Abstract. We consider the KAM theory for rotational flows on an $n$-dimensional torus. We show that if its frequencies are diophantine of type $n - 1$, then Moser’s KAM theory with parameters applies to small perturbations of weaker regularity than $C^n$. Derivatives of order $n$ need not be continuous, but rather $L^2$ in a certain strong sense. This disproves the long standing conjecture that $C^n$ is the minimal regularity assumption for KAM to apply in this setting while still allowing for Herman’s $C^{n-\varepsilon}$-counterexamples.

We consider one of the model problems of KAM theory, namely the perturbation of a constant vector field

$$N = \sum_{i=1}^{n} \omega_i \partial_i$$

on an $n$-dimensional torus given by $n$ frequencies

$$\omega = (\omega_1, \ldots, \omega_n).$$

The flow of this vector field is very simple, being the superposition of uniform rotations on each circle of the torus.

Upon perturbation this simple flow is usually destroyed, and chaotic behaviour may set in. However, it is one of the fundamental results of KAM theory that if the frequencies are strongly nonresonant, and if the perturbation $P$ is sufficiently smooth and sufficiently small, then there exists a modifying constant vector field $Y$ so that $N + P - Y$ is indeed conjugate to $N$. Or to put it succinctly, ‘KAM applies’. 
We address the question of how smooth that perturbation has to be. As is well known the answer depends on the diophantine properties of the frequencies $\omega$. So we assume that $\omega$ is diophantine with exponent $\tau$, that is,

$$|\langle k, \omega \rangle| \geq \alpha |k|^{-\tau}, \quad 0 \neq k \in \mathbb{Z}^n,$$

with some $\alpha > 0$. This holds for almost all frequencies when $\tau > n - 1$, and on a non-empty zero set when $\tau = n - 1$.

Now consider a perturbation $P$ of $N$. Writing

$$P = \sum_{0 \neq k \in \mathbb{Z}^n} p_k e_k, \quad e_k = e^{\langle k, x \rangle},$$

we measure the size of $P$ with respect to the norms

$$\|P\|_{r,b} := \sum_{\nu \geq 0} \left( \sum_{b^{\nu-1} < |k| \leq b^{\nu}} |p_k|^2 |k|^{2r} \right)^{1/2}.$$ 

where $b \geq 2$ is an integer. These norms interpolate between the standard weighted $\ell^1$- and $\ell^2$-norms

$$\|P\|_{r,1} := \sum_{k \neq 0} |p_k| |k|^r, \quad \|P\|_{r,\infty} := \left( \sum_{k \neq 0} |p_k|^2 |k|^{2r} \right)^{1/2}.$$

Indeed, as one easily verifies, $\|P\|_{r,1} \geq \|P\|_{r,b} \searrow \|P\|_{r,\infty}$ as $b \to \infty$.

**Differentiable Kam Theorem** Suppose the frequencies $\omega$ of $N$ are diophantine with exponent $\tau$. Then KAM applies to perturbations $P$ of $N$, whenever $\|P\|_{\tau+1,b}$ is sufficiently small for some integer $b \geq 2$.

This result is new and, it seems, unexpected. Consider the case of $n$ degrees of freedom and $\tau = n - 1$. It was conjectured for a long time, and sometimes even stated as fact, that $C^n$ is the minimal regularity requirement for KAM to apply in this case [1, 2, 3]. This, however, is not true. For $\|P\|_{n,b}$ to be small for any $b \geq 2$, the $n - 1$-derivative of $P$ need not be continuous, but only of strong $L^2$-type in the sense of this norm.

We point out that this result does not contradict the counterexamples of class $C^{n-\varepsilon}$ given by Herman [3]. It rather fits precisely into the gap between the class of $C^n$-perturbations, for which KAM was known to apply, and the class of $C^{n-\varepsilon}$-counterexamples of Herman.

The rest of this paper is devoted to the proof of this theorem. We combine a succession of coordinate transformations familiar in KAM theory with the analytic
smoothing approach of Moser [5]. This way no analytic perturbation theorem as
an intermediate step is required. The convergence speed of this procedure is not of
Newton type, but rather slow in the spirit of Rüssmann [8, 6]. Indeed, at each step
of the iteration, we do not linearize the conjugacy equation under consideration, but
rather solve a nonlinear equation with the help of the Brouwer fixed point theorem.

