A Renormalization Group Fixed Point Associated with the Breakup of Golden Invariant Tori

Hans Koch

Department of Mathematics, University of Texas at Austin
Austin, TX 78712

Abstract. We give a computer-assisted proof for the existence of a renormalization group fixed point (Hamiltonian) with non-trivial scaling, associated with the breakup of invariant tori with rotation number equal to the golden mean.

1. Introduction and main results

Renormalization group methods for Hamiltonian flows and related maps were introduced first in [1,2], as a tool for explaining the breakup of golden invariant tori. Since then, these methods have been developed further and extended to other systems [3–30], but one of the main problems has remained open: The existence of a non-trivial renormalization group (RG) fixed point. In this paper, we prove that such a fixed point exists. Its connection with critical invariant tori was discussed earlier in [30].

The general form of our RG transformation $\mathcal{R}$ can be described as follows. Consider the matrix $T=\left[\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right]$, whose eigenvalues are the golden mean $\vartheta=\frac{1}{2}+\frac{1}{2}\sqrt{5}$, and $-\vartheta^{-1}$. This matrix defines a map on the torus $\mathbb{T}^2$, and it can be extended to a symplectic map $\mathcal{T}$ on $\mathbb{T}^2\times\mathbb{R}^2$, by setting $\mathcal{T}(q,p)=\bigl(Tq,T^{-1}p\bigr)$. If H is a Hamiltonian function on $\mathbb{T}^2\times\mathbb{R}^2$, then

$$\mathcal{R}(H)=H\circ\mathcal{T}\pmod{\mathcal{G}}\,. \tag{1.1}$$

Here, $\mathcal{G}$ is a group (ideally) of transformations that preserves rotation numbers. This includes coordinate changes $H\mapsto H\circ U$, with U canonical and homotopic to the identity, a scaling of of the action variables $H\mapsto\mu^{-1}H(.,\mu.)$, and a scaling of the energy $H\mapsto\tau H-\epsilon$.

Thus, a specific RG transformation is obtained by choosing a map $H\mapsto G_H$ with values in $\mathcal{G}$, and defining $\mathcal{R}(H)=G_H(H\circ\mathcal{T})$. The goal is to do this in such a way that $\mathcal{R}$ defines a hyperbolic dynamical system on some manifold of "normalized" Hamiltonians, which is contracting in all but a finite number of directions. By contrast to other RG schemes in dynamical systems, this seems possible only if we allow $\mathcal{G}$ to be infinite dimensional. This is one of the interesting and challenging aspect of a renormalization group for Hamiltonians. We will now define the notion of a "resonant" Hamiltonian, and after introducing an appropriate function space, the resonant subspace will serve as our manifold of normalized Hamiltonians.

Consider the eigenvectors $\omega=(\vartheta^{-1},1)$ and $\Omega=(1,-\vartheta^{-1})$ of the matrix T, associated with the eigenvalues ϑ and $-\vartheta^{-1}$, respectively. In what follows, a Hamiltonian is always a function that admits a representation

$$H(q,p)=\omega\cdot p+\sum_{(\nu,k)\in I}H_{\nu,k}\cos(\nu\cdot q)(\Omega\cdot p)^k\,, \tag{1.2}$$

This work was supported in part by the National Science Foundation under Grants No. DMS-9705095 and DMS-0088935.

where $I = \mathcal{V} \times \mathbb{N}$ and $\mathcal{V} = \{\nu \in \mathbb{Z} \times \mathbb{N} : \nu_2 > 0 \text{ or } \nu_1 \geq 0\}$. We call H *real* if the coefficients $H_{\nu,k}$ are all real. Given a set $J \subseteq I$, denote by $\mathbb{J}H$ the Hamiltonian obtained from H by restricting the sum in (1.2) to the index set J. In particular, consider the set

$$I^{+} = \bigl\{(\nu, k) \in I : |\omega \cdot \nu| \leq \sigma |\Omega \cdot \nu| \text{ or } |\omega \cdot \nu| < \kappa k \bigr\}\,, \tag{1.3}$$

and the corresponding projection $\mathbb{I}^{+}$, where σ and κ are fixed but arbitrary positive real numbers. The Hamiltonian $\mathbb{I}^{+}H$ will be called the *resonant* part of H, and $\mathbb{I}^{-}H = H - \mathbb{I}^{+}H$ the *non-resonant* part. The (non)resonant part of a function

$$F(q, p) = \sum_{(\nu,k)\in I} \bigl[F_{\nu,k} \cos(\nu \cdot q) + F'_{\nu,k} \sin(\nu \cdot q)\bigr] (\Omega \cdot p)^k \tag{1.4}$$

is defined similarly. In addition, if the coefficients $F'_{\nu,k}$ or $F_{\nu,k}$ are all zero, we will call F an *even* or *odd* function, respectively.

As mentioned above, "normalizing" a Hamiltonian involves the elimination of its non-resonant part. This will be done by means of a canonical change of variables. A canonical transformations $U_\phi : (q, p) \mapsto (q + Q, p + P)$, associated with a generating function ϕ, is defined implicitly by the equation

$$\begin{aligned} Q(q, p) &= \nabla_p \phi\bigl(q, p + P(q, p)\bigr)\,, \\ P(q, p) &= -\nabla_q \phi\bigl(q, p + P(q, p)\bigr)\,. \end{aligned} \tag{1.5}$$

In what follows, a generating function ϕ is always assumed to be an odd function of the form (1.4). Notice that if ϕ is small, then $H \circ U_\phi = H + \{H, \phi\} +$ "higher order terms".

We are now ready to give a formal definition of our RG transformation $\mathcal{R}$. In what follows, two Hamiltonians that differ only by an additive constant will be considered equivalent. (We will do this by simply ignoring such constants.) Let ϕ_0 be a fixed generating function, to be chosen later. Define $\mathcal{T}_\mu(q, p) = \bigl(Tq, \mu T^{-1}p\bigr)$ and $\mu_0 = \vartheta^{-3}$. Then we set

$$\mathcal{R} = \mathcal{N} \circ \mathcal{L} \circ \mathcal{S}\,, \tag{1.6}$$

where

$$\begin{aligned} (\mathcal{S}H)(q, p) &= c_H H\bigl(q, p/c_H\bigr)\,, \\ \mathcal{L}H &= \tau_H \mu_0^{-1} H \circ \mathcal{T}_{\mu_0} \circ U_{\phi_0}\,, \\ \mathcal{N}(H) &= H \circ \mathcal{U}_H\,. \end{aligned} \tag{1.7}$$

Here, $c_H = 2H_{0,2}$, $\tau_H = \vartheta$, and $\mathcal{U}_H$ is a canonical transformation, determined by solving the equation

$$\mathbb{I}^{-}\bigl(H \circ \mathcal{U}_H\bigr) = 0\,. \tag{1.8}$$

If c_H is equal to zero, or if (1.8) cannot be solved in the way described below, then $\mathcal{R}(H)$ remains undefined.

Our choice $\tau_H = \vartheta$ guarantees that if H is of the form (1.2), then so is $\mathcal{L}H$, i.e., the coefficient of $\omega \cdot p$ remains 1. The constant c_H is determined by the condition $(\mathcal{S}H)_{0,2} =$

1/2. Another H-dependent constant could be added to $\mathcal{R}(H)$, in order to guarantee e.g. that $(\mathcal{R}(H))_{0,0} = 0$. These (somewhat arbitrary) normalization conditions ensure that $\mathcal{R}$ contracts along the orbits of scaling transformations. Similarly, we expect the normalization map $\mathcal{N}$ to make $\mathcal{R}$ a contraction along the orbits of coordinate changes $H \mapsto H \circ U_\phi$. We will comment later on the corresponding choice of normal form (resonant subspace).

Equation (1.8) will be solved iteratively. We now give a formal description of the first step. Assume that $H = H_1 + f_1$ is a small perturbation of some resonant Hamiltonian H_1. If we regard (1.8) as an equation for the generating function of $\mathcal{U}_H$, then its first order approximation is

$$\mathbb{I}^{-}\bigl(f_1 + \{H_1, \phi_1\}\bigr) = 0\,. \tag{1.9}$$

In order to get a unique solution, we also impose the condition that ϕ_1 be non-resonant. Now define H_2 and f_2 to be the resonant and nonresonant part, respectively, of $H \circ U_{\phi_1}$. Having solved equation (1.8) to first order in f_1, we can expect f_2 to be much smaller than f_1. Thus, we iterate the procedure, and define $\mathcal{U}_H = U_{\phi_1} \circ U_{\phi_2} \circ \ldots$, if the iteration converges.

The Hamiltonian H_1 that will be used in this procedure is an approximate fixed point of $\mathcal{R}$. To be more precise, we have determined numerically a resonant Hamiltonian H_1, together with a generating function ϕ_0, in such a way that $\mathcal{L}\mathcal{S}H_1 \approx H_1$. In what follows, we will limit the domain of $\mathcal{R}$ to a small neighborhood of H_1, and thus $\mathcal{N}$ needs to be defined only near H_1. This is why we can obtain $\mathcal{U}_H$ by perturbation theory. Notice that the combination of $\mathcal{N}$ with the other parts of $\mathcal{R}$ involves a change of coordinates with $U_{\phi_0} \circ U_{\phi_1} \circ U_{\phi_2} \circ \ldots$, where all generating functions except ϕ_0 depend on the argument of $\mathcal{R}$. The transformation U_{ϕ_0} can be viewed as a "common factor", and we have chosen to incorporate this factor into the definition of $\mathcal{L}$, instead of making it part of $\mathcal{U}_H$. For a RG analysis of near-integrable Hamiltonians [22,25], this common factor is simply the identity.

