{aligned}$$

Combining these results with (5.1),

$$\|\nabla\chi\|_{\rho-\delta/2} \leq \frac{K\epsilon}{c^2\delta^{2\tau}} \leq \frac{\delta}{4}$$

So, by the existence theorem for ordinary differential equations,

$$\|\phi - \mathrm{Id}\|_{\rho-\delta/2} \leq \frac{\delta}{4}$$

Which implies that

$$\phi : D_{\rho-\delta} \to D_{\rho-3\delta/4}$$

and

$$\phi(p,q) = (p,q) + g(p,q)$$

where $\|g\|_{\rho-\delta} \leq \frac{K\epsilon}{c^2\delta^{2\tau}}$.

Remark: In Theorem 1 we discuss the situation where there are no "varying" variables. The estimates above can clearly be improved in that case since then $E^v = 0$. These improvements do not, however, seem to improve the measure estimates in Theorem 1. They would certainly reduce the size of many of the constants, K which appear, but this is not particularly helpful since this paper does not attempt to estimate these constants.

Continuing with the proof of Lemma 2 we address the uniformity results (5.8) and (5.9). We begin by letting Z be the average zero solution to

$$\breve{\omega} \cdot \frac{\partial Z}{\partial q} = A \tag{6.17}$$

Note that $Z - \breve{X}$ is a solution to

$$\sum_{i=1}^{n} \breve{\omega}_i \frac{\partial}{\partial q_i}(Z - \breve{X}) = \tilde{A} - \tilde{\breve{A}}$$

By Lemma 4 this is the unique average zero solution. Then using Lemma 4 (i) and (ii) we have

$$\begin{aligned}
\left\| \frac{\partial X}{\partial q_i} - \frac{\partial \breve{X}}{\partial q_i} \right\|^{\cap}_{\rho-\delta/4} &\leq \left\| \frac{\partial X}{\partial q_i} - \frac{\partial Z}{\partial q_i} \right\|^{\cap}_{\rho-\delta/4} + \left\| \frac{\partial Z}{\partial q_i} - \frac{\partial \breve{X}}{\partial q_i} \right\|^{\cap}_{\rho-\delta/4} \\
&\leq \frac{K\|\tilde{A}\|_\rho \|\omega - \breve{\omega}\|}{c\delta^\tau} + \frac{K\|\tilde{A} - \tilde{\breve{A}}\|^{\cap}_\rho}{c\delta^\tau} \qquad (6.18)
\end{aligned}$$

Now, since $\breve{\omega} = \frac{\partial \bar{H}}{\partial p}(\Delta p)$

$$\|\omega - \breve{\omega}\| \leq \frac{K}{\rho_0^2}\|\Delta p\|$$

$\breve{A}$, can be written

$$\breve{A}(q) = H(\Delta p, q) - \bar{H}(\Delta p) = R(\Delta p, q) - \bar{R}(\Delta p)$$

So

$$\|\tilde{A} - \tilde{\breve{A}}\|_\rho \leq \|A - \breve{A}\|_\rho \leq \frac{K\epsilon}{\rho}\|\Delta p\|$$

Substituting these results into (6.18) gives

$$\left\| \frac{\partial X}{\partial q_i} - \frac{\partial \breve{X}}{\partial q_i} \right\|^\cap_{\rho-\delta/4} \leq \frac{K\epsilon}{c\delta^\tau}\|\Delta p\|$$

Similarly,

$$\left\| \frac{\partial Y}{\partial q_i} - \frac{\partial \breve{Y}}{\partial q_i} \right\|^\cap_{\rho-\delta/2} \leq \frac{K\epsilon\|\Delta p\|}{c^2\delta^{2\tau}}$$

Combining these results we have

$$\|\nabla\chi - \nabla\breve{\chi}\|^\cap_{\rho-\delta/2} \leq \frac{K\epsilon\|\Delta p\|}{c^2\delta^{2\tau}}$$

The theory of ordinary differential equations gives us, then,

$$\begin{aligned} \|\phi - \breve{\phi}\|^\cap_{\rho-\delta} &\leq [1 - \|\nabla^2\chi\|^\cap_{\rho-\delta/2}]^{-1}\, \|\nabla\chi - \nabla\breve{\chi}\|^\cap_{\rho-\delta/2} \\ &\leq [1-\delta]^{-1}\frac{K\epsilon\|\Delta p\|}{c^2\delta^{2\tau'}} \\ &\leq \frac{K\epsilon\|\Delta p\|}{c^2\delta^{2\tau}} \leq \frac{\delta^2\|\Delta p\|}{4} \end{aligned} \tag{6.19}$$

which is what was claimed. Using this we can easily prove (5.8).

Now, to bound the size of A^+ and B^+ we need to bound the only terms which may still have a dependence on q which is not $O(\|p\|^2)$. These are the terms in (6.14). This can be done with the simple application of Cauchy's inequality.

$$\begin{aligned} \|\{\chi, A + B\cdot p\}\|_{\rho-\delta} &\leq \frac{K}{\delta}\|\nabla\chi\|_{\rho-\delta/2}\|A + B\cdot p\|_\rho &&\leq \frac{K\epsilon^2}{c^2\delta^{2\tau+1}} \\ \|\{\chi, \frac{1}{2}Ep\cdot p\}\|_{\rho-\delta} &\leq \frac{K}{\delta}\|\nabla\chi\|_{\rho-\delta/2}\|E\|_\rho\rho &&\leq \frac{K\epsilon^2}{c^2\delta^{2\tau+1}} \\ \|\{\chi,\{\chi, H\}\}\circ\phi\|_{\rho-\delta} &\leq \|\{\chi,\{\chi,H\}\}\|_{\rho-3\delta/4} && \\ &\leq \frac{K}{(\delta/4)^2}\|\nabla\chi\|^2_{\rho-\delta/2}\|H\|_\rho &&\leq \frac{K\epsilon^2}{c^4\delta^{4\tau+2}} \end{aligned} \tag{6.20}$$

These inequalities, combined with (4.4), give us (5.3).