1 Norms

The proof of the main theorem makes use of the interplay of two different
norms for analytic functions and maps. For a function or map \( f \) on the torus with
Fourier series expansion \( f = \sum_{k \in \mathbb{Z}^n} f_k e_k \) and \( s \geq 0 \) we set
\[
|f|_s := \sum_{k \in \mathbb{Z}^n} |f_k| e^{s|k|},
\]
and, with \( U_s : |\text{Im } z| \leq s \),
\[
\|f\|_s^2 := \frac{1}{|U_s|} \int_{U_s} |f|^2 \quad = \quad \frac{1}{(4\pi s)^n} \int_{[-s,s]^n} \int_{[-\pi,\pi]^n} |f(x + \beta y)|^2 \, dx \, dy.
\]
This norm also has a representation with particular weights.

**Lemma 1.1** For \( f = \sum_k f_k e_k \),
\[
\|f\|_s^2 = \sum_{k \in \mathbb{Z}^n} |f_k|^2 w_k(2s)
\]
with
\[
w_k(t) = \prod_{i=1}^n \frac{\sinh(tk_i)}{tk_i}
\]
and the convention that \( t^{-1} \sinh t = 1 \) for \( t = 0 \).

**Proof** The general term of \( |f|^2 \) at \( z = x + \beta y \) is
\[
f_k e^{\beta \langle k, z \rangle} \bar{f}_l e^{-\beta \langle l, \bar{z} \rangle} = f_k \bar{f}_l e^{\beta \langle k-l, x \rangle} e^{-\langle k+l, y \rangle}.
\]
Its mean integral over \( U_s \) vanishes whenever \( k \neq l \). For \( k = l \) we get
\[
\frac{1}{(2s)^n} \int_{[-s,s]^n} |f_k|^2 e^{-\langle 2k, y \rangle} \, dy = |f_k|^2 \prod_{i : k_i \neq 0} \frac{\sinh(2sk_i)}{2sk_i},
\]
which gives the claimed identity. 

We note that for constant functions \( Z, |Z| = |Z|_0 = |Z|_s = \|Z\|_s \) for all \( s \geq 0 \). Moreover,

\[
|f|_{U_s} := \sup_{U_s} |f| \leq |f|_s ,
\]

so we have

\[
\|fg\|_s \leq |f|_s \|g\|_s
\]

and in particular \( \|f\|_s \leq |f|_s \). Finally, the exponential weights are submultiplicative, so we have the Banach algebra property

\[
|fg|_s \leq |f|_s |g|_s .
\]

The usual Neumann series argument then gives us

**Lemma 1.2** Let \( \varphi = I + \hat{\varphi} \) be a self map the \( n \)-torus. If \( \mu = |D\hat{\varphi}|_s < 1 \), then \( \varphi \) is a diffeomorphism, and for its inverse \( \psi = I + \hat{\psi} \) one has

\[
|D\hat{\psi}|_s \leq \frac{\mu}{1 - \mu}, \quad |D\psi|_s \leq \frac{1}{1 - \mu}.
\]

A central role is played by the behaviour of these norms under coordinate transformations.

**Lemma 1.3** For a diffeomorphismus \( \varphi \) of the \( n \)-torus extending to \( U_s \),

\[
\|f \circ \varphi\|_s \leq \lambda \|f\|_{\lambda s}, \quad \lambda := |D\varphi|_{U_s} \vee |D\varphi^{-1}|_{\varphi(U_s)} \geq 1,
\]

where \( \vee \) denotes the maximum operator. Moreover, if \( |\varphi - I|_s \leq \sigma \), then

\[
|f \circ \varphi|_s \leq |f|_{s+\sigma} .
\]

**Proof** We have \( \varphi(U_s) \subset U_{\lambda s} \) by the reality of \( \varphi \) and the bound on \( D\varphi \). Therefore,

\[
\int_{U_s} |f \circ \varphi|^2 \leq \int_{\varphi(U_s)} |f|^2 |D\varphi^{-1}|^2 \leq |D\varphi^{-1}|_{\varphi(U_s)}^2 \int_{U_s} |f|^2 \leq \lambda^2 \int_{U_{\lambda s}} |f|^2 ,
\]

proving the first claim. Writing \( \varphi = I + \hat{\varphi} \) we have \( e_k \circ \varphi = e_k \cdot e_k \circ \hat{\varphi} \) and thus

\[
|e_k \circ \varphi|_s \leq |e_k|_s |e_k \circ \hat{\varphi}|_s \leq |e_k|_s \exp(|\langle k, \hat{\varphi}\rangle|_s) = e^{(s+\sigma)|k|}.
\]
For general $f$ we thus obtain

$$|f \circ \varphi|_s \leq \sum_{k \in \mathbb{Z}^n} |f_k| |e_k \circ \varphi|_s \leq \sum_{k \in \mathbb{Z}^n} |f_k| e^{(s+\sigma)|k|} = |f|_{s+\sigma}. \quad \blacksquare$$

We also have a Cauchy inequality connecting the two norms $|\cdot|_s$ and $\|\cdot\|_s$.