Definition 1.1. *Given any vector $\rho = (\rho_1, \rho_2)$ with positive components, and any $\alpha \geq 0$, define $\mathcal{F}_{\rho,\alpha}$ to be the (complex) Banach space of functions (1.4), for which the norm*

$$\|F\|_{\rho,\alpha} = \sum_{(\nu,k)\in I} \bigl(|F_{\nu,k}| + |F'_{\nu,k}|\bigr) e^{\alpha|\omega\cdot\nu|} \cosh(\rho_1 \Omega \cdot \nu) \rho_2^k \tag{1.10}$$

is finite. (We assume of course that $F'_{0,k} = 0$ for all k.) The odd an even subspaces of $\mathcal{F}_{\rho,\alpha}$ will be denoted by $\mathcal{A}_{\rho,\alpha}$ and $\mathcal{B}_{\rho,\alpha}$, respectively. Furthermore, the affine space $H_0 + \mathcal{B}_{\rho,\alpha}$, where $H_0(q,p) = \omega \cdot p$, will be denoted by $\mathcal{B}'_{\rho,\alpha}$. In order to simplify notation, we will drop the subscript α, whenever $\alpha = 0$.

If $\alpha > 0$, then the functions in $\mathcal{F}_{\rho,\alpha}$ are analytic on the complex domain $D_{\rho,\alpha}$, defined by $|\mathrm{Im} x| < \alpha$, $|\mathrm{Im} y| < \rho_1$, and $|z| < \rho_2$, where

$$x = \omega' \cdot q\,, \qquad y = \Omega' \cdot q\,, \qquad z = \Omega \cdot p\,, \tag{1.11}$$

and where ω' and Ω' are constant multiples of ω and Ω, respectively, such that $\omega \cdot \omega' = \Omega \cdot \Omega' = 1$. The functions in $\mathcal{F}_\rho$ are defined in the complex domain D_ρ, characterized by $\mathrm{Im} x = 0$, $|\mathrm{Im} y| < \rho_1$, and $|z| < \rho_2$, where they are continuous, and analytic in y and z.

The main result of this paper is the following.

Theorem 1.2. *There exists an even real resonant Fourier-Taylor polynomial h_1, an odd real Fourier-Taylor polynomial ϕ_0, and a choice of the parameters σ, κ, ρ, such that $\mathcal{R}$ defines a compact analytic map, from some open neighborhood B of $H_1 = H_0 + h_1$ in $\mathcal{B}'_\rho$, to the resonant subspace of $\mathcal{B}'_\rho$. This map $\mathcal{R}$ has a unique fixed point H_* in B, which is real, belongs to $\mathcal{B}'_{\rho,\alpha}$ for some $\alpha > 0$, and is non-trivial, in the sense that $c_{H_*} > 1$.*

Our proof of this theorem is computer-assisted. As a by-product, we also obtain rigorous bounds on the Fourier-Taylor coefficients of H_*, and other quantities of interest. For the normalization constant c_{H_*}, we obtain the bound $c_{H_*} = 1.024332969\ldots$, which is consistent with the numerical results in [3,21]. On a purely numerical level, the fixed point H_*, and the corresponding scaling map $\Lambda_* = \mathcal{T}_{\mu_*} \circ U_{\phi_0} \circ \mathcal{U}_{\mathcal{L}H_*}$, also satisfy a set of conditions that was shown in [30] to imply the existence of a "critical" invariant torus. Here, $\mu_* = \mu_0 / c_{H_*}$. We did not (yet) verify these conditions rigorously, since they require bounds in different function spaces.

We note that the choice of the resonant projection $\mathbb{I}^+$ is somewhat arbitrary, like other normalization conditions. The properties of $\mathbb{I}^\pm$ that play an important role in our analysis are that resonant Hamiltonians in $\mathcal{B}_\rho$ are analytic (see Proposition 3.4), and that the condition (1.9) defines a continuous linear map $f_1 \mapsto \phi_1$ on the non-resonant subspace of $\mathcal{B}_\rho$.

Theorem 1.2 is proved by converting the fixed point problem for $\mathcal{R}$, to a fixed point problem for an approximate Newton map $\mathcal{M}$ associated with $\mathcal{R}$. To be more precise, we choose an approximate inverse M of $\mathbb{I} - D\mathcal{R}(H_1)$, and define

$$\mathcal{M}(h) = h + \mathcal{R}(H_1 + Mh) - (H_1 + Mh)\,. \tag{1.12}$$

The goal is then to show that $\mathcal{M}$ is a contraction on a small ball centered at zero in $\mathcal{B}_\rho$. This is a common strategy in computer-assisted proofs; see e.g. [32] and references therein. It usually involves explicit bounds on a finite dimensional truncation of $\mathcal{M}$, and error bounds for the terms that were truncated. In the case at hand, however, working with a truncation of the full map $\mathcal{M}$ turned out to be computationally prohibitive. We solve this problem by using the fact that $\mathcal{M}$ only needs to be estimated on a small open set. More specifically, we approximate the most complex part of $\mathcal{M}$, which is the map $\mathcal{N}$, by a much simpler affine map $\mathcal{N}_1$, and then estimate the difference between the two.

Let H_1 be a fixed non-resonant Hamiltonian, and define

$$\mathcal{N}_1(H_1 + f_1) = H_1 + \mathbb{I}^+\bigl(f_1 + \{H_1, \phi_1\}\bigr)\,, \tag{1.13}$$

where ϕ_1 is the solution of equation (1.9). Then $\mathcal{N}_1$ is an approximation of $\mathcal{N}$, in the sense that $\mathcal{N}(H_1 + f_1) - \mathcal{N}_1(H_1 + f_1)$ vanishes to first order in f_1. The map $\mathcal{N}_1$ can now be used to define an approximate RG transformation

$$\mathcal{R}_1 = \mathcal{N}_1 \circ \mathcal{L} \circ \mathcal{S}\,. \tag{1.14}$$

Notice that $\mathcal{R}_1$ is nonlinear, but only due to the trivial scaling transformation $\mathcal{S}$. In order to estimate the nonlinear term, let

$$\widetilde{\mathcal{S}}(H_1, f) = \mathcal{S}(H_1 + f) - \mathcal{S}(H_1) - D\mathcal{S}(H_1) f\,. \tag{1.15}$$

Then, by using that $\mathcal{N}_1$ is affine, we can decompose $\mathcal{R}_1$ as follows:

$$\begin{aligned}\mathcal{R}_1(H_1+f) &= \mathcal{N}_1\bigl(\mathcal{L}\mathcal{S}(H_1)+\mathcal{L}D\mathcal{S}(H_1)f+\mathcal{L}\widetilde{\mathcal{S}}(H_1,f)\bigr)\\ &= \mathcal{R}_1(H_1)+D\mathcal{R}_1(H_1)f+\widetilde{\mathcal{R}}(H_1,f)\,,\end{aligned} \tag{1.16}$$

where

$$\widetilde{\mathcal{R}}(H_1,f) = D\mathcal{N}_1(H_1)\mathcal{L}\widetilde{\mathcal{S}}(H_1,f)\,. \tag{1.17}$$

This leads to the following representation for the map $\mathcal{M}$.

$$\begin{aligned}\mathcal{M}(h) = \bigl(\mathcal{R}_1(H_1)-H_1\bigr)&+\bigl[\mathbb{I}-\bigl(\mathbb{I}-D\mathcal{R}_1(H_1)\bigr)M\bigr]h\\ &+\widetilde{\mathcal{R}}(H_1,Mh)+\bigl(\mathcal{R}-\mathcal{R}_1\bigr)(H_1+Mh)\,.\end{aligned} \tag{1.18}$$

The term in square brackets defines a linear operator on $\mathcal{B}_\rho$, which we will show to be of norm $K<1$. Since the Hamiltonian H_1 is chosen to be an approximate fixed point of $\mathcal{R}$, and thus of $\mathcal{R}_1$, the term $\mathcal{R}_1(H_1)-H_1$ in equation (1.18) is small, say of norm $\varepsilon>0$. If $\mathcal{M}$ is restricted to a ball B of radius $R\varepsilon$, centered at $h=0$, then formally, the nonlinear (last two) terms on the right hand side of (1.18) are of the oder $\mathcal{O}(\varepsilon^2)$. Thus, if $(1-K)R>1$, and $\varepsilon>0$ is sufficiently small, we expect $\mathcal{M}$ to map the ball B into itself. In order to prove that $\mathcal{M}$ is a contraction, we also need to estimate the derivatives of the nonlinear terms. But by using analyticity, it suffices to estimate their norm, on a slightly larger domain, say $2B$. Formally, this yields a bound of the order $\mathcal{O}(\varepsilon)$. What needs to be proved is that this bound is less than $1-K$.

The remaining part of this paper is organized as follows. In the next two sections, we estimate the maps $\mathcal{N}_1$ and $\mathcal{R}_1$. Section 4 describes the iterative solution of equation (1.8), which defines $\mathcal{N}$. The RG transformation $\mathcal{R}$, and the contraction $\mathcal{M}$, are discussed in Section 5. In particular, we show that Theorem 1.2 follows from a set of inequalities, which are formulated in Theorem 5.2. Section 6 contains an outline of our (computer-assisted) proof of Theorem 5.2. The details of this proof are written in the programming language Ada95 [33], and can be found in Section.7 [31].