To control the nondegeneracy condition (5.4) we need show first that the change in the $O(\|p\|^2)$ terms in the averaged Hamiltonian is small.

$$H \circ \phi - H = \{\chi, C_v\} + \{\chi, A + B \cdot p + R\} + \frac{1}{2}\{\chi, \{\chi, H\}\} \circ \phi_t$$

Since $\overline{\{\chi, C_v\}} = 0$,

$$\frac{\partial^2 [\overline{H \circ \phi - H}]}{\partial p_i \partial p_j} = \frac{\partial}{\partial p_i \partial p_j}[\overline{\{\chi, R\}} + \frac{1}{2}\overline{\{\chi, \{\chi, H\}\} \circ \phi_t}]$$

Setting $p = 0$, we can use Cauchy's inequality and (6.20) to bound this quantity.

$$\left\| \frac{\partial^2 [\overline{H \circ \phi - H}]}{\partial p_i \partial p_j} \right\| \leq \frac{K\epsilon^2}{\rho^2 c^4 \delta^{4\tau+2}}$$

With this in hand we may address (5.4)

$$\begin{aligned} \left\| \left[\frac{\partial^2 \overline{H \circ \phi}}{\partial p_i \partial p_j} \right] v \right\| &\geq \left\| \left[\frac{\partial^2 \overline{H}}{\partial p_i \partial p_j} \right] v \right\| - \left\| \left[\frac{\partial^2 \overline{H \circ \phi - H}}{\partial p_i \partial p_j} \right] v \right\| \\ &\geq \begin{cases} \mathtt{m}\|v\| - \frac{K\epsilon^2}{c^4\delta^{4\tau+2}}\|v\| & \forall v \in \mathbb{R}^l \times \{0\}^{n-l} \\ \mathtt{m}\epsilon\|v\| - \frac{K\epsilon^2}{c^4\delta^{4\tau+2}}\|v\| & \forall v \in \mathbb{R}^n \end{cases} \end{aligned}$$

which, with (5.1), implies (5.4).

For (5.5), from (6.12) and (6.14) we have

$$R^+ - R = \{Y \cdot p, \frac{1}{2}(E^v + E)p \cdot p\} + \{\chi, F^v + F\} + G$$

where G is the $O(\|p\|^2)$ part of $\frac{1}{2}\{\chi, \{\chi, H\}\} \circ \phi_t$. This gives the bound

$$\begin{aligned} \|R^+ - R\|_{\rho-\delta} &\leq \frac{K\rho^2}{\delta^2}(\|Y\|_{\rho-\delta/2}\|E^v + E\|_\rho + \|\chi\|_{\rho-\delta/2}\|F^v + F\|_\rho) + \|G\|_{\rho-\delta} \\ &\leq \frac{K\epsilon}{c^2\delta^{2\tau+2}} \end{aligned}$$

So

$$\begin{aligned} \|R^+\|_{\rho-\delta} &\leq \|R\|_\rho + \|R^+ - R\|_{\rho-\delta} \\ &\leq \left[\frac{K}{c^2\delta^{2\tau+2}} \right] \epsilon \end{aligned}$$

All that remains is to show (5.7). A^+ is just the terms in (6.14) evaluated at $p = 0$, and likewise B^+ is the derivitive of these terms with respect to p at $p = 0$. It is sufficient, then, to show that the terms in (6.14) minus the corresponding terms for $\breve{H}$ satisfy the right hand side of (5.7). This can be accomplished–albeit tediously–using bounds already calculated. We show this explicitly only for two terms.

$$
\begin{aligned}
&\|\{\chi, A + B\cdot p\} - \{\breve{\chi}, \breve{A} + \breve{B}\cdot p\}\|^{\cap}_{\rho-\delta} \\
&\quad \le \|\{\chi - \breve{\chi}, A + B\cdot p\}\|^{\cap}_{\rho-\delta/2} + \|\{\breve{\chi}, A - \breve{A} + (B - \breve{B})\cdot p\}\|^{\cap}_{\rho-\delta/2} \\
&\quad \le \|\nabla(\chi - \breve{\chi})\|^{\cap}_{\rho-\delta/2}\frac{\|A + B\cdot p\|_{\rho}}{\delta} + \frac{\|\nabla\breve{\chi}\|_{\rho}}{\delta}\left[\|A - \breve{A}\|^{\cap}_{\rho} + \rho\|B - \breve{B}\|^{\cap}_{\rho}\right] \\
&\quad \le \frac{K\epsilon^2\|\Delta p\|}{c^2\delta^{2\tau+1}}
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&\|\{\chi, \{\chi, H\}\}\circ\phi - \{\breve{\chi}, \{\breve{\chi}, \breve{H}\}\}\circ\breve{\phi}\|^{\cap}_{\rho-\delta} \\
&\quad \le \|\nabla\{\chi, \{\chi, H\}\}\|_{\rho-3\delta/4}\|\phi - \breve{\phi}\|^{\cap}_{\rho-\delta} + \|\{\chi - \breve{\chi}, \{\chi, H\}\}\|^{\cap}_{\rho-3\delta/4} \\
&\qquad\quad + \|\{\breve{\chi}, \{\chi - \breve{\chi}, H\}\}\|^{\cap}_{\rho-3\delta/4} + \|\{\breve{\chi}, \{\breve{\chi}, H - \breve{H}\}\}\|^{\cap}_{\rho-3\delta/4} \\
&\quad \le \frac{K\epsilon^2\|\Delta p\|}{c^4\delta^{4\tau+2}}
\end{aligned}
$$