**Lemma 1.4** For compatible maps $f$ and $\varphi$ on the $n$-torus,

$$\|Df \cdot \varphi\|_{\alpha s} \leq \frac{1}{e^{\alpha n/2}} \frac{1}{(1-\alpha)s} \|f\|_s |\varphi|_s, \quad 0 < \alpha < 1.$$  

**Proof** It is easy to verify that $\frac{\sinh t}{t} e^{-t}$ is decreasing for all $t$. Hence,

$$\frac{\sinh s}{s} \leq \frac{\sinh(s-t)}{s-t} e^{t}, \quad 0 \leq t \leq s.$$  

It follows that for all $k$ and $l$ we have the submultiplicativity property

$$w_k(s) \leq w_{k-l}(s) e^{s|l|}, \quad s \geq 0,$$

connecting the weights of $\|\cdot\|_s$ and $|\cdot|_s$.

Now consider the Fourier series expansion

$$\frac{1}{\beta} Df \cdot \varphi = \sum_{k \in \mathbb{Z}^n} \langle k, \varphi \rangle f_k e_k = \sum_{k,l \in \mathbb{Z}^n} \langle k-l, \varphi_l \rangle f_{k-l} e_l.$$  

By the triangle inequality,

$$\|Df \cdot \varphi\|_{\alpha s} \leq \left( \sum_{k \in \mathbb{Z}^n} \left| \sum_{l \in \mathbb{Z}^n} \langle k-l, \varphi_l \rangle f_{k-l} \right|^2 w_k(2\alpha s) \right)^{1/2}$$

$$\leq \sum_{l \in \mathbb{Z}^n} \left( \sum_{k \in \mathbb{Z}^n} |k-l|^2 |\varphi_l|^2 |f_{k-l}|^2 w_k(2\alpha s) \right)^{1/2}.$$  

With inequality (1) and a subsequent re-indexing it follows that

$$\|Df \cdot \varphi\|_{\alpha s} \leq \sum_{l \in \mathbb{Z}^n} |\varphi_l| e^{\alpha s|l|} \left( \sum_{k \in \mathbb{Z}^n} |k|^2 |f_k|^2 w_k(2\alpha s) \right)^{1/2}$$

$$\leq \sup_{k \in \mathbb{Z}^n} |k| \left( \frac{w_k(2\alpha s)}{w_k(2s)} \right)^{1/2} \|f\|_s |\varphi|_s.$$  

Now, with \( \frac{\sinh s}{\sinh t} \leq e^{s-t} \) for \( s \leq t \) we get

\[
\frac{w_k(s)}{w_k(t)} = \left( \frac{s}{t} \right)^n \prod_{1 \leq i \leq n} \frac{\sinh(sk_i)}{\sinh(tk_i)} \leq \left( \frac{s}{t} \right)^n e^{s-t} |k|,
\]

so the supremum is bounded by

\[
\sup_{t \geq 0} \frac{t}{\alpha^{n/2}} e^{(\alpha-1)st} = \frac{1}{s\alpha^{n/2}} \sup_{t \geq 0} r e^{(\alpha-1)r} = \frac{1}{e\alpha^{n/2} (1-\alpha)s}.
\]

This proves the claim. \( \Box \)

2 Small Divisors and Cut Offs

Next we consider the solution of the typical small divisor equation \( \partial_s \phi = f \) with a nonresonant frequency vector \( \omega \). Formally, for \( f = \sum_{k \neq 0} f_k e_k \), the unique solution with mean value zero is

\[
\phi = Lf := \sum_{k \neq 0} \frac{f_k}{\beta(k,\omega)} e_k.
\]

**Lemma 2.1** Suppose that

\[
\Omega := \max_{0 < |k| \leq K} |\langle k, \omega \rangle|^{-1} < \infty.
\]

Then, for a trigonometric polynomial \( f \) of order \( K \) without constant term,

\[
|Lf|_s \leq C\Omega \|f\|_s, \quad C = 2^ne^{sK}.
\]

**Proof** With Cauchy-Schwarz,

\[
|Lf|_0^2 \leq \sum_{0 < |k| \leq K} \frac{1}{\langle k, \omega \rangle^2} \sum_{0 < |k| \leq K} |f_k|^2.
\]

By an elegant estimate due to Rüssmann [7],

\[
\sum_{0 < |k| \leq K} \frac{1}{\langle k, \omega \rangle^2} \leq 2^{n+2} \Omega^2,
\]

while the second sum is just \( \|f\|_0^2 \). Since the exponential weights in \( |Lf|_s \) are all bounded by \( e^{sK} \) and the weights in \( \|f\|_s \) are all \( \geq 1 \) we obtain

\[
|Lf|_s \leq e^{sK} |Lf|_0 \leq C\Omega \|f\|_0 \leq C\Omega \|f\|_s. \quad \Box
\]
We also need an estimate for ultra-violent cut offs.