2. The transformation $\mathcal{N}_1$

In this section, we discuss the map $\mathcal{N}_1$, and give some basic estimates for compositions and derivatives of functions in $\mathcal{F}_\rho$. Here, and in what follows, ρ is a fixed but arbitrary domain parameter in $\mathbb{R}^2_+$. If L is a bounded linear operator on $\mathcal{F}_\rho$, then $\|L\|_\rho$ denotes the operator norm of L. Define

$$\begin{aligned}\partial_x = \omega\cdot\nabla_q\,,\qquad \partial_y &= \Omega\cdot\nabla_q\,,\qquad \partial_z = \Omega'\cdot\nabla_p\,,\\ \mathcal{D}_y &= \partial_x^{-1}\partial_y\,,\qquad \mathcal{D}_z = \partial_x^{-1}\partial_z\,.\end{aligned} \tag{2.1}$$

Proposition 2.1. *$\mathcal{D}_y\mathbb{I}^-$ and $\mathcal{D}_z\mathbb{I}^-$ are bounded linear operators on $\mathcal{F}_\rho$, with norms*

$$\bigl\|\mathcal{D}_y\mathbb{I}^-\bigr\|_\rho \le \sigma^{-1}\,,\qquad \bigl\|\mathcal{D}_z\mathbb{I}^-\bigr\|_\rho \le (\kappa\rho_2)^{-1}\,. \tag{2.2}$$

Proof. Let $\psi = \cos(\nu \cdot Q)Z^k$ or $\psi = \sin(\nu \cdot Q)Z^k$, with (v, k) belonging to the complement of I^+ in I. Here, $Q(q,p) = q$, and $Z(q,p) = \Omega \cdot p$. By using that $|\Omega \cdot \nu|/|\omega \cdot \nu| < \sigma^{-1}$ and $k/|\omega \cdot \nu| < \kappa^{-1}$, we immediately get the bounds $\|\mathcal{D}_y\psi\|_\rho \le \sigma^{-1}\|\psi\|_\rho$ and $\|\mathcal{D}_z\psi\|_\rho \le (\kappa\rho_2)^{-1}\|\psi\|_\rho$. These bounds extend by linearity to all of $\mathbb{I}^-\mathcal{F}_\rho$, which proves the claim. **QED**

Next, we consider the map $\mathcal{N}_1$ described in equation (1.13). Consider a function h_1 whose partial derivatives $\partial_z h_1$ and $\partial_y h_1$ belong to $\mathcal{F}_\rho$, and define

$$\mathcal{D}(h_1) = \bigl[(\partial_z h_1)\mathcal{D}_y - (\partial_y h_1)\mathcal{D}_z\bigr]\mathbb{I}^- . \tag{2.3}$$

Let $H_1 = H_0 + h_1$, and assume that

$$\|\mathcal{D}(h_1)\|_\rho \le A < 1 . \tag{2.4}$$

Then for every function $f_1 \in \mathcal{B}_\rho$, the equation (1.9) has a solution given by

$$\phi_1 = \partial_x^{-1}\psi_1 , \qquad \psi_1 = \bigl[\mathbb{I} + \mathbb{I}^-\mathcal{D}(h_1)\bigr]^{-1}\mathbb{I}^- f_1 . \tag{2.5}$$

Clearly, the function $\psi_1 \in \mathbb{I}^-\mathcal{B}_\rho$ is uniquely determined by f_1 and satisfies the bound

$$\|\psi_1\|_\rho \le (1 - A)^{-1}\|f_1\|_\rho . \tag{2.6}$$

The map $\mathcal{N}_1$ can now be written as follows:

$$\mathcal{N}_1(H_1, f_1) = H_1 + \mathbb{I}^+ f_1 - \mathbb{I}^+\mathcal{D}(h_1)\psi_1 . \tag{2.7}$$

Here, we have re-defined $\mathcal{N}_1$ as a function of two variables, in order to emphasize the dependence on the choice of H_1. Estimating $\mathcal{N}_1$ requires a bound on the last term in equation (2.7). We will do this by using a numerical approximation $\psi \in \mathbb{I}^-\mathcal{B}_\rho$ for the function ψ_1 defined in (2.5). Since

$$\psi_1 - \psi = \bigl[\mathbb{I} + \mathbb{I}^-\mathcal{D}(h_1)\bigr]^{-1}\bigl(\mathbb{I}^- f_1 - \mathbb{I}^-[\mathbb{I} + \mathcal{D}(h_1)]\psi\bigr) , \tag{2.8}$$

we have immediately the following

Proposition 2.2. *If ψ is any non-resonant function in $\mathcal{B}_\rho$, then*

$$\bigl\|\mathbb{I}^+\mathcal{D}(h_1)\psi_1 - \mathbb{I}^+\mathcal{D}(h_1)\psi\bigr\|_\rho \le \frac{A}{1 - A}\bigl\|\mathbb{I}^- f_1 - \mathbb{I}^-[\mathbb{I} + \mathcal{D}(h_1)]\psi\bigr\|_\rho . \tag{2.9}$$

This inequality shows in particular (setting $\psi = 0$) that $f \mapsto \mathcal{N}_1(H_1, f)$ is a continuous linear operator on $\mathcal{B}_\rho$.

We conclude this section with a bound on $\mathcal{D}(h)$, and some other basic estimates that will be used later. Define

$$\Gamma(s) = \Gamma(s, 0, 0), \qquad \Gamma(s, t, v) = \sup_{u \ge v} \frac{\cosh(su)}{\cosh(u)} e^{(1-s-t)u} , \qquad s, t, v \ge 0 . \tag{2.10}$$

Proposition 2.3. *Consider functions $f, G \in \mathcal{F}_\rho$ and $h \in \mathcal{F}_{\rho'}$, and the change of variables $U(q,p) = (q + f(q,p)\Omega, G(q,p)\Omega')$. Then*
(a) $\|fG\|_\rho \le \|f\|_\rho \|G\|_\rho$.
(b) $\|h \circ U\|_\rho \le \Gamma(\rho_1/\rho_1')\|h\|_{\rho'}$, *if* $\rho_1' \ge \rho_1 + \|f\|_\rho$ *and* $\rho_2' \ge \|g\|_\rho$.
(c) $\|\mathcal{D}(h)\|_\rho \le s_0^{-1}\Gamma(\rho_1/\rho_1')\|h\|_{\rho'}$, *if* $\rho_1' \ge \rho_1 + s_0/(\kappa\rho_2)$ *and* $\rho_2' \ge \rho_2 + s_0/\sigma$ *and* $s_0 > 0$.
(d) $\|\mathcal{D}(h)\|_\rho \le \sigma^{-1}\|\partial_z h\|_{\rho'} + (\kappa\rho_2)^{-1}\|\partial_y h\|_{\rho'}$, *if* $\rho_1' \ge \rho_1$ *and* $\rho_2' \ge \rho_2$.
(e) $\|\partial_y h\|_\rho \le (e\delta_1)^{-1}\Gamma(\rho_1/\rho_1')\|h\|_{\rho'}$, *if* $\rho_1' = \rho_1 + \delta_1 > \rho_1$ *and* $\rho_2' \ge \rho_2$.
(f) $\|\partial_z h\|_\rho \le \delta_2^{-1}\|h\|_{\rho'}$, *if* $\rho_1' \ge \rho_1$ *and* $\rho_2' = \rho_2 + \delta_2 > \rho_2$.

Proof. The first and last bounds are straightforward to check. In order to prove (*b*), consider first the case $h = \cos(\nu \cdot Q)Z^k$, where $Q(q,p) = q$ and $Z(q,p) = \Omega \cdot p$. Then

$$\begin{aligned} \|h \circ U\|_\rho &= \bigl\|\bigl[\cos(\nu\cdot Q)\cos(\nu\cdot\Omega f) - \sin(\nu\cdot Q)\sin(\nu\cdot\Omega f)\bigr]g^k\bigr\|_\rho \\ &\le \cosh(\rho_1\Omega\cdot\nu)\bigl[\cosh\bigl(\|\nu\cdot\Omega f\|_\rho\bigr) + \sinh\bigl(\|\nu\cdot\Omega f\|_\rho\bigr)\bigr]\|g\|_\rho^k \\ &\le \frac{\cosh(\rho_1\Omega\cdot\nu)}{\cosh(\rho_1'\Omega\cdot\nu)} e^{|\Omega\cdot\nu|\|f\|_\rho}\cosh(\rho_1'\Omega\cdot\nu)\bigl(\rho_2'\bigr)^k \\ &\le \Gamma(\rho_1/\rho_1')\|h\|_{\rho'} , \end{aligned} \tag{2.11}$$

and an analogous bound holds for $h = \sin(\nu \cdot Q)Z^k$. Since we are dealing with weighted ℓ^1 spaces, the claim (*b*) now follows by taking linear combinations and limits.

The bound (*c*) is obtained from Proposition 2.1 and (*b*), by taking a general non-resonant ψ in the unit ball of $\mathcal{B}_\rho$, and estimating the derivative in

$$\bigl\|\mathcal{D}(h_1)\psi\bigr\|_\rho = \left\|\frac{d}{ds}h\bigl(Q + s[\mathcal{D}_z\psi]\Omega, [Z - s\mathcal{D}_y\psi]\Omega'\bigr)\Bigr|_{s=0}\right\|_\rho , \tag{2.12}$$

using Cauchy's formula with contour $|s| = s_0$. The inequality (*d*) follows immediately from Proposition 2.1 and (*a*).

In order to prove (*e*), consider again $h = \cos(\nu \cdot Q)Z^k$. By using that $|\Omega \cdot \nu| \le (e\delta_1)^{-1}\exp(\delta_1|\Omega\cdot\nu|)$, we obtain the bound

$$\begin{aligned} \|\partial_y h\|_\rho &= \|(\Omega\cdot\nu)\sin(\nu\cdot Q)Z^k\|_\rho = |\Omega\cdot\nu|\cosh(\rho_1\Omega\cdot\nu)\rho_2^k \\ &\le \frac{\cosh(\rho_1\Omega\cdot\nu)}{\cosh(\rho_1'\Omega\cdot\nu)}|\Omega\cdot\nu|\|h\|_{\rho'} \le (e\delta_1)^{-1}\Gamma(\rho_1/\rho_1')\|h\|_{\rho'} . \end{aligned} \tag{2.13}$$

Similarly for $h = \sin(\nu \cdot Q)Z^k$. The claim (*e*) now follows by taking linear combinations and limits.
QED

These bounds will typically be used for ρ' close to ρ. In order to give an idea of the size of $\Gamma(s)$ in such cases, we note that for $0 < s \le 1$,

$$\Gamma(s) \le s^{-\gamma} , \qquad \gamma = \max_{u\ge 0} u[1 - \tanh(u)] = 0.2784\ldots , \tag{2.14}$$

as was shown in [30].

Remark. The results of this section hold more generally in the spaces $\mathcal{F}_{\rho,\alpha}$, for any fixed $\alpha \ge 0$. The only change (besides notation) are norm factors $e^{\alpha|\omega\cdot\nu|}$ in the intermediate steps of the two estimates (2.11) and (2.13).