Hence we have (5.7). □

The proof of Lemma 3 is similar, but complicated by the need to deal with terms in the perturbation whose average over the fast and varying angles is not zero (these terms could be ignored in the previous lemma since they were small by hypothesis). Again, the canonical transformation is the time one flow according to a Hamiltonian of the form:

$$
\begin{aligned}
\chi(p, q) &= \chi_1(p, q) + \chi_2(p, q_s) + \xi_v \oplus \xi_f \oplus \xi_s \cdot q \\
&= X_1(q) + Y_1(q)\cdot p + X_2(q_s) + Y_2(q_s)\cdot p + \xi_v \oplus \xi_f \oplus \xi_s \cdot q
\end{aligned}
$$

Here ξ is a constant vector, so χ is not actually a periodic function in q. Nevertheless the flow is well-defined, so we ignore this point. χ_1 will be chosen to remove the part of the perturbation which has zero average over the fast and varying angles, while χ_2 is chosen to handle the remainder. To facilitate this we decompose each term into three components which satisfy averaging conditions (as described in section 2). We will have a much weaker bound on χ_2, however since $\{\chi_2, f(p_v, p_f)\} = 0$ for any function f, the error terms involving χ_2 are still under control. As before ϕ_t is the Hamiltonian flow induced by χ. Expanding in powers of t,

$$H \circ \phi_t = H + t\{\chi, H\} + \frac{t^2}{2}\{\chi, \{\chi, H\}\} \circ \phi_{\bar{t}} \tag{6.21}$$

Now, following the notation and procedure of the previous lemma (that is, let $t = 1$, $\phi = \phi_1$, and expand in powers of p), noting that

$$\overline{\{X_1, E^v\}} = \widecheck{\{X_1, E^v\}} = 0$$

we have

$$
\begin{aligned}
H \circ \phi = {} & \omega_v \cdot p_v + \omega_f \cdot p_f + \omega_s \cdot p_s && (6.22)\\
& + (E^v + \bar{E})\xi \cdot p + \bar{B} \cdot p + \overline{\{X_1 + X_2, Ep \cdot p/2\}} && (6.23)\\
& + [\{X_1, \omega \cdot p\} + \tilde{A}] + [\{X_2, \omega_s \cdot p_s\} + \check{A}] && (6.24)\\
& + \{Y_1 \cdot p, \omega \cdot p\} + \tilde{E}\xi \cdot p + \tilde{B} \cdot p \\
& + \frac{1}{2}\widetilde{\{X_1, (E^v + E)p \cdot p\} + \{X_2, Ep \cdot p\}} && (6.25)\\
& + \{Y_2 \cdot p, \omega_s \cdot p_s\} + \check{E}\xi \cdot p + \check{B} \cdot p + \frac{1}{2}\widecheck{\{X_1, Ep \cdot p\} + \{X_2, Ep \cdot p\}} && (6.26)\\
& + C_v + R + \{(Y_1 + Y_2) \cdot p, (E^v + E)p \cdot p/2\} + \{\chi, F^v + F\} && (6.27)\\
& + \{\chi, A + B \cdot p\} + \frac{1}{2}\{\chi, \{\chi, H\}\} \circ \phi_{\bar{t}} && (6.28)
\end{aligned}
$$

where $0 < \bar{t} \leq 1$. As before we will chose χ so that (6.24), (6.25), (6.26) as well as (6.23), are set to zero. Beginning with (6.24), we chose X_1 by applying Lemma 4 to

$$\sum_{i=1}^{n} \omega_i \frac{\partial X_1}{\partial q_i} = \tilde{A}$$

with $\mathcal{K} = \{0\}^{l+m} \times \mathbb{Z}^{n-l-m}$. Then

$$\|X_1\|_{\rho-\delta/4} \leq \tfrac{K\eta}{c\delta^{\tau-1}} \qquad \left\|\tfrac{\partial X_1}{\partial q_i}\right\|_{\rho-\delta/4} \leq \tfrac{K\eta}{c\delta^{\tau}} \tag{6.29}$$

Similarly we chose X_2 by applying Lemma 4 to

$$\sum_{i=n-l-m}^{n} \omega_i \frac{\partial X_2}{\partial q_i}(q_s) = \check{A}(q_s)$$

with $\mathcal{K} = \{0\}^{n-l-m}$. Then

$$\|X_2\|_{\rho-\delta/4} \leq \tfrac{K\eta}{c\epsilon^{\alpha}\delta^{\tau-1}} \qquad \left\|\tfrac{\partial X_2}{\partial q_i}\right\|_{\rho-\delta/4} \leq \tfrac{K\eta}{c\epsilon^{\alpha}\delta^{\tau}} \tag{6.30}$$

The terms in (6.23) have no angular dependence and are therefore not obviously threatening. However these terms produce a shift in the frequencies, ω. These shifted frequencies–if the shift is not carefully chosen–would not necessarily satisfy a nonresonance condition like (5.10), thus destroying our ability to repeatedly apply Lemma 3. Therefore we shift the action variables, p, by a vector, ξ, chosen precisely to set (6.23) to zero. It is in this step that the degeneracy of H is most troublesome. Variations in p_s produce inconveniently small variations in ω_s. Solving for the necessary ξ,

$$\left[(E^v + \bar{E})\xi + \bar{B} + \overline{E\left(\frac{\partial X_1}{\partial q_i} + \frac{\partial X_2}{\partial q_i}\right)}\right] \cdot p = 0$$