**Lemma 2.2** Suppose \( f \) contains no Fourier coefficients up to order \( K \). Then

\[
\|f\|_{\alpha s} \leq \alpha^{-n/2} e^{(\alpha-1)sK} \|f\|_s, \quad 0 < \alpha \leq 1.
\]

**Proof** For \( f = \sum_{|k| \geq K} f_k e_k \) we get

\[
\|f\|_{\alpha s}^2 = \sum_{|k| \geq K} |f_k|^2 w_k(2\alpha s) \leq \sup_{|k| \geq K} \frac{w_k(2\alpha s)}{w_k(2s)} \|f\|_s^2.
\]

By the same argument as in (2)

\[
\sup_{|k| \geq K} \frac{w_k(2\alpha s)}{w_k(2s)} \leq \alpha^{-n/2} e^{(\alpha-1)sK}.
\]

---

3 Outline of Proof and Step Lemma

Suppose we already found a modifying term \( Y \) and a coordinate transformation \( \Psi \) so that

\[
\Psi^*(N + P - Y) = N + Q,
\]

where \( P \) is an analytic approximation – indeed a trigonometric polynomial – to the original smooth perturbation we are aiming at, and \( Q \) is smaller than \( P \). We then construct another modifying term \( Z \) and a transformation \( \Phi \) so that

\[
\Phi^*(N + Q - \Psi^*Z) = N + Q^+
\]

improves on \( Q \). Setting \( \Psi_+ = \Psi \circ \Phi \) and \( Y_+ = Y + Z \) we obtain

\[
\Psi_+^*(N + P - Y_+) = \Phi^*(\Psi^*(N + P - Y) - \Psi^*Z)
\]

\[
= \Phi^*(N + Q - \Psi^*Z)
\]

\[
= : N + Q^+.
\]

Passing from \( P \) to the next approximation \( P_+ \) we arrive at

\[
\Psi_+^*(N + P_+ - Y_+) = N + Q^+ + \Psi_+^*(P_+ - P) =: N + Q_+,
\]

which completes one cycle of the iterative procedure. The following lemma describes the quantitative details of this construction.
Step Lemma  Consider $\Psi^*(N + P) = N + Q$. Assume that

$$4\Delta \|Q\|_s \leq \frac{1}{4}, \quad \|D\Psi - I\|_s \leq \frac{1}{7},$$

where $\Delta = CK\Omega$ with $C = 2^n e^{sK}$ and

$$\Omega = \max_{0 < |k| \leq K} |\langle k, \omega \rangle|^{-1}.$$

Also assume that $sK \geq (4/3)^{n/2}$. Then there exists a unique modifying term $Z$ and a unique coordinate transformation $\Phi = I + \hat{\Phi}$ with

$$\Delta |Z|_0 \vee K \|\hat{\Phi}\|_{s/2} \leq 4\Delta \|Q\|_s,$$

so that $\Phi^*\Psi^*(N + P - Z) = N + Q^+$ with

$$\|Q^+\|_{s/4} \leq 4 \|(I - \Pi)Q\|_{s/2},$$

where $\Pi$ denotes truncation of Fourier series at order $K$.

Proof  Consider equation (3) which with $\Phi = I + \hat{\Phi}$ is equivalent to

$$D\hat{\Phi} \cdot N + D\Phi \cdot Q^+ = (Q - \Psi^*Z) \circ \Phi.$$  \hspace{1cm} (4)

Instead of this functional equation we solve the finite dimensional equation

$$D\hat{\Phi} \cdot N = \Pi((Q - \Psi^*Z) \circ \Phi),$$  \hspace{1cm} (5)

where $\Pi$ denotes truncation of a Fourier series at order $K$. Writing

$$\Psi^*Z = Z - \Theta Z, \quad \Theta = D\Psi^{-1}(D\Psi - I),$$

its right hand side becomes $\Pi((Q - \Psi^*Z) \circ \Phi) = \Pi((Q + \Theta Z) \circ \Phi) - Z$ so that this equation amounts to

$$D\hat{\Phi} \cdot N + Z = \Pi T(Z, \hat{\Phi})$$  \hspace{1cm} (6)