3. The transformation $\mathcal{R}_1$

We will define $\mathcal{R}_1$ as a map from $\mathcal{B}_\rho$ to $\mathcal{B}_{\rho''}$, with $\rho'' > \rho$. One of the ingredients of $\mathcal{R}_1$ is a canonical change of variables $H \mapsto H \circ U_{\phi_0}$, determined via equation (1.5) from its generating function ϕ_0. We assume that $\phi_0(q,p)$ is odd, and that the first partial derivatives of ψ_0, including $\psi_0 = \partial_x \phi_0$, belong to $\mathcal{F}_{\rho'}$, for ρ' sufficiently large (explicit conditions will be given below).

The second equation in (1.5) yields the following equation for $g = \Omega \cdot P$.

$$g(q,p) = -[\mathcal{D}_y \psi_0]\bigl(q, p + g(q,p)\Omega'\bigr)\,. \tag{3.1}$$

We note that the argument $p+g(q,p)\Omega'$ above could be replaced by $G(q,p)\Omega'$ with $G(q,p) = \Omega \cdot p + g(q,p)$. Thus, Proposition 2.3.b applies to compositions of this type. In order to solve equation (3.1), we first determine an approximate solution $g_0 \in \mathcal{B}_{\rho''}$, and then estimate the error. To this end, define

$$a = \bigl\|[\partial_z \mathcal{D}_y \psi_0](.,. + g_0\Omega')\bigr\|_{\rho''}\,, \qquad \varepsilon = \bigl\|[\mathcal{D}_y \psi_0](.,. + g_0\Omega') + g_0\bigr\|_{\rho''}\,, \tag{3.2}$$

and if $a < 1$, choose a real number b such that

$$b \ge \bigl\|\bigl[\partial_z^2 \mathcal{D}_y \psi_0\bigr](.,. + g\Omega')\bigr\|_{\rho''} \quad \text{whenever} \quad \|g - g_0\|_{\rho''} \le 2\varepsilon/(1-a)\,. \tag{3.3}$$

Proposition 3.1. *If $a < 1$ and $4b\varepsilon \le (1-a)^2$, then the equation (3.1) has a unique solution in $g \in \mathcal{B}_{\rho''}$ within a distance $r = 2\varepsilon/(1-a)$ of g_0, assuming that $\mathcal{D}_y\psi_0$ and $\mathcal{D}_z\psi_0$ belong to $\mathcal{A}_{\rho'}$, with $\rho_1' = \rho_1''$ and $\rho_2' > \rho_2'' + \|g_0\|_{\rho''} + r$.*

Proof. G solves equation (3.1) if and only if $h = g - g_0$ is a fixed point of the map $\mathcal{K}$, defined by

$$\mathcal{K}(h) = h_0 + K(h)h\,, \qquad K(h) = -\int_0^1 [\partial_z \mathcal{D}_y \psi_0]\bigl(.,. + (g_0 + sh)\Omega'\bigr) ds\,, \tag{3.4}$$

where $h_0 = -[\mathcal{D}_y \psi_0](.,. + g_0\Omega') - g_0$. Denote by B the closed ball of radius r in $\mathcal{B}_{\rho''}$, centered at the origin. By the mean value theorem,

$$\|K(h) - K(h')\|_{\rho''} \le \frac{b}{2}\|h - h'\|_{\rho''}\,, \tag{3.5}$$

whenever $h, h' \in B$, and by taking $h' = 0$, we obtain the bound $\|K(h)\|_{\rho''} \le a + \frac{b}{2}\|h\|_{\rho''}$. Thus, if h and h' belong to B, then

$$\begin{aligned}
\|\mathcal{K}(h) - \mathcal{K}(h')\|_{\rho''} &= \|[K(h) - K(h')]h + K(h')[h - h']\|_{\rho''} \\
&\le \frac{b}{2}\|h - h'\|_{\rho''}\|h\|_{\rho''} + \left(a + \frac{b}{2}\|h'\|_{\rho''}\right)\|h - h'\|_{\rho''} \\
&\le (a + br)\|h - h'\|_{\rho''} \le \frac{a+1}{2}\|h - h'\|_{\rho''}\,.
\end{aligned} \tag{3.6}$$

By taking again $h' = 0$, we find that $\|\mathcal{K}(h)\|_{\rho''} \le r$. Thus, $\mathcal{K}$ is a contraction on B, and the assertion follows.

For reference later on, we note that the fixed point $h \in B$ of $\mathcal{K}$ satisfies the equation $h = h_0/(1 - K(h))$. This leads to the bound

$$\|h\|_{\rho''} \le \varepsilon \big\|1/\big(1 - K(h)\big)\big\|_{\rho''} . \tag{3.7}$$

QED

In what follows, we assume that equation (3.1) has a solution $g \in \mathcal{B}_{\rho''}$. The canonical transformation U_{ϕ_0} can now be written explicitly as

$$U_{\phi_0}(q,p) = \Big(q + [\mathcal{D}_z\psi_0]\big(q, p + g(q,p)\Omega'\big)\Omega,\, p + g(q,p)\Omega' + \gamma(q,p)\omega'\Big) , \tag{3.8}$$

with $\gamma(q,p) = -\psi_0(q, p + g(q,p)\Omega')$. As a result, we have the following expression for the linear transformation $\mathcal{L}$ defined in (1.7). Let $G(q,p) = \Omega \cdot p + g(q,p)$. Then

$$(\mathcal{L}H)(q,p) = \omega \cdot p - \psi_0\big(q, G(q,p)\Omega'\big) + (Lh)(q,p) , \tag{3.9}$$

where $h = H - H_0$ and

$$(Lh)(q,p) = \vartheta^4 h\big(Tq - \vartheta^{-1}[\mathcal{D}_z\psi_0](q, G(q,p)\Omega')\Omega, -\vartheta^{-2}G(q,p)\Omega'\big) . \tag{3.10}$$

Define

$$b_2 = \|G\|_{\rho''} , \qquad b_1 = \|\mathcal{D}_z\psi_0\|_{(\rho_1'',b_2)} , \tag{3.11}$$

assuming that $\mathcal{D}_z\psi_0$ belongs to $\mathcal{A}_{(\rho_1'',b_2)}$.

Proposition 3.2. *Let $h \in \mathcal{B}_{\rho'}$ with $\rho_1' > \vartheta^{-1}(\rho_1'' + b_1)$ and $\rho_2' = \vartheta^{-2}b_2$. Then Lh belongs to $\mathcal{B}_{\rho''}$, and if the Fourier-Taylor series (1.4) of h only includes frequencies ν satisfying a bound $|\Omega \cdot \nu| \ge \lambda \ge 0$, then*

$$\|Lh\|_{\rho''} \le \vartheta^4 \Gamma\big(\rho_1''/\rho_1', 1 - (\rho_1'' + b_1)/(\vartheta\rho_1'), \lambda/\rho_1'\big)\|h\|_{\rho'} . \tag{3.12}$$

Proof. Consider first $h = \cos(\nu \cdot Q)Z^k$, where $Q(q,p) = q$ and $Z(q,p) = \Omega \cdot p$. Define $f(q,p) = [\mathcal{D}_z\psi_0](q, G(q,p)\Omega')$. This function f belongs to $\mathcal{A}_{\rho''}$ and is bounded in norm by b_1, as Proposition 2.3.b shows. Thus, we have the bound

$$\begin{aligned}
\|Lh\|_{\rho''} &= \Big\|\vartheta^4\big[\cos\big((T\nu)\cdot Q\big)\cos\big(\vartheta^{-1}(\Omega\cdot\nu)f\big) \\
&\qquad + \sin\big((T\nu)\cdot Q\big)\sin\big(\vartheta^{-1}(\Omega\cdot\nu)f\big)\big]\big(\vartheta^{-2}G\big)^k\Big\|_{\rho''} \\
&\le \vartheta^4\cosh\big({\rho''}_1\vartheta^{-1}\Omega\cdot\nu\big)e^{\vartheta^{-1}|\Omega\cdot\nu|\|f\|_{\rho''}}\big(\vartheta^{-2}\|G\|_{\rho''}\big)^k \\
&\le \vartheta^4\|h\|_{\rho'}\frac{\cosh(\rho_1''\vartheta^{-1}\Omega\cdot\nu)}{\cosh(\rho_1'\Omega\cdot\nu)}e^{\vartheta^{-1}|\Omega\cdot\nu|b_1} ,
\end{aligned}$$

which implies (3.12). The generalization to arbitrary $h \in \mathcal{B}_{\rho'}$ follows by taking linear combinations and limits. **QED**

Our next goal is to define $\mathcal{R}_1$ as the composition

$$\mathcal{R}_1 : \mathcal{B}'_\rho \xrightarrow{\mathcal{S}} \mathcal{B}'_{\rho'} \xrightarrow{\mathcal{L}} \mathcal{B}'_{\rho''} \xrightarrow{\mathcal{N}_1} \mathcal{B}'_{\rho''}\,, \tag{3.13}$$

with some additional restriction on the domain. Here, $\rho' = (\rho_1, \vartheta^{-2}b_2)$. By combining Proposition 3.2 and Proposition 2.2, we obtain immediately the following

Corollary 3.3. *Let $\rho'' > \rho > 0$. Denote by B the set of all functions $H \in \mathcal{B}'_\rho$ satisfying*

$$2|H_{0,2}|\rho_2 > \vartheta^{-2}b_2\,. \tag{3.14}$$

Assume that $b_1 < \vartheta\rho_1 - \rho''_1$, and that $H_1 = H_0 + h_1$ is a Hamiltonian in B for which $\|\mathcal{D}(h_1)\|_{\rho''}$ is finite, and less than 1. Then $\mathcal{R}_1 = \mathcal{N}_1 \circ \mathcal{L} \circ \mathcal{S}$ is well defined and analytic as a map from B to $\mathcal{B}'_{\rho''}$.