Then using (5.13), (6.29), and (6.30) we bound ξ as

$$\begin{aligned}
\|\xi\| &\leq \frac{\|(E^v + \bar{E})\xi\|}{c^2 \mathtt{m}\epsilon} \\
&\leq \frac{1}{c^2\mathtt{m}\epsilon}\left[\|B\|_\rho + \|E\|_\rho\left(\left\|\frac{\partial X_1}{\partial q}\right\|_{\rho-\delta/4} + \left\|\frac{\partial X_2}{\partial q_i}\right\|_{\rho-\delta/4}\right)\right] \\
&\leq \frac{1}{c^2\mathtt{m}\epsilon}\left[\eta + \frac{K\eta}{c\epsilon^{\alpha-1}\delta^{s+m+1}}\right] \\
&\leq \frac{K\eta}{c^3\epsilon^{\alpha}\delta^{\tau}}
\end{aligned} \tag{6.31}$$

It is possible to estimate a part of ξ more closely by noticing that $E^v(0 \oplus \xi_f \oplus \xi_s) = 0$, so

$$\begin{aligned}
\|\xi_v\| &\leq \frac{\|(E^v + \bar{E})(\xi_v \oplus 0 \oplus 0)\|}{\mathtt{m}} \\
&\leq \frac{\|(E^v + \bar{E})\xi\|}{\mathtt{m}} + \frac{\|\bar{E}(0 \oplus \xi_f \oplus \xi_s)\|}{\mathtt{m}} \\
&\leq \frac{K\eta}{c^3\epsilon^{\alpha-1}\delta^{\tau}}
\end{aligned} \tag{6.32}$$

The choices of Y_1 and Y_2 setting (6.25) and (6.26) to zero proceed similarly to the choices of X_1 and X_2. Applying Lemma 4 to the following equation yields Y_1.

$$\left[\sum_{i=1}^{n} \omega_i \frac{\partial Y_1}{\partial q_i}\right] \cdot p = \left[\tilde{E}\xi + \tilde{B} + \widetilde{\overbrace{(E^v + E)\frac{\partial X_1}{\partial q}}} + E\widetilde{\frac{\partial X_2}{\partial q}}\right] \cdot p$$

Thus, using (6.29), (6.30), and (6.31),

$$\begin{aligned}
\|Y_1\|_{\rho-\delta/2} &\leq \frac{K}{c\delta^{\tau-1}} \left[\|B\|_\rho + \|E\|_\rho \|\xi\| + \|E^v + E\|_\rho \left\| \frac{\partial X_1}{\partial q} \right\|_{\rho-\delta/4} \right. \\
&\qquad \left. + \|E\|_\rho \left\| \frac{\partial X_2}{\partial q_i} \right\|_{\rho-\delta/4} \right] \\
&\leq \frac{K\eta}{c^4 \epsilon^{\alpha-1} \delta^{2\tau-1}} && (6.33) \\
\left\| \frac{\partial Y_1}{\partial q_i} \right\|_{\rho-\delta/2} &\leq \frac{K\eta}{c^4 \epsilon^{\alpha-1} \delta^{2\tau}} && (6.34)
\end{aligned}$$

Proceding in the same way gives

$$\begin{aligned}
\|Y_2\|_{\rho-\delta/2} &\leq \frac{K\eta}{c^4 \epsilon^{2\alpha-1} \delta^{2\tau-1}} && (6.35) \\
\left\| \frac{\partial Y_2}{\partial q_i} \right\|_{\rho-\delta/2} &\leq \frac{K\eta}{c^4 \epsilon^{2\alpha-1} \delta^{2\tau}} && (6.36)
\end{aligned}$$

¿From these estimates we have

$$\begin{aligned}
\|\nabla \chi_1\|_{\rho-\delta/2} &\leq \frac{K\eta}{c^4 \epsilon^{\alpha-1} \delta^{2\tau}} && (6.37) \\
\|\nabla \chi\|_{\rho-\delta/2} &\leq \frac{K\eta}{c^4 \epsilon^{2\alpha-1} \delta^{2\tau}} && (6.38)
\end{aligned}$$

Now, by the bound on the size of η, (5.14),

$$\|\nabla \chi\|_{\rho-\delta/2} \leq \frac{\delta}{4}$$

So,

$$\phi : D_{\rho-\delta} \to D_{\rho-3\delta/4} \tag{6.39}$$

Proceeding to the main estimate (5.16), note first that the terms in (6.27) are all $O(\|p\|^2)$, and so only contribute to R^+. As before, the only terms which have an angular dependence less than $O(\|p\|^2)$ are in (6.28). The first term is easily bounded,

$$\begin{aligned} \|\{\chi, A + B\cdot p\}\|_{\rho-\delta} &\leq K\|\nabla\chi\|_{\rho-\delta}\frac{\|A + B\cdot p\|_\rho}{\delta} \\ &\leq \frac{K\eta^2}{c^4\epsilon^{2\alpha-1}\delta^{2\tau+1}} \end{aligned}$$

The other term in (6.28) must be treated with more care. Proceeding naively, one would arrive at the bound

$$\begin{aligned} \|\{\chi, \{\chi, H\}\}\|_{\rho-\delta} &\leq \frac{K\|\nabla\chi\|^2_{\rho-\delta}\|H\|_\rho}{\delta^2} \\ &\leq \frac{K\eta^2}{c^8\epsilon^{4\alpha-2}\delta^{4\tau+2}} \end{aligned}$$

which is not strong enough. This can be seen by letting $\eta = \epsilon^2$ which is typical in the second step. The error after the second step would then be on the order of $\epsilon^{6-4\alpha}$, which is no improvement.