with the nonlinear operator

$$T(Z, \hat{\Phi}) := (Q + \Theta Z) \circ (I + \hat{\Phi}).$$

Here, $D\hat{\Phi} \cdot N$ is the familiar linear differential operator $\partial_\omega \hat{\Phi}$ giving rise to small divisors. Its inverse on the space of trigonometric polynomials of order $K$ with vanishing mean value is the operator $L$ considered in Lemma 2.1. Thus, a
solution of equation (6) is a fixed point of the mapping
\[ Z_1 = \Pi_0 T(Z, \hat{\Phi}), \quad \hat{\Phi}_1 = L \Pi_1 T(Z, \hat{\Phi}), \]
where $\Pi_0$ denotes the operator of taking the mean value over the $n$-torus and $\Pi_1 = \Pi - \Pi_0$. This we solve with the Banach contraction principle.

To this end let $\alpha = 1/2$ for brevity and consider the ball
\[ \mathcal{B} : \quad \Delta |Z|_0 \vee K |\hat{\Phi}|_{\alpha s} \leq 4\Delta \|Q\|_s \leq \frac{1}{4} \]
within the space of constant terms times trigonometric polynomials of order $K$ without constant terms. By assumption on $\Psi$ and Lemma 1.2,

\[ |\Theta|_s = |D\Psi^{-1}|_s |D\Psi - I|_s \leq \frac{1/7}{1 - 1/7} = \frac{1}{6}. \]
Hence, $\|\Theta Z\|_s \leq |\Theta Z|_s \leq |\Theta|_s |Z|_0 \leq \|Q\|_s$ and thus
\[ \|Q + \Theta Z\|_s \leq 2 \|Q\|_s. \]
Similarly, $|D\hat{\Phi}|_{\alpha s} \leq K |\hat{\Phi}|_{\alpha s} \leq 1/4$ and therefore
\[ |D\Phi|_{\alpha s} \vee |D\Phi^{-1}|_{\alpha s} \leq \frac{3}{2}. \]
So with Lemma 1.3
\[ \|T(Z, \hat{\Phi})\|_{\alpha s} = \|(Q + \Theta Z) \circ (I + \hat{\Phi})\|_{\alpha s} \leq 2 \|Q + \Theta Z\|_s \leq 4 \|Q\|_s. \]
It follows that
\[ \Delta |Z_1|_0 = \Delta |\Pi_0 T(Z, \hat{\Phi})|_0 \leq 4\Delta \|Q\|_s, \]
\[ K |\hat{\Phi}_1|_{\alpha s} = K |L \Pi_1 T(Z, \hat{\Phi})|_{\alpha s} \leq CK \Omega \|T(Z, \hat{\Phi})\|_{\alpha s} \leq 4\Delta \|Q\|_s. \]
So our ball $\mathcal{B}$ is mapped into itself, and continuously so. So the Brouwer fixed point theorem applies.

Indeed, this map is a contraction. Consider
\[ T(Z, \hat{\Phi}) - T(Z', \hat{\Phi}') = (Q + \Theta Z) \circ \Phi - (Q + \Theta Z) \circ \Phi' \]
\[ + (\Theta(Z - Z')) \circ \Phi' =: A + B. \]
In view of (8) any map in $\mathcal{B}$ maps $U_{\alpha s}$ into $U_{s-\sigma}$ with $\sigma = s/4$. So we can apply the Cauchy estimate of Lemma 1.4 – here with $\alpha = 3/4$ – to $A$ to obtain

$$\|A\|_{\alpha s} \leq \left(\frac{4}{3}\right)^{n/2} \frac{2}{s} \|Q + \Theta Z\|_s |\hat{\Phi} - \hat{\Phi}'|_{\alpha s}$$

$$\leq \left(\frac{4}{3}\right)^{n/2} \frac{4}{sK} \|Q\|_s \|\hat{\Phi} - \hat{\Phi}'|_{\alpha s} \leq 4 \|Q\|_s K |\hat{\Phi} - \hat{\Phi}'|_{\alpha s}$$

by assuming that $sK \geq (4/3)^n$. Similarly, by Lemma 1.3

$$\|B\|_{\alpha s} = \|(\Theta(Z - Z')) \circ \Phi'\|_{\alpha s} \leq \frac{3}{2} \|\Theta(Z - Z')\|_s \leq \frac{1}{4} |Z - Z'|_0.$$

Since we assume that $4\Delta \|Q\|_s \leq 1/4$ it follows that

$$K |\hat{\Phi}_1 - \hat{\Phi}'_1|_{\alpha s} \leq \Delta \|A\|_{\alpha s} + \Delta \|B\|_{\alpha s} \leq \frac{1}{4} K |\hat{\Phi} - \hat{\Phi}'|_{\alpha s} + \frac{1}{4} \Delta |Z - Z'|_0.$$

Exactly the same estimate holds for $\Delta |Z_1 - Z'_1|_0$. But this means that with respect to the norm $\Delta |Z - Z'|_0 \vee K |\hat{\Phi} - \hat{\Phi}'|_{\alpha s}$

we obtain a contraction by the factor $1/2$.