The analyticity of $\mathcal{R}_1$ follows from the analyticity of the map $\mathcal{S}$, together with the fact that $\mathcal{N}_1$ and $\mathcal{L}$ are affine and continuous on the spaces indicated in (3.13).

We note that $\mathcal{R}_1(H)$ is always resonant. The following fact shows that, as a consequence, $\mathcal{R}_1$ is analyticity improving, and thus also compact, when considered as a map from $\mathcal{B}'_\rho$ to itself.

Proposition 3.4. *If $\rho^* = (\rho_1 + \alpha\sigma, e^{\alpha\kappa}\rho_2)$ with $\alpha \ge 0$, then $\mathbb{I}^+\mathcal{F}_{\rho^*}$ is contained in $F_{\rho,\alpha}$.*

This proposition follows from a simple but important property of the projection $\mathbb{I}^+$: Under the given assumption,

$$e^{\alpha|\omega\cdot\nu|}\,\cosh(\rho_1\Omega\cdot\nu)\rho_2^k \le \Gamma(\rho_1/\rho_1^*)\cosh(\rho_1^*\Omega\cdot\nu)\big(\rho_2^*\big)^k\,, \qquad (\nu,k)\in I^+\,. \tag{3.15}$$

4. The transformation $\mathcal{N}$

As mentioned in the introduction, the transformation $\mathcal{U}_H$ is defined as a composition

$$\mathcal{U}_H = U_{\phi_1} \circ U_{\phi_2} \circ U_{\phi_3} \circ \ldots\,, \tag{4.1}$$

where U_{ϕ_1} is determined in such as way that $H \mapsto H \circ U_{\phi_1}$ eliminates the leading non-resonant part of $H = H_1 + f_1$, and the other canonical transformations U_{ϕ_k} are obtained by iterating the same procedure. The precise definition of ϕ_1 is given by equation (2.5).

Proposition 4.1. *Let $\rho' = (\rho_1, \rho_2 + \delta_2)$ with $\delta_2 > 0$. Let ψ_1 be a non-resonant function in $\mathcal{B}_{\rho'}$ which satisfies*

$$(\sigma\delta_2)^{-1}\|\psi_1\|_{\rho'} \le \tau < 1/2\,. \tag{4.2}$$

Then the equation (3.1), with $\psi_0 = \psi_1$, has a unique solution g in the ball B of radius $\delta_2/2$ in $\mathcal{B}_\rho$, centered at the origin, and

$$\|g\|_\rho \le \|\mathcal{D}_y\psi_1\|_{\rho'} \le \sigma^{-1}\|\psi_1\|_{\rho'}\,. \tag{4.3}$$

Proof. Let $r = (\rho_1, \rho_2 + \delta_2/2)$. Denote by $\mathcal{K}$ the map that assigns to $g \in B$ the function on the right hand side of (3.1), with $\psi_0 = \psi_1$. Then the bound (2.2) and Proposition 2.3 imply that

$$\begin{aligned} \|\mathcal{K}(g)\|_\rho &\le \|\mathcal{D}_y\psi_1\|_{\rho'} \le \sigma^{-1}\|\psi_1\|_{\rho'} < \delta_2/2\,, \\ \|\mathcal{K}(g)-\mathcal{K}(g')\|_\rho &\le \|\partial_z\mathcal{D}_y\psi_1\|_r\|g-g'\|_\rho \le 2\tau\|g-g'\|_\rho\,, \end{aligned} \tag{4.4}$$

for any functions $g, g' \in B$. This shows that $\mathcal{K}$ is a contraction on B, and thus has a unique fixed point in B. The bound (4.3) follows from the first part of (4.4). **QED**

Proposition 4.2. *Let $\rho' = (\rho_1+\delta_1, \rho_2+\delta_2)$ with $\delta_1 \ge (\kappa\rho_2')^{-1}\sigma\delta_2 > 0$. Let h_1, f_1, and ψ_1 be functions in $\mathcal{B}_{\rho'}$, and define $H_1 = H_0 + h_1$. If ψ_1 is non-resonant and satisfies (4.2), then*

$$\begin{aligned} \bigl\|f_1\circ U_{\phi_1} - f_1\bigr\|_\rho &\le \frac{(\sigma\delta_2)^{-1}}{1-\tau}\Gamma\bigl(\rho_1/\rho_1'\bigr)\|f_1\|_{\rho'}\|\psi_1\|_{\rho'}\,, \\ \bigl\|H_1\circ U_{\phi_1} - H_1 - \{H_1,\phi_1\}\bigr\|_\rho &\le \frac{(\sigma\delta_2)^{-2}}{1-\tau}\Bigl(\sigma\delta_2 + \Gamma\bigl(\rho_1/\rho_1'\bigr)\|h_1\|_{\rho'}\Bigr)\|\psi_1\|_{\rho'}^2\,. \end{aligned} \tag{4.5}$$

Proof. The case $\psi_1 = 0$ is trivial. Assume now that $\psi_1 \ne 0$, and consider the family of functions

$$F(s)(q,p) = f_1\bigl(q + s[\mathcal{D}_z\psi_0]\bigl(q, p+g(q,p)\Omega'\bigr)\Omega,\ p + sg(q,p)\Omega'\bigr)\,,$$

where g is the solution of (3.1) described in Proposition 4.1. Let $1 < s_0 < \tau^{-1}$. If s is any complex number of modulus s_0 or less, we have

$$\begin{aligned} \|s\mathcal{D}_z\psi_1\|_{\rho'} &\le s_0(\kappa\rho_2')^{-1}\|\psi_1\|_{\rho'} < \delta_1\,, \\ \|sg\|_\rho \le \|s\mathcal{D}_y\psi_1\|_{\rho'} &\le s_0\sigma^{-1}\|\psi_1\|_{\rho'} < \delta_2\,, \end{aligned} \tag{4.6}$$

as a consequence of (2.2) and (4.3). Thus, by using Proposition 2.3.b, we obtain

$$\begin{aligned} \bigl\|f_1\circ U_{\phi_1} - f_1\bigr\|_\rho = \|F(1)-F(0)\|_\rho &= \left\|\frac{1}{2\pi i}\oint_\gamma \frac{F(s)ds}{s(s-1)}\right\|_\rho \\ &\le (s_0-1)^{-1}\Gamma\bigl(\rho_1/\rho_1'\bigr)\|f_1\|_{\rho'}\,, \end{aligned} \tag{4.7}$$

where γ denotes the complex contour $|s| = s_0$. This yields the first part of (4.5), in the limit $s_0 \to \tau^{-1}$. The second part is obtained in a similar way, by using the family

$$\begin{aligned} F(s)(q,p) &= -s\psi_1\bigl(q, p+sg(q,p)\Omega'\bigr) \\ &\quad + h_1\Bigl(q + s[\mathcal{D}_z\psi_1]\bigl(q,p+sg(q,p)\Omega'\bigr)\Omega,\ ,p - s[\mathcal{D}_y\psi_1]\bigl(q,p+sg(q,p)\Omega'\bigr)\Bigr), \end{aligned}$$

which leads to the bound

$$\begin{aligned} \bigl\|H_1\circ U_{\phi_1} - H_1 - \{H_1,\phi_1\}\bigr\|_\rho &= \|F(1)-F(0)-F'(0)\|_\rho \\ &= \left\|\frac{1}{2\pi i}\oint_\gamma \frac{F(s)ds}{s^2(s-1)}\right\|_\rho \\ &\le s_0^{-1}(s_0-1)^{-1}\Bigl(s_0\|\psi_1\|_{\rho'} + \Gamma\bigl(\rho_1/\rho_1'\bigr)\|h_1\|_{\rho'}\Bigr)\,. \end{aligned} \tag{4.8}$$

QED

In order to simplify notation later on, we will denote our starting Hamiltonian by $H_n + f_n$, instead of $H_1 + f_1$, where n is some fixed positive integer. Assume now that $\rho' = \rho + \delta$ with $\delta > 0$ component-wise, and that

$$\|f_n\|_{\rho'} \le C_n\,, \quad \|H_n - H_0\|_{\rho'} \le B_n\,, \quad \|\mathcal{D}(H_n - H_0)\|_{\rho'} \le A_n < 1\,. \tag{4.9}$$

Implicit in this assumption is that f_n , $h_n = H_n - H_0$, and $\partial_z h_n$ belong to $\mathcal{B}_{\rho'}$, and that $\partial_y h_n$ belong to $\mathcal{A}_{\rho'}$. This will not always be mentioned.

Our goal is to obtain analogous bounds, in $\mathcal{B}_\rho$, for the functions

$$H_{n+1} = \mathbb{I}^{+}\big[(H_n + f_n)\circ U_{\phi_n}\big]\,, \qquad f_{n+1} = \mathbb{I}^{-}\big[(H_n + f_n)\circ U_{\phi_n}\big]\,, \tag{4.10}$$

where ϕ_n is the generating function given by the analogue of equation (2.5), where all indices 1 have been replaced by n. To this end, define

$$\tau_n(x) = \frac{C_n}{\sigma x(1-A_n)}\,, \quad K_n(x) = \frac{(2-A_n)\sigma x + B_n}{\sigma^2\big(1-\tau_n(x)\big)(1-A_n)^2}\Gamma\big(\rho_1/\rho_1'\big)\,, \tag{4.11}$$

for $0 < x < \rho_1'$.