Consider first $\{\chi, H\}$. Let $h = H - \omega_v\cdot p_v - \omega_f\cdot p_f - C_v$. Then, ignoring the constant $\{\xi\cdot q, \omega_v\cdot p_v + \omega_f\cdot p_f\}$ and noticing that $\{\chi_2, \omega_v\cdot p_v + \omega_f\cdot p_f + C_v\} = 0$,

$$\{\chi, H\} = \{\chi_1, \omega_v\cdot p_v + \omega_f\cdot p_f + C_v\} + \frac{\partial C_v}{\partial p}\xi_v + \{\chi, h\}$$

Now, using (6.32), (6.37), and (6.38), since $\|h\|_\rho \leq K\epsilon$

$$\begin{aligned} \|\{\chi, H\}\|_{\rho-\delta/2} &\leq K\|\nabla\chi_1\|_{\rho-\delta/2}\frac{\|\omega_v\cdot p_v + \omega_f\cdot p_f + C_v\|}{\delta} \\ &\quad + \frac{K\|C_v\|_\rho}{\delta}\|\xi_v\| + K\|\nabla\chi\|_{\rho-\delta/2}\frac{\|h\|_\rho}{\delta} \\ &\leq \frac{K\eta}{c^4\epsilon^{\alpha-1}\delta^{2\tau+1}} + \frac{K\eta}{c^3\epsilon^{\alpha-1}\delta^{\tau+1}} + \frac{K\eta}{c^4\epsilon^{2\alpha-2}\delta^{2\tau+1}} \\ &\leq \frac{K\eta}{c^4\epsilon^{2\alpha-2}\delta^{2\tau+1}} \end{aligned} \tag{6.40}$$

The second term in (6.28) can now be bounded

$$\begin{aligned}
\|\{\chi,\{\chi,H\}\}\circ\phi_t\|_{\rho-\delta} &\le \|\{\chi,\{\chi,H\}\}\|_{\rho-3\delta/4} \\
&\le K\|\nabla\chi\|_{\rho-\delta/2}\frac{\|\{\chi,H\}\|_{\rho-\delta/2}}{\delta} \\
&\le \frac{K\eta^2}{c^8\epsilon^{4\alpha-3}\delta^{4\tau+2}} \qquad (6.41)
\end{aligned}$$

Combining these two results with Cauchy's inequality gives (5.16).

For (5.17) note that from (6.27) and (6.28) we have

$$R^+ - R = \{(Y_1+Y_2)\cdot p, (E^v+E)p\cdot p/2\} + \{\chi, F^v+F\} + G$$

where G is the $O(\|p\|^2)$ part of $\{\chi,\{\chi,H\}\}\circ\phi_t$.

So, applying (5.14),

$$\begin{aligned}
\|R^+ - R\|_{\rho-\delta} \le & \frac{K}{\delta^2}\left[\rho^2\|Y_1+Y_2\|_{\rho-\delta/2}\|E^v+E\|_\rho \right.\\
& \left. + \|\chi\|_{\rho-\delta/2}\|F^v+F\|_\rho\right] + \|G\|_{\rho-\delta} \\
\le & \frac{K\eta}{c^4\epsilon^{2\alpha-2}\delta^{2\tau+1}} \le \delta\epsilon
\end{aligned}$$

Then

$$\|R^+\|_{\rho-\delta} \le \|R\|_\rho + \epsilon\delta \le [N+\delta]\epsilon$$

For (5.18), first we have that

$$\left\|\frac{\partial^2[\overline{H\circ\phi - H}]}{\partial p_i\partial p_j}(0)\right\| \le \frac{\|\{\chi,H\}\|_{\rho-\delta}}{(\rho-\delta)^2} \le c^2\epsilon\delta$$

So, as before,

$$\begin{aligned}
\left\|\left[\frac{\partial^2\overline{H\circ\phi}}{\partial p_i\partial p_j}\right]v\right\| &\ge \left\|\left[\frac{\partial^2\overline{H}}{\partial p_i\partial p_j}\right]v\right\| - \left\|\left[\frac{\partial^2\overline{H\circ\phi - H}}{\partial p_i\partial p_j}\right]v\right\| \\
&\ge \begin{cases} \mathtt{m}\|v\| - c^2\epsilon\delta\|v\| & \forall v\in\mathbb{R}^l\times\{0\}^{n-l} \\ c^2\mathtt{m}\epsilon\|v\| - c^2\epsilon\delta\|v\| & \forall v\in\mathbb{R}^n \end{cases}
\end{aligned}$$

And for (5.19),

$$\begin{aligned}\|g\|_{\rho-\delta} &\leq \|\phi - Id\|_{\rho-\delta} \\ &\leq \|\nabla\chi\|_{\rho-\delta} \leq \frac{K\eta}{c^4\epsilon^{2\alpha-1}\delta^{2\tau}}\end{aligned}$$

Now for the uniformity conclusions, (5.22), (5.23) and (5.24), we need to show that the change in the error terms (6.28) induced by a small change in the actions, Δp, is both small and proportional to Δp.

Let $\breve{\chi} = \breve{X}_1 + \breve{X}_2 + \breve{\xi} \cdot q + (\breve{Y}_1 + \breve{Y}_2) \cdot (p - \Delta p)$, be constructed analogously to χ for $\breve{H} = H(p - \Delta p, q)$. To estimate the difference in the errors arising from the two transformations we will need to estimate $\nabla(\chi - \breve{\chi})$. To that end let X_1^* be the solution to

$$\breve{\omega} \cdot \frac{\partial X_1^*}{\partial q} = \tilde{A}$$

By Lemma 4 (iii), we have

$$\begin{aligned}\left\| \frac{\partial}{\partial q}(X_1^* - X_1) \right\|_{\rho-\delta/4} &\leq \frac{K\|A\|_\rho \|\omega - \breve{\omega}\|}{c\delta^\tau} \\ &\leq \frac{K\eta\|\Delta p\|}{c\delta^\tau}\end{aligned}$$