The estimates for the unique fixed point $Z, \hat{\Phi}$ are the same as for the ball $\mathcal{B}$. So it remains to bound $Q^+$. From (4) and (5) we deduce that

$$D\Phi \cdot Q^+ = ((I - \Pi)(Q - \Psi^*Z) \circ \Phi).$$

With $\|\Psi^*Z\|_s \leq \|D\Psi^{-1}\|_s |Z|_0 \leq \|Q\|_s$, estimate (8) and Lemma 2.2 we arrive at

$$\|Q^+\|_{\alpha 2s} \leq \|D\Phi^{-1} \circ \Phi\|_{\alpha s} \|(I - \Pi)(Q - \Psi^*Z) \circ \Phi)\|_{\alpha 2s}$$

$$\leq 2 \|(I - \Pi)(Q - \Psi^*Z)\|_{\alpha s}$$

$$\leq 4 \|(I - \Pi)Q\|_{\alpha s}.$$

This finishes the proof of the Step Lemma.  

4 Iteration

We now assume the frequency vector $\omega$ of the vector field $N$ to be diophantine with exponent $\tau$. Scaling time, we may even assume that

$$|\langle k, \omega \rangle| \geq |k|^{-\tau}, \quad k \neq 0.$$

We recall that $\tau \geq n - 1$, since otherwise no such frequencies exist.

To simplify the exposition we now assume with loss of generality that $\|P\|_{r,b}$ is small with $b \geq 4$. For $\nu \geq 0$ we set

$$K_\nu = b^\nu, \quad s_\nu = \frac{r}{b^\nu}, \quad \Delta_\nu = K_\nu^{\tau + 1} = b^{\nu(\tau + 1)}.$$

Thus, $s_\nu K_\nu = r$ for all $\nu \geq 0$, and we may simplify the estimates of section 2

$$K_\nu |Lf|_{s_\nu} \vee |DLf|_{s_\nu} \leq 2^n e^r \Delta_\nu \|f\|_{s_\nu},$$

$$\|F\|_{s_\nu/2} \leq 2^{n/2} e^{-r/2} \|F\|_{s_\nu},$$

for $F$ containing no Fourier coefficients up to order $K_\nu$. We then fix $r$ so that

$$\beta := 4^n e^{-r/2} \leq \frac{1}{2. b^{\tau + 1}}.$$

We approximate the given perturbation $P = \sum_{k \neq 0} p_k e_k$ by the sequence of trigonometric polynomials

$$P_\nu = \sum_{|k| \leq K_\nu} p_k e_k, \quad \nu \geq 0.$$

To make the exposition more transparent with respect to the norm of $P$ we introduce the weighted $L^2$-norm

$$\|P\|_m^2 = \sum_{k \neq 0} |p_k|^2 m_k^2.$$

**Lemma 4.1** For $\Delta P_0 = P_0$ and $\Delta P_\nu = P_\nu - P_{\nu - 1}$ for $\nu \geq 1$ we have

$$\|\Delta P_\nu\|_{s_\nu} \leq \frac{e^r}{m_\nu} \|\Delta P_\nu\|_m, \quad \nu \geq 0,$$

with $m_0 = 1$ and $m_\nu = \min_{|k| > K_{\nu - 1}} m_k$ for $\nu \geq 1$.  

**Proof** With $w_k(s) \leq e^{s|k|}$ and $\mathcal{K}_\nu: K_{\nu - 1} < |k| \leq K_\nu$ we get

$$\|\Delta P_\nu\|_{s_\nu}^2 = \sum_{k \in \mathcal{K}_\nu} |p_k|^2 w_k(2s_\nu)$$