Proposition 4.3. *Let $0 < s < 1$. Let $k = 0$, or $k = 1$ and $\mathbb{I}^{+} f_n = 0$. Define $\gamma = \Gamma\big(\rho_1/\rho_1'\big)$ and $d = \min\{\delta_1\kappa\rho_2'/\sigma, \delta_2\}$. If $\tau_n(sd) < 1/2$, then the conditions (4.9) imply*

$$\begin{aligned}
\|f_{n+1}\|_\rho &\le C_{n+1} = K_n(d)d^{-2}C_n^2\,,\\
\|H_{n+1} - H_n\|_\rho &\le \Delta B_n = (1-A_n)^{-1}A_n^k C_n + C_{n+1}\,,\\
\big\|\mathcal{D}(H_{n+1} - H_n)\big\|_\rho &\le \Delta A_n = \frac{(1-A_n)^{-1}A_n^k C_n + K_n(sd)(sd)^{-2}C_n^2}{\sigma(1-s)d}\,\gamma\,.
\end{aligned} \tag{4.12}$$

Proof. It is useful to rewrite (4.10) as follows:

$$H_{n+1} = \mathcal{N}_1(H_n, f_n) + \mathbb{I}^{+} R(H_n, f_n)\,, \qquad f_{n+1} = \mathbb{I}^{-} R(H_n, f_n)\,, \tag{4.13}$$

where

$$R(H_n, f_n) = \big[H_n\circ U_{\phi_n} - H_n - \{H_n, \phi_n\}\big] + \big[f_n\circ U_{\phi_n} - f_n\big]\,. \tag{4.14}$$

Assume now that δ is equal to $\delta' = ((\kappa\rho_2')^{-1}\sigma d, d)$. It suffices to prove the assertion in this case, since it corresponds to replacing $\rho = \rho' - \delta$ by $\rho' - \delta'$, which only strengthens the conclusion of our proposition.

The condition $\tau_n(sd) < 1/2$, together with (2.6), guarantees that (4.2) holds, so that Proposition 4.2 applies. By adding up the two inequalities in (4.5), we obtain

$$\|R(H_n, f_n)\|_\rho \le K_n(d)d^{-2}C_n^2\,, \tag{4.15}$$

which proves the claim regarding f_{n+1}. The second bound in (4.12) is now obtained from Proposition 2.2 (with $\psi = 0$), by using the representation (2.7) of $\mathcal{N}_1$.

Consider now $r = \rho' - s\delta$. Then arguments analogous to the ones used above yield

$$\|H_{n+1} - H_n\|_r \leq (1 - A_n)^{-1} A_n^m C_n + K_n(sd)(sd)^{-2} C_n^2 . \tag{4.16}$$

The third inequality in (4.12) now follows from Proposition 2.3.c. **QED**

Since $C_{n+1} = \mathcal{O}(C_n^2)$, we expect that an iteration of $(H_n, f_n) \mapsto (H_{n+1}, f_{n+1})$, with appropriately chosen parameters

$$s = s_n , \qquad \rho' = \rho^{(n)} , \qquad \rho = \rho^{(n+1)} \tag{4.17}$$

in Proposition 4.3, will produce a sequence of Hamiltonians $H_n + f_n$, whose non-resonant part f_n tends to zero for large n. Define $A_{n+1} = A_n + \Delta A_n$ and $B_{n+1} = B_n + \Delta B_n$. In order to get good estimates on this procedure, we will determine

$$0 < s_n < 1 , \qquad 0 < \rho^* < \rho^{(n+1)} < \rho^{(n)} , \qquad 1 \leq n < m , \tag{4.18}$$

as explained later, and then control the evolution of (A_n, B_n, C_n) explicitly for $m-1$ steps, up to a point (determining m) where C_m is is so small that little is lost by continuing with a more crude estimate, such as the one given below. Define

$$s_n = 1/2 , \quad \rho^{(n+1)} = \rho^* + \delta^{(n)} , \quad \delta^{(n)} = 2^{-n+m-1}\bigl(\rho^{(m)} - \rho^*\bigr) , \quad n \geq m , \tag{4.19}$$

In what follows, we always assume that $4/5 \leq \sigma \leq 1$.

Proposition 4.4. *Assume $\Gamma\bigl(\rho_1^{(m+1)}/\rho_1^{(m)}\bigr) = \gamma \leq 6/5$. Let $d_m = \min\bigl\{\delta_1^{(m)}\kappa\rho_2^*/\sigma, \delta_2^{(m)}\bigr\}$, and let f_m be a non-resonant function in $\mathcal{B}_{\rho^{(m)}}$, such that*

$$\|f_m\|_{\rho^{(m)}} \leq C_m \leq \frac{1}{24} d_m^2 . \tag{4.20}$$

Let H_m be a resonant Hamiltonian with the property that

$$\|H_m - H_0\|_{\rho^{(m)}} \leq B_m \leq \frac{1}{6} - 2d_m , \qquad \|\mathcal{D}(H_m - H_0)\|_{\rho^{(m)}} \leq A_m \leq \frac{2}{3} , \tag{4.21}$$

Then for every $n > m$, the functions $H_n - H_0$ and f_n belong to $\mathcal{B}_{\rho^{(n)}}$ and satisfy the bounds

$$\|f_n\|_{\rho^{(n)}} \leq C_n \leq 4^{-n+m} C_m , \qquad \|H_n - H_m\|_{\rho^{(n)}} \leq B_n - B_m \leq \frac{13}{3}\, C_m . \tag{4.22}$$

Proof. We proceed by induction, using Proposition 4.3 with the parameters (4.17). In fact, we will weaken the bounds (4.12) slightly, by substituting ρ_2^* for ρ_2' in the definition of the parameter d. Let now $i \geq m$ be fixed, and consider $m \leq n \leq i$. Assume that the first bound in (4.22) holds, together with the inequalities

$$\|H_n - H_0\|_{\rho^{(n)}} \leq B_n \leq \frac{1}{6}, \qquad \|\mathcal{D}(H_n - H_0)\|_{\rho^{(n)}} \leq A_n \leq \frac{3}{4}. \tag{4.23}$$

Define $d_n = 2^{-n+m} d_m$. From the assumptions (4.21) and (4.20), it follows e.g. that $d_m < 1/6$ and $C_m < d_m/120$. By using these two inequalities, and (4.23), we find that for both $x = d_n$ and $x = s_n d_n$, the constants $\tau_n(x)$ and $K_n(x)$, defined in (4.11), are bounded by 1/6 and 6, respectively. Thus, by using Proposition 4.3, together with the assumption (4.20), and the bound (4.22) on C_n, we have

$$\|f_{n+1}\|_{\rho^{(n+1)}} \leq C_{n+1} \leq K_n(d_n) d_n^{-2} C_n^2 \leq 6 d_m^{-2} C_m C_n \leq \frac{1}{4} C_n. \tag{4.24}$$

This shows that the first part of (4.22) extends to $n = i + 1$.

Let again $m \leq n \leq i$. By using that $A_n \leq 3/4$ and $C_{n+1} \leq C_n/4$, we obtain from Proposition 4.3 the bound

$$\|H_{n+1} - H_n\|_{\rho^{(n+1)}} \leq \Delta B_n = \frac{A_n}{1 - A_n} C_n + C_{n+1} \leq \frac{13}{4} C_n \leq \frac{13}{4} C_m 4^{-n+m}. \tag{4.25}$$

Summing over $n = m, \ldots, i$ shows that the second bound in (4.22) holds for $n = i + 1$. This, together with the assumption (4.21), and the fact that $C_m < d_m/120$, implies that the first inequality in (4.23) extends to $n = i + 1$.

By using the same facts that lead to (4.24), we obtain for $m \leq n \leq i$ the bound

$$\bigl\|\mathcal{D}(H_{n+1} - H_n)\bigr\|_{\rho^{(n+1)}} \leq \Delta A_n \leq \frac{(1 - A_n)^{-1} A_n C_n + C_n}{\sigma(1 - s_n) d_n} \gamma \leq 10 \frac{C_m}{d_m} 2^{-n+m}. \tag{4.26}$$

Summing over $n = m, \ldots, i$, and adding the result to the first inequality in (4.21), using that $C_m < d_m/120$, we find that the second bound in (4.23) holds for $n = i + 1$. This concludes the induction step, and thus the proof. **QED**

Consider now a fixed set of parameters (4.18), such that $\Gamma\bigl(\rho_1^{(m+1)}/\rho_1^{(m)}\bigr) \leq 6/5$. In addition, we fix a resonant Hamiltonian $H_1 = H_0 + h_1$, and positive constants A_1, B_1, satisfying

$$\|h_1\|_{\rho^{(1)}} < B_1 < \frac{1}{6} - 2\delta_2^{(m)}, \qquad \|\mathcal{D}(h_1)\|_{\rho^{(1)}} < A_1 < 2/3. \tag{4.27}$$

Given $C_1 > 0$, we define a sequence of triples (A_n, B_n, C_n), using equation (4.12), with $k = 0$ for $n = 1$, and $k = 1$ otherwise.

Corollary 4.5. *If $\|f_1\|_{\rho^{(1)}} \leq C_1$, and C_m satisfies the second inequality in (4.20), then the limit*

$$\mathcal{N}(H_1 + f_1) = \lim_{n \to \infty} H_n \tag{4.28}$$

exists in $\mathcal{B}_{\rho^*}$ and satisfies the bound

$$\|\mathcal{N}(H_1+f_1)-\mathcal{N}_1(H_1,f_1)\|_{\rho^*}\le C_2+\sum_{n=2}^{m-1}\Delta B_n+\frac{13}{3}C_m\,. \tag{4.29}$$

Proof. The inequality (4.29) follows from the identity

$$\mathcal{N}(H_1+f_1)-\mathcal{N}_1(H_1,f_1)=\mathbb{I}^{+}R(H_1,f_1)+\sum_{n=2}^{\infty}\bigl(H_{n+1}-H_n\bigr)\,, \tag{4.30}$$

using Proposition 4.4, and the bound (4.15). **QED**

At this point, we take equation (4.28) as our definition of the map $\mathcal{N}$. Some properties of $\mathcal{N}$ are summarized in the

Lemma 4.6. *Assume that $C_1>0$ is sufficiently small, such that C_m satisfies the second inequality in (4.20). Consider an open ball B in $\mathcal{B}_{\rho^{(1)}}$ of radius less than C_1, centered at H_1. Then $\mathcal{N}$ is well defined and analytic, as a map from B to $\mathbb{I}^{+}\mathcal{B}_{\rho^*}$. For every Hamiltonian H in B, the equation (4.1) defines a map $\mathcal{U}_H$ from D_{ρ^*} to $D_{\rho^{(1)}}$, and we have*

$$\mathcal{N}(H)=H\circ\mathcal{U}_H\,. \tag{4.31}$$

If H_1 and $H\in B$ belong to $\mathcal{B}_{\rho^{(1)},\varepsilon}$, for some $\varepsilon>0$, then there exists $\alpha>0$, such that $\mathcal{U}_H$ extends to an analytic canonical transformation from $D_{\rho^*,\alpha}$ to $D_{\rho^{(1)},\alpha}$.