Since $X_1^* - \breve{X}_1$ is the unique solution with average zero (over the fast and varying angles) to

$$\breve{\omega} \cdot \frac{\partial}{\partial q}(X_1^* - \breve{X}_1) = \tilde{A} - \tilde{\breve{A}}$$

by Lemma 4 (ii), it satisfies the bound

$$\left\| \frac{\partial}{\partial q}(X_1^* - \breve{X}_1) \right\|_{\rho-\delta/4} \leq \frac{K\|A - \breve{A}\|_\rho}{c\delta^\tau} \leq \frac{K\eta\|\Delta p\|}{c\delta^\tau}$$

Combining these two results, we have

$$\|\nabla(X_1 - \breve{X}_1)\|_{\rho-\delta/4} \leq \frac{K\eta\|\Delta p\|}{c\delta^\tau}$$

Similarly,

$$\|\nabla(X_2 - \breve{X}_2)\|_{\rho-\delta/4} \leq \frac{K\eta\|\Delta p\|}{c\epsilon^\alpha\delta^\tau}$$

Comparing the two shifts,

$$\begin{aligned}
\|\xi - \breve{\xi}\| \leq & \frac{\|(E^v + \bar{E})(\xi - \breve{\xi})\|}{c^2\mathtt{m}\epsilon} \\
\leq & \frac{1}{c^2\mathtt{m}\epsilon}\left[\|(E^v - \bar{E})\xi - (\breve{E}^v - \breve{\bar{E}})\breve{\xi}\| + \|(E^v - \breve{E}^v + \bar{E} - \breve{\bar{E}})\breve{\xi}\|\right] \\
\leq & \frac{1}{c^2\mathtt{m}\epsilon}\left[\|B - \breve{B}\|_\rho^\cap + \|\overline{E\frac{\partial}{\partial q}(X_1 - X_2)} - \overline{\breve{E}\frac{\partial}{\partial q}(\breve{X}_1 - \breve{X}_2)}\|_{\rho-\delta}\right. \\
& \left. + \|(E^v - \breve{E}^v + \bar{E} - \breve{\bar{E}})\|\|\breve{\xi}_v\| + \|\bar{E} - \breve{\bar{E}}\|\|\breve{\xi}\|\right] \\
\leq & \frac{K\eta\|\Delta p\|}{c^3\delta^\tau\epsilon^\alpha}
\end{aligned}$$

As before, we can also establish

$$\|\xi_v - \breve{\xi}_v\| \leq \frac{K\eta\|\Delta p\|}{c^3\delta^\tau\epsilon^{\alpha-1}}$$

We estimate $Y - \breve{Y}$ by the same method used for $X - \breve{X}$ above. Let Y_1^* be the average zero solution to

$$\breve{\omega}\cdot\frac{\partial Y_1^*}{\partial q} = \tilde{E}\xi + \tilde{B} + \overbrace{(E^v + E)\frac{\partial(X_1 + X_2)}{\partial q}\cdot p + E\frac{\partial X_2}{\partial q}\cdot p}^{\sim}$$

and proceed as before, establishing

$$\begin{aligned}
\|\nabla(Y_1 - \breve{Y}_1)\|_{\rho-\delta/2}^\cap &\leq \frac{K\eta\|\Delta p\|}{c^4\epsilon^{\alpha-1}\delta^{2\tau}} \\
\|\nabla(Y_2 - \breve{Y}_2)\|_{\rho-\delta/2}^\cap &\leq \frac{K\eta\|\Delta p\|}{c^4\epsilon^{2\alpha-1}\delta^{2\tau}}
\end{aligned}$$

Combining these results, and following arguments given earlier, we have

$$\begin{aligned}
\|\nabla(\chi - \breve{\chi})\|^{\cap}_{\rho-\delta/2} &\leq \frac{K\eta\|\Delta p\|}{c^4\epsilon^{2\alpha-1}\delta^{2\tau}} \\
\|\{(\chi - \breve{\chi}), H\}\|^{\cap}_{\rho-\delta/2} &\leq \frac{K\eta\|\Delta p\|}{c^4\epsilon^{2\alpha-2}\delta^{2\tau+1}} \\
\|\{\breve{\chi}, H - \breve{H}\}\|^{\cap}_{\rho-\delta/2} &\leq \frac{K\eta\|\Delta p\|}{c^4\epsilon^{2\alpha-2}\delta^{2\tau+1}} \\
\|\phi - \breve{\phi}\|^{\cap}_{\rho-\delta/2} &\leq \frac{K\eta\|\Delta p\|}{c^4\epsilon^{2\alpha-1}\delta^{2\tau}} \leq \epsilon\delta^2\|\Delta p\|
\end{aligned}$$

Using the last of these estimates one can easily establish (5.23).

With these estimates in hand we may address (5.22). Again we need only show that the difference between the error terms (6.28) and the corresponding terms for $\breve{H}$ is small and proportional to Δp.

$$\begin{aligned}
&\|\{\chi, A + B\cdot p\} - \{\breve{\chi}, \breve{A} + \breve{B}\cdot(p - \Delta p)\}\|^{\cap}_{\rho-\delta} \\
&\qquad \leq \|\{\chi - \breve{\chi}, A + B\cdot p\}\|^{\cap}_{\rho-\delta} \\
&\qquad\qquad + \|\{\breve{\chi}, A - \breve{A} + (B - \breve{B})\cdot p + \breve{B}\cdot\Delta p\}\|^{\cap}_{\rho-\delta} \\
&\qquad \leq \frac{K\eta^2\|\Delta p\|}{c^4\epsilon^{2\alpha-1}\delta^{4\tau+2}}
\end{aligned}$$