$$\leq \max_{k \in \mathcal{K}_\nu} \frac{e^{2s_\nu|k|}}{m_k^2} \sum_{k \in \mathcal{K}_\nu} |p_k|^2 m_k^2 \leq \frac{e^{2r}}{m_\nu^2} \|\Delta P_\nu\|_m^2. \quad \blacksquare$$
Iterative Lemma  Suppose that \( \sup_{\nu \geq 0} \frac{\Delta_{\nu}}{m_{\nu}} \leq A < \infty \) and that
\[
\varepsilon = \sum_{\nu \geq 0} \|\Delta P_{\nu}\|_m
\]
is sufficiently small. Then for each \( P_{\nu} \) there exists a modifying term \( Y_{\nu} \) and a transformation \( \Psi_{\nu} \) such that
\[
\Psi_{\nu}^*(N + P_{\nu} - Y_{\nu}) = N + Q_{\nu}
\]
with
\[
\|Q_{\nu}\|_{s_{\nu}} \leq \varepsilon_{\nu} := B \sum_{\mu \leq \nu} \frac{b^{\nu-\mu}}{m_{\mu}} \|\Delta P_{\mu}\|_m,
\]
\[
|D\hat{\Psi}_{\nu}|_{s_{\nu}} \leq \delta_{\nu} := \prod_{0 \leq \mu < \nu} (1 + 4\Delta_{\nu}\varepsilon_{\nu}) - 1,
\]
where \( B = 2e^r \). Moreover, \( |Y_{\nu+1} - Y_{\nu}| \leq 4\varepsilon_{\nu} \) and \( |D\hat{\Psi}_{\nu+1} - D\hat{\Psi}_{\nu}|_0 \leq 8\Delta_{\nu}\varepsilon_{\nu} \).

Proof  For \( \nu = 0 \) we can choose \( Y_0 = 0 \) and \( \Psi_0 = I \). Then \( Q_0 = P_0 \), and the estimate for \( Q_0 \) is satisfied by the preceding lemma by choice of \( B \). So we may proceed by induction.

By our choice of \( b \), the sequence \( b^\nu \Delta_{\nu} \) decays geometrically so that
\[
\sum_{\nu \geq \mu} b^\nu \Delta_{\nu} \leq 2 \cdot b^\mu \Delta_{\mu}.
\]
Therefore, with the abbreviation \( \rho_{\mu} := \|\Delta P_{\mu}\|_m \) and \( \Delta_{\nu}/m_{\nu} \leq A \),
\[
\sum_{\nu \geq 0} \Delta_{\nu}\varepsilon_{\nu} = B \sum_{\mu \geq 0} \frac{b^{-\mu}}{m_{\mu}} \sum_{\nu \geq \mu} b^\nu \Delta_{\nu} \leq 2B \sum_{\mu \geq 0} \frac{\Delta_{\mu}}{m_{\mu}} \rho_{\mu} \leq 2AB \sum_{\mu \geq 0} \rho_{\mu} = 2AB \varepsilon.
\]
It follows that for \( \varepsilon \) small enough, the smallness conditions of the Step Lemma are satisfied by \( \Psi_{\nu} \) and \( Q_{\nu} \) for all \( \nu \geq 0 \).

We obtain a modifying term \( Z_{\nu} \) and a transformation \( \Phi_{\nu} \) with
\[
|Z_{\nu}|_0 \leq 4\varepsilon_{\nu}, \quad K_{\nu} \|\hat{\Phi}_{\nu}\|_{s_{\nu}/2} \vee \|D\hat{\Phi}_{\nu}\|_{s_{\nu}/2} \leq 4\Delta_{\nu}\varepsilon_{\nu},
\]
With \( Y_{\nu+1} = Y_{\nu} + Z_{\nu} \) and \( \Psi_{\nu+1} = \Psi_{\nu} \circ \Phi_{\nu} \) we have
\[
\Psi_{\nu+1}^*(N + P_{\nu+1} - Y_{\nu+1}) = N + Q_{\nu+1}^* + \Psi_{\nu+1}^* \Delta P_{\nu+1}.
\]
For $Q_\nu^+$ we have
\[ \|Q_\nu^+\|_{s_\nu/4} \leq 4\| (I - \Pi)Q_\nu \|_{s_\nu/2} \leq 4^n e^{-r/2} \|Q_\nu\|_{s_\nu}. \]

For the other term we have
\[ \|\Psi_{\nu+1}^* \Delta P_{\nu+1}\|_{s_\nu/4} \leq 2\| \Delta P_{\nu+1}\|_{s_\nu/2} \leq \frac{2e^r}{m_{\nu+1}} \|\Delta P_{\nu+1}\|_m. \]