Proof. The analyticity of $\mathcal{N}$ follows from the uniform convergence of our iteration schemes. The convergence of the infinite composition (4.1), uniformly on compact subsets of D_{ρ^*}, can be proved as in [25]. This yields a map $\mathcal{U}_H: D_{\rho^*}\to D_{\rho^{(1)}}$, which satisfies equation (4.31).

At this point, it is useful to note that we could have chosen any fixed $\alpha\ge 0$ in the definition of (the norm in) the spaces $\mathcal{F}_\rho$, without affecting any of the estimates in this section. The only change, when increasing α, is that the results apply to fewer Hamiltonians.

Assume now that H_1 belongs to $\mathcal{B}_{\rho^{(1)},\varepsilon}$, for some $\varepsilon>0$. Then we can find $\alpha>0$, such that the bounds (4.27) remain satisfied, if $\|.\|_{\rho^{(1)}}$ is replaced by $\|.\|_{\rho^{(1)},\alpha}$, and such that B is contained in an open ball in $\mathcal{B}_{\rho,\alpha}$ of radius less than C_1, centered at H_1. For a Hamiltonian H in this ball, we now obtain convergence of (4.1) on compact subsets of $D_{\rho^*,\alpha}$, which preserves analyticity. **QED**

5. The transformations $\mathcal{R}$ and $\mathcal{M}$

The transformation $\mathcal{M}$ will be estimated by using the decomposition (1.18). Controlling the affine part (first two terms) is relatively straightforward, albeit tedious, using the estimates from Section 3 and the formula for $D\mathcal{S}(H_1)$ given below. What remains to be discussed are the nonlinear (last two) terms in the decomposition (1.18) of $\mathcal{M}$. We start by considering the normalization transformations $\mathcal{S}$, defined by equation (1.7). In what follows, the approximate RG fixed point H_1 is always assumed to be a Fourier-Taylor polynomial. This is not really necessary, but it is sufficient and simplifies the discussion.

Given $H_1 = H_0 + h_1$, let $c = 2(h_1)_{0,2}$. If f is any function of q and p, define

$$\widehat{f} = \mathcal{S}_c f\,, \qquad (\mathcal{S}_\lambda f)(q,p) = \lambda f(q, p/\lambda)\,, \qquad \lambda \in \mathbb{C} \setminus \{0\}\,. \tag{5.1}$$

Then the derivative of $\mathcal{S}$ at H_1, and the nonlinear part (1.15) of $\mathcal{S}$, can be written as

$$\begin{aligned} D\mathcal{S}(H_1)f &= \widehat{f} + \epsilon\bigl(\widehat{H_1} - Z\partial_z \widehat{H_1}\bigr)\,, \\ \widetilde{\mathcal{S}}(H_1, f) &= \bigl[\mathcal{S}_{1+\epsilon}\widehat{H_1} - \widehat{H_1} - \epsilon\bigl(\widehat{H_1} - Z\partial_z\widehat{H_1}\bigr)\bigr] + \bigl(\mathcal{S}_{1+\epsilon}\widehat{f} - \widehat{f}\;\bigr)\,, \end{aligned} \tag{5.2}$$

where $\epsilon = 2\widehat{f}_{0,2}$ and $Z(q,p) = \Omega \cdot p$. Due to our assumption on H_1, the derivative $D\mathcal{S}(H_1)$ is bounded as a linear operator from $\mathcal{B}_\rho$ to $\mathcal{B}_{\hat\rho}$, for $\hat\rho = (\rho_1, c\rho_2)$. Furthermore, the term in square brackets in equation (5.2) can be controlled by an explicit computation. The following proposition can be used to estimate the remaining part of $\widetilde{\mathcal{S}}(H_1, f)$.

Proposition 5.1. *Assume $\rho'_1 = \hat\rho_1$ and $\rho'_2 < \hat\rho_2$. If $\widehat{f}$ belongs to $\mathcal{B}_{\hat\rho}$ and $|\epsilon|$ is less than $\epsilon_0 = 1 - \rho'_2/\hat\rho_2$, then*

$$\bigl\|\mathcal{S}_{1+\epsilon}\widehat{f} - \widehat{f}\;\bigr\|_{\rho'} \le |\epsilon|\left(1 + \frac{1}{\epsilon_0 - |\epsilon|}\right)\bigl\|\widehat{f}\;\bigr\|_{\hat\rho}\,. \tag{5.3}$$

Proof. If we define $\widehat{f}_\lambda(q,p) = \widehat{f}(q, p/\lambda)$, then

$$\bigl\|\widehat{f}_{1+\epsilon} - \widehat{f}\;\bigr\|_{\rho'} \le \left\| \frac{1}{2\pi i} \oint\limits_{|\lambda|=\lambda_0} \frac{\epsilon}{\lambda(\lambda-\epsilon)} \widehat{f}_\lambda \, d\lambda \right\|_{\rho'} \le \frac{|\epsilon|}{\epsilon_0 - |\epsilon|}\bigl\|\widehat{f}\;\bigr\|_{\hat\rho}\,, \tag{5.4}$$

and by adding to this $\bigl\|\epsilon\widehat{f}_{1+\epsilon}\bigr\|_{\rho'} \le |\epsilon|\bigl\|\widehat{f}\;\bigr\|_{\rho}$, the bound (5.3) follows. **QED**

What remains to be done is to put all the pieces together. By using that $\mathcal{S} = \mathcal{S} \circ \mathcal{S}_c$, we define $\mathcal{R}$ to be the following composition of maps:

$$\mathcal{R}:\ \mathcal{B}'_\rho \xrightarrow{\mathcal{S}_c} \mathcal{B}'_{\hat\rho} \xrightarrow{\mathcal{S}} \mathcal{B}'_{\rho'} \xrightarrow{\mathcal{L}} \mathcal{B}'_{\rho''} \xrightarrow{\mathcal{N}} \mathcal{B}'_{\rho^*}\,, \tag{5.5}$$

with some additional restriction on its domain. To be more precise about the domain, let us fix $\rho'' > \rho^* > \rho$. Assume that the function ψ_0 defining the canonical transformation U_{ϕ_0} satisfies the conditions given in Section 3, so that $\mathcal{L}$ is a well defined and continuous map between the spaces indicated in (5.5). Here, $\hat{\rho} = (\rho_1, c\rho_2)$ and $\rho' = (\rho_1, \vartheta^{-2}b_2)$, with b_2 the constant defined by equation (3.11). In order to benefit from Proposition 5.1, we restrict $\mathcal{R}$ to Hamiltonians $H = H_1 + f$ for which

$$2\bigl|\widehat{f}_{0,2}\bigr| < 1 - \vartheta^{-2}b_2/(c\rho_2)\,, \tag{5.6}$$

assuming that this inequality holds for $f = 0$. Additional choices, and conditions, are involved in the definition of $\mathcal{N}$, as described in Section 4. This includes the domain parameter $\rho^{(1)}$, which we take to be ρ'', and the bounds (4.27) on the Hamiltonian H_1.

For the purpose of this section, the domain $D_{\mathcal{R}}$ of $\mathcal{R}$ is now defined as the set of all Hamiltonians $H = H_1 + f$ in $\mathcal{B}'_\rho$ that satisfy the condition (5.6), and are mapped into the domain of $\mathcal{N}$ by $\mathcal{L} \circ \mathcal{S}$. This in turn defines the domain of the map $\mathcal{M}$, given by equation (1.12), as the inverse image of $D_{\mathcal{R}}$ under $h \mapsto H_1 + Mh$. Here, M is a fixed linear isomorphism of $\mathcal{B}_\rho$, chosen to be an approximate inverse of $\mathbb{I} - D\mathcal{R}_1(H_1)$.

Denote by $B(r)$ the open ball of radius r in $\mathcal{B}_\rho$, centered at the origin. Given $r > 0$ such that $B(2r)$ is contained in the domain of $\mathcal{M}$, we can use Proposition 5.1, together with Proposition 3.2 (case $\lambda = 0$) and Proposition 2.2 (case $\psi = 0$), to determine positive constants K_1 and K_2, such that

$$\bigl\|\widetilde{\mathcal{R}}(H_1, Mh)\bigr\|_{\rho^*} \le K_n\,, \qquad h \in B(nr)\,, \quad n = 1, 2\,, \tag{5.7}$$

where $\widetilde{\mathcal{R}}$ is the nonlinear part of $\mathcal{R}_1$, defined by equation (1.17). Similarly for the last term in the representation (1.18), which involves the difference

$$\mathcal{R} - \mathcal{R}_1 = \bigl(\mathcal{N} - \mathcal{N}_1\bigr) \circ \mathcal{L} \circ \mathcal{S}\,. \tag{5.8}$$

By using Proposition 3.2 (case $\lambda = 0$) and the bound (4.29) on $\mathcal{N} - \mathcal{N}_1$, we can determine constants K'_1 and K'_2, such that

$$\bigl\|\bigl(\mathcal{R} - \mathcal{R}_1\bigr)(H_1 + Mh)\bigr\|_{\rho^*} \le K'_n\,, \qquad h \in B(nr)\,, \quad n = 1, 2\,. \tag{5.9}$$

Theorem 5.2. *There exists an odd real Fourier-Taylor polynomial ψ_0, an even real resonant Fourier-Taylor polynomial h_1, an invertible bounded linear operator M on $\mathcal{B}_\rho$, a set of parameters as described after (5.5), and a positive real number r, such that $B(2r)$ is contained in the domain of $\mathcal{M}$. Let $H_1 = H_0 + h_1$ and*

$$\varepsilon = \|\mathcal{R}_1(H_1) - H_1\|_{\rho^*}\,, \qquad K = \bigl\|\mathbb{I} - \bigl(\mathbb{I} - D\mathcal{R}_1(H_1)\bigr)M\bigr\|_\rho\,. \tag{5.10}$$

Then there are positive constants K_1, K_2, K'_1, and K'_2, satisfying

$$\varepsilon + Kr + K_1 + K'_1 < r\,, \qquad K + \bigl(K_2 + K'_2\bigr)/r < 1\,, \tag{5.11}$$

such that the inequalities (5.7) and (5.9) hold. Furthermore, for every Hamiltonian H in $H_0 + MB(r)$, the normalization constant $c_H = 2H_{0,2}$ is larger than 1.