and

$$\begin{aligned}
&\|\{\chi,\{\chi, H\}\}\circ\phi - \{\breve{\chi}, \{\breve{\chi}, \breve{H}\}\}\circ\breve{\phi}\|^{\cap}_{\rho-\delta} \\
&\qquad \leq \|\nabla\{\breve{\chi}, \{\breve{\chi}, \breve{H}\}\}\|^{\cap}_{\rho-\delta}\|\phi - \breve{\phi}\|^{\cap}_{\rho-\delta} + \|\{\chi - \breve{\chi}, \{\chi, H\}\}\circ\breve{\phi}\|^{\cap}_{\rho-\delta} \\
&\qquad\qquad + \|\{\breve{\chi}, \{\chi - \breve{\chi}, H\}\}\circ\breve{\phi}\|^{\cap}_{\rho-\delta} + \|\{\breve{\chi}, \{\breve{\chi}, H - \breve{H}\}\}\circ\breve{\phi}\|^{\cap}_{\rho-\delta} \\
&\qquad \leq \frac{K\eta^2\|\Delta p\|}{c^8\epsilon^{4\alpha-3}\delta^{4\tau+2}}
\end{aligned}$$

establishing (5.22), and completing the proof. □

7 Application to Weinberg's Model

We now apply our main result to a system constructed by Weinberg to test the current, linear theory of quantum mechanics. A full discussion of the physical motivations of this theory is beyond the scope of this paper, however we point out that non-linear modifications of the Schrodinger equation appear naturally as "effective theories" incorporating, for example, radiative effects due to quantum electrodynamics (these are, however, slightly nonlocal). Weinberg's motivation was closely related to this.

The use of this non-linear evolution as a fundamental theory with universal applicability (rather than an effective theory with a limited domain) is problematic. For example Polchincski [13] suggests that these non-linear effects could lead to an Einstein-Podolsky-Rosen paradox even more troubling than that in standard quantum mechanics. Nevertheless we feel that the study of an effective theory–within its range of applicability–is worthwhile; after all, the Schrodinger equation itself is only a phenomenological approximation which ignores relativistic effects, the quantized description of the external fields, etc.

In any event this system provides an opportunity to demonstrate the application of Theorem 1. The system describes a spin 3/2 particle in a weak magnetic field. Using the nonlinear theory, the Hamiltonian for this system is first expressed in terms of complex variables, $\psi_{+3/2}$, $\psi_{+1/2}$, $\psi_{-1/2}$, and $\psi_{-3/2}$.

$$\begin{aligned} H = {} & \omega_{+3/2}|\psi_{+3/2}|^2 + \omega_{+1/2}|\psi_{+1/2}|^2 + \omega_{-1/2}|\psi_{-1/2}|^2 + \omega_{-3/2}|\psi_{-3/2}|^2 \\ & - \frac{9n\epsilon}{4} + \frac{\epsilon}{n}\Big[2|\sqrt{3}\psi_{+3/2}\psi_{-1/2} - \psi_{+1/2}^2|^2 + |3\psi_{+3/2}\psi_{-3/2} - \psi_{+1/2}\psi_{-1/2}|^2 \\ & + 2|\sqrt{3}\psi_{-3/2}\psi_{+1/2} - \psi_{-1/2}^2|^2\Big] \end{aligned}$$

where $n = |\psi_{+3/2}|^2 + |\psi_{+1/2}|^2 + |\psi_{-1/2}|^2 + |\psi_{-3/2}|^2$. The time evolution of the system is determined by the equations

$$\frac{d\psi_k}{dt} = -i\frac{\partial H}{\partial \psi_k^*} \qquad \frac{d\psi_k^*}{dt} = i\frac{\partial H}{\partial \psi_k} \tag{7.1}$$

Changing variables so that

$$J_k = |\psi_k|^2 \qquad \theta_k = \mathrm{Arg}(\psi_k)$$

we have

$$
\begin{aligned}
H &= \omega_{+3/2}J_{+3/2} + \omega_{+1/2}J_{+1/2} + \omega_{-1/2}J_{-1/2} + \omega_{-3/2}J_{-3/2} - \frac{9n\epsilon}{4} \\
&+ \frac{\epsilon}{n}\big[6(J_{+3/2}J_{-1/2} + J_{-3/2}J_{+1/2}) + 2(J_{+1/2}^2 + J_{-1/2}^2) + 9J_{+3/2}J_{-3/2} \\
&+ J_{+1/2}J_{-1/2} - 4J_{+1/2}\sqrt{3J_{+3/2}J_{-1/2}}\cos(\theta_{+3/2} + \theta_{-1/2} - 2\theta_{+1/2}) \\
&- 4J_{-1/2}\sqrt{3J_{-3/2}J_{+1/2}}\cos(\theta_{-3/2} + \theta_{+1/2} - 2\theta_{-1/2}) \\
&- 6\sqrt{J_{+3/2}J_{+1/2}J_{-1/2}J_{-3/2}}\cos(\theta_{+3/2} + \theta_{-3/2} - \theta_{+1/2} - \theta_{-1/2})\big]
\end{aligned}
$$

Under this transformation the equations, (7.1), become

$$
\frac{dJ_k}{dt} = -\frac{\partial H}{\partial \theta_k} \qquad\qquad \frac{d\theta_k}{dt} = \frac{\partial H}{\partial J_k}
$$

The parameter representing the size of the nonlinear contribution, ϵ, is small so this Hamiltonian has the form of a linear function on the actions perturbed by a small function of both the actions and angles. To check that the hypotheses of our theorem apply we first choose some more appropriate coordinates.