Putting both estimates together and taking into account the definitions of $b$ and $B$ as well as the fact that $s_{\nu+1} \leq s_\nu/4$, we arrive at
\[ \|Q_{\nu+1}\|_{s_{\nu+1}} \leq \|Q_\nu\|_{s_\nu} + \|\Psi_{\nu+1}^* \Delta P_{\nu+1}\|_{s_\nu/4} \]
\[ \leq b\|Q_\nu\|_{s_\nu} + \frac{B}{m_{\nu+1}} \|\Delta P_{\nu+1}\|_m \]
\[ \leq B \sum_{\mu \leq \nu} \frac{b^{\nu+1-\mu}}{m_\mu} \rho_\mu + \frac{B}{m_{\nu+1}} \rho_{\nu+1} = \frac{B}{m_{\nu+1}} \rho_{\nu+1} = \frac{B}{m_{\nu+1}} \rho_{\nu+1} \]
as required. — Finally, consider $\psi_{\nu+1} = I + \hat{\psi}_{\nu+1}$. We have
\[ D\hat{\psi}_{\nu+1} = D\hat{\psi}_\nu \circ \Phi_\nu + D\hat{\psi}_\nu \circ \Phi_\nu \cdot \hat{D}\Phi_\nu. \]

So in view of $|\hat{D}\Phi_\nu|_{s_\nu} \leq \alpha s_\nu$ by estimate (9) we have
\[ |D\hat{\psi}_{\nu+1}|_{s_{\nu+1}} \leq |D\hat{\psi}_\nu|_{s_\nu} + |D\hat{\psi}_\nu|_{s_\nu} |D\hat{\psi}_\nu|_{s_{\nu+1}}. \]

Setting $\eta_\nu = 4\Delta_\nu \varepsilon_\nu$ we thus get
\[ |D\hat{\psi}_{\nu+1}|_{s_{\nu+1}} \leq \delta_\nu + (1 + \delta_\nu)\eta_\nu \]
\[ = (1 + \delta_\nu)(1 + \eta_\nu) - 1 = \prod_{0 \leq \mu < \nu + 1} (1 + \eta_\mu) - 1 = \delta_{\nu+1} \]
as required in the Iterative Lemma. Moreover,
\[ D\Psi_{\nu+1} - D\Psi_\nu = D\Psi_\nu \circ \Phi_\nu - D\Psi_\nu + D\Psi_\nu \circ \Phi_\nu \cdot \hat{D}\Phi_\nu. \]

It is not difficult to see that
\[ |D^2\Psi_\nu|_0 \leq \frac{4}{s_\nu} |D\Psi_\nu|_{s_\nu} \leq \frac{8}{s_\nu} \]
for $\varepsilon$ sufficiently small. As $s_\nu K_\nu = r$ is sufficiently large, we conclude that
\[ |D\Psi_\nu \circ \Phi_\nu - D\Psi_\nu|_0 \leq |D^2\Psi_\nu|_0 |\hat{\Phi}_\nu|_0 \leq \frac{8}{s_\nu} \frac{4\Delta_\nu \varepsilon_\nu}{K_\nu} \leq 4\Delta_\nu \varepsilon_\nu. \]
Similary, $|D\Psi_\nu \circ \Phi_\nu \cdot D\Phi_\nu|_0 \leq 4\Delta_\nu \varepsilon_\nu$. This completes the proof of the Iterative Lemma.

5 Convergence

The convergence of the scheme described in the Iterative Lemma is obvious. The modifying terms $Y_\nu$ have a limit $Y$, the transformations $\Psi_\nu$ have a limit $\Psi$ in the $C^1$-norm, and $Q_\nu$ vanishes with respect to $\|\cdot\|_0$. Hence, passing to the limit in equation

$$\Psi_\nu^*(N + P_\nu - Y_\nu) = N + Q_\nu$$

we obtain

$$\Psi^*(N + P - Y) = N$$

as promised.

References

[1] J. Albrecht, On the existence of invariant tori in nearly-integrable hamiltonian systems with finitely differentiable perturbations. Regular and Chaotic Dynamics 12 (2007) 281–320.

[2] Chong-Qing Chen & Lin Wang, Destruction of Lagrangian torus for positive definite hamiltonian systems. Geom. funct. Anal. 23 (2013) 848–866.

[3] M. Herman, Sur les courbes invariantes par les difféomorphismes de l’anneau. Astérisque 144 (1986).

[4] C. E. Koudjinan, A KAM theorem for finitely differentiable Hamiltonian systems. J. Diff. Equ. 269 (2020) 4720–4750.

[5] J. Moser, Convergent series expansions for quasi-periodic motions. Math. Ann. 169 (1967) 136–176.

[6] J. Pöschel, KAM à la R. Regul. Chaotic Dyn. 16 (2011) 17–23.

[7] H. Rüssmann, Note on Sums Containing Small Divisors. Comm. Pure Appl. Math. 29 (1976) 755–758.

[8] H. Rüssmann, KAM-iteration with nearly infinitely small steps in dynamical systems of polynomial character. Discrete Contin. Dynam. Syst. Ser. S 3 (2010) 683–718.