This theorem is proved with the aid of a computer, as described in Section 6.

As a corollary, we find that $\mathcal{M}$ is a contraction on the closure of the ball $B(r)$, and thus has a unique fixed point in $\mathcal{B}_\rho$ of norm $\le r$. The fact that $B(r)$ is mapped into itself by $\mathcal{M}$ follows from the first inequality in (5.11), using the representation (1.18) of $\mathcal{M}$. The contraction property follows from the second inequality in (5.11), and from Proposition 5.3 below, which implies that the derivative of $\mathcal{M}$ is bounded in norm by $K+\bigl(K_2+K_2'\bigr)/r<1$ on the ball $B(r)$, and thus also on the closure of this ball. Here, we have used that $\mathcal{M}$ is analytic on $B(2r)$. This is a consequence of the fact that $\mathcal{M}$ is composed of maps that are analytic on their domains; see Corollary 3.3 and Lemma 4.6.

Proposition 5.3. *Let X and Y be complex Banach spaces, and let F be an analytic map from $B=\{x\in X:\|x\|<2r\}$ to Y, with values in a ball $\{y\in Y:\|y\|\le k\}$. If $2x$ belongs to B, then the derivative of F at x satisfies the bound*

$$\|DF(x)\|\le k/r\,. \tag{5.12}$$

Proof. Assume that $2x$ belongs to B. Given any unit vector h in X, choose a continuous linear functional $\ell:Y\to\mathbb{C}$ of norm 1, such that $\ell(DF(x)h)=\|DF(x)h\|$. For every $z\in\mathbb{C}$ of modulus $<r$, define $f(z)=\ell(F(x+zh))$. Then $|f(z)|\le k$ whenever $|z|<r$, and by using Cauchy's formula, we obtain $\|DF(x)h\|=|f'(0)|\le k/r$. **QED**

Proof of Theorem 1.2. Let $B=H_1+MB(r)$, with $B(r)$ the ball described in Theorem 5.2 (which we assume to hold). By the definition of the map $\mathcal{M}$ and its domain, $\mathcal{R}$ is well defined on the (open) set B, and it is analytic as a map from B to $\mathcal{B}'_{\hat\rho}$, since it is composed of maps that are analytic on their domains; see Corollary 3.3 and Lemma 4.6. Furthermore, as equation (4.28) shows, $\mathcal{R}(H)$ is resonant, for every $H\in B$. By Proposition 3.4, it follows that the range of $\mathcal{R}$ is contained in $\mathcal{B}'_{\rho,\alpha}$ for some $\alpha>0$, and that $\mathcal{R}$ is compact.

As mentioned earlier, Theorem 5.2 implies the existence of a unique fixed point for $\mathcal{M}$ in the closure of $B(r)$. In fact, since the first inequality in (5.11) is strict, this fixed point h lies in $B(r)$. The corresponding Hamiltonian $H=H_1+Mh$ is clearly a fixed point of $\mathcal{R}$. It satisfies $c_H>1$ by Theorem 5.2, and given that M is invertible, $\mathcal{R}$ has no other fixed point in B. This concludes the proof of Theorem 1.2. **QED**

6. Outline of the proof of Theorem 5.2

The proof of Theorem 5.2 is based on a discretization of the problem, made possible by the compactness property of $\mathcal{R}$, and by the fact that only strict inequalities need to be proved. We will describe here an algorithm that can be (and has been) used to implement this discretization, and to give rigorous estimates on the discretization errors. This description only covers the main ideas and choices; for details, the reader is referred to the source code of our computer program [31].

In order to simplify notation, we will from now on ignore the component $\omega\cdot p$ of the action variable p. This does not represent a restriction, since a function f in $\mathcal{F}_\rho$ only depend on the component $z=\Omega\cdot p$. Similarly, if f does not depend on the variable q or z, we will write $f(z)$ or $f(q)$, respectively, in place of $f(q,z)$.

6.1. Standard sets

Estimates on a real number usually take the form of upper and lower bounds, i.e., intervals. In order to mechanize the process, we will limit such estimates to intervals $[c-r, c+r]$ whose center c and radius $r \geq 0$ belong to a finite set of real numbers that are "representable" on a computer, in a standard IEEE floating point format (described later). The collection of all these intervals, also referred to as "standard sets" for $\mathbb{R}$, will be denoted by $\operatorname{std}(\mathbb{R})$.

Objects in other spaces will be estimated similarly, by locating them in standard sets for those spaces. We start by defining the standard sets for $\mathcal{B}_\rho$. This involves the the choice of four positive integers $d' \leq d$ and $\ell < N$. For every non-negative integer $n \leq N$, let

$$V_n = \bigl\{\nu \in \mathcal{V} : |\omega \cdot \nu| < n\,,\ |\Omega \cdot \nu| < n\bigr\}\,, \qquad V_n^c = \mathcal{V} \setminus V_n\,, \tag{6.1}$$

where $\mathcal{V}$ is the index set described after equation (1.2). Let $\mathcal{I}$ be the collection of all balls

$$[c-r; c+r] = \bigl\{f \in \mathcal{B}_\rho : \|f - c\|_\rho \leq r\bigr\}\,, \tag{6.2}$$

with representable center c and radius r. If C is such a ball, then the center of C will be denoted by $\overline{C}$. In addition, if J a non-empty subset of I, we denote by $C(J)$ the intersection of C with the subspace $\mathbb{J}\mathcal{B}_\rho$. Given any non-empty set $V \subseteq \mathcal{V}$, and any element $\mathcal{C} = (E, C_0, \ldots, C_d)$ in $\mathcal{I}^{d+2}$, such that $\overline{E} = 0$, we define

$$\mathcal{P}(\mathcal{C}, V) = \sum_{k=0}^{d'-1} C_k(V \times \{0\})Z^k + E(V \times \mathbb{N}) + \sum_{k=d'}^{d} C_k(V \times \mathbb{N})Z^k\,, \tag{6.3}$$

where $Z(q,p) = \Omega \cdot p$. This defines a closed convex star-shaped set in $\mathcal{B}_\rho$, centered at the polynomial $P = \sum_{k=0}^{d} \overline{C_k} Z^k$. The first component E of $\mathcal{C}$ will be referred to as its "error interval".

Consider now functions from $V_\ell \cup \{0, 1, \ldots, N\}$ to $\mathcal{I}^{d+2}$, with the property that all intervals in $\mathcal{C}_1, \ldots, \mathcal{C}_N$, and all error intervals, are centered at zero. To every such function $\mathcal{C}$, we associate a set

$$\mathcal{H}(\mathcal{C}) = \sum_{\nu \in V_\ell} \mathcal{P}(\mathcal{C}_\nu, \{0\}) \cos(\nu \cdot Q) + \sum_{n=0}^{N} \mathcal{P}(\mathcal{C}_n, V_n^c)\,, \tag{6.4}$$

where $Q(q,p) = q$, and we define $\operatorname{std}(\mathcal{B}_\rho)$ to be the collection of these sets (for now. Some simple "polynomial sets" will be included later). The first sum in this equation represents a finite Fourier series with coefficients that are functions of z. The terms in the second sum will be referred to as "higher order errors in q", since the functions in $\mathcal{P}(\mathcal{C}_n, V_n^c)$ have only frequencies outside the rectangle V_n. However, it should be noted that for $n = 0$, there is no restriction on the frequency.

Standard sets for the spaces $\mathcal{A}_\rho$ are defined analogously. Unless mentioned otherwise, the standard sets for a product space $X_1 \times X_2$ are taken to be the products $C_1 \times C_2$ with $C_1 \in \operatorname{std}(X_1)$ and $C_2 \in \operatorname{std}(X_2)$.

6.2. Basic bounds

Our next goal is to bound maps $x \mapsto y$ in a way that yields estimates on y from estimates on x. Given any set X, denote by 2^X the collection of all subsets of X. If f is a map from $D_f \subseteq X$ to a set Y, and F is a map from $D_F \subseteq 2^{D_f}$ to 2^Y, we say that F is a "bound" on f, if $f(x) \in F(C)$ whenever $x \in C \in D_F$.

If F_1 and F_2 are bounds on two maps f_1 and f_2, then $(C_1, C_2) \mapsto (F(C_1), F(C_2))$ is a bound on $(x_1, x_2) \mapsto (f_1(x_1), f_2(x_2))$. Furthermore, the composition $F_1 \circ F_2$, if it exists, is a bound on $f_1 \circ f_2$. These two properties of bounds, together with our choice of standards sets (for the domains and ranges of our bounds), allow us to estimate complicated maps such as $\mathcal{R}$, in terms of very simple ones.

As an example of a very simple map, consider the sum of two real numbers, $f(x,y) = x+y$. Our bound on f uses a tool that is available on most modern computer hardware: A function S that assigns to any pair (c,c') of IEEE numbers, the smallest IEEE number larger than or equal to $c+c'$, if such a number exists; otherwise, (c,c') is by definition outside the domain of S, and the program is aborted. Using this function S, we obtain a bound F on f by setting $F([c_1-r_1, c_1+r_1], [c_2-r_2, c_2+r_2]) = [c_3-r_3, c_3+r_3]$, where $c_3 = S(c_1, c_2)$ and $r_3 = S(r_1, S(r_2, S(c_3, S(-c_1, -c_2))))$, provided that all arguments of S are in the domain of S. Here, we have used also that $x \mapsto -x$ is computed precisely