This system can be reduced from four degrees of freedom to two by a linear, canonical transformation

$$
\begin{aligned}
\Phi : &(\Theta_x, \Theta_y, \Theta_z, \Theta_w, x, y, z, w) \\
&\to (\theta_{+3/2}, \theta_{+1/2}, \theta_{-1/2}, \theta_{-3/2}, J_{+3/2}, J_{+1/2}, J_{-1/2}, J_{-3/2})
\end{aligned}
$$

of the form

$$
\Phi(q, p) = (Aq, Bp)
$$

where $A \in SL_4(\mathbb{Z})$, and B is the inverse transpose of A. In particular,

$$
A = \begin{bmatrix} 1 & -1 & 0 & 1 \\ 1 & -2 & 1 & 0 \\ 1 & -3 & 3 & -1 \\ 1 & 0 & -1 & 1 \end{bmatrix}
$$

The transformed Hamiltonian is

$$\begin{aligned}
H &= (\omega_{+3/2} - \omega_{+1/2} + \omega_{-3/2})x + (\omega_{+3/2} - \omega_{-1/2} + \omega_{-3/2})w \\
&+(\omega_{+3/2} - 2\omega_{+1/2} + \omega_{-1/2})y + (\omega_{+3/2} - 3\omega_{+1/2} + 3\omega_{-1/2} - \omega_{-3/2})z \\
&-\frac{9\epsilon(x+w)}{4} + \frac{\epsilon}{x+w}\big[14y^2 + 54z^2 + 54yz + (10x - 5w)y + 15(x-w)z \\
&+7xw + 5(x^2+w^2) + 4(x+2y+3z)\sqrt{3(x+y+z+w)(y+3z-w)}\cos\Theta_y \\
&-4(y+3z-w)\sqrt{3(x-z+w)(-x-2y-3z)}\cos(\Theta_y - \Theta_z) \\
&-6\sqrt{(x+y+z+w)(x+2y+3z)(w-y-3z)(x-z+w)}\cos(2\Theta_y + \Theta_z)\big]
\end{aligned} \tag{7.2}$$

x and w are now constants of motion, so we now need only consider the momenta y and z and their conjugate angles.

The frequencies of the unperturbed, linear Hamiltonian are assumed to be of the form

$$\omega_m = bm + cm^2 \tag{7.3}$$

for constants b, c. When (7.3) is used to simplify the unperturbed Hamiltonian, we find that the frequency associated to the angle Θ_z is zero. This implies that the given frequencies are in resonance. In the language of our result, then, H has the form

$$H(y, z, \Theta_y, \Theta_z) = 4cy + \epsilon[h_0(y,z) + h(y, z, \Theta_y, \Theta_z)]$$

where there are no varying variables ($l = 0$), one "fast" pair (y, Θ_y), and one "slow" pair, (z, Θ_z). Here h_0 consists of the terms in (7.2) proportional to $\epsilon/(x+w)$ which are independent of the angles; h consists of those proportional to $\epsilon/(x+w)$ but dependent on the angles. Checking hypotheses, we first note that the nonresonance condition is trivially satisfied. For the nondegeneracy condition

$$\mathrm{Hess}(\bar{H}) = \mathrm{Hess}(\epsilon h_0) = \frac{\epsilon}{x+w}\begin{pmatrix} 28 & 54 \\ 54 & 108 \end{pmatrix}$$

which is clearly a nonsingular matrix. H is analytic for all x, y, z, and w which lie in the image of the region $\{J_k > 0\}$ under the linear tranformation. Finally, the smallness condition on the perturbation averaged over the fast angles is satisfied since the average of $h - \bar{h}$ over Θ_y is zero. Hence, for sufficiently small ϵ this system will have a large measure of invariant tori.

We hope that our main result can be applied to a diverse collection of problems arising from this theory. It could be possible, using similar methods, to provide lower bounds for the discrepancy between linear and non-linear theory for specific models. One can hope that such calculations may help to clarify the requirements for experimental verification of these theories.

8 Acknowledgements

The author wishes to thank Rafael de la Llave for the comments, discussions, and encouragement without which this paper would never have been completed.

References

[1] J. Pöschel. Integrability of Hamiltonian systems on Cantor sets. *Comm. Pure Appl. Math.*, 35:653–696, 1982.

[2] V.I. Arnold. Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian. *Uspehi Mat. Nauk*, 18(5):13–40, 1963.

[3] A.I. Neishtadt. Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions. *J. Appl. Math. Mech.*, 45(6):766–772, 1981.

[4] Rüssman. Nondegeneracy in the perturbation theory of integrable dynamical systems. *Math. Appl.*, 59:211–223, 1991.

[5] M.B. Sevryuk H.W. Broer, G.B. Huitema. *Quasi-Periodic Motions in Families of Dynamical Systems: Order amidst Chaos*. Springer, 1996.

[6] M.B. Sevryuk. KAM-stable Hamiltonians. *J. Dynam. Control Systems*, 1(3):351–366, 1995.

[7] Y. Sun C. Cheng. Existence of KAM tori in degenerate Hamiltonian systems. *J.Diff.Equ*, 114:288–335, 1994.

[8] M. Herman. Lecture notes. from a talk given at the University of Texas at Austin.

[9] V. Arnold. Small denominators. *Russ. Math Survey*, 18, N6:85–192, 1963.

[10] S. Weinberg. Testing quantum mechanics. *Annals of Physics*, 194:336–386, 1989.

[11] A. N. Kolmogorov. Preservation of conditionally periodic movements with small change in the Hamiltonian functions. *Akad. Nauk. Dokl.*, 98:527–530, 1954.

[12] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn. A proof of Kolmogorov's theorem on invariant tori using canonical transformations defined by the Lie method. *Il Nuovo Cimento*, 79 B(2):201–223, Febbraio 1984.

[13] J. Polchincski. Weinberg's nonlinear quantum mechanics and the Einstein-Podolsky-Rosen paradox. *Phys. Rev. Let.*, 66, n4:397–400, 1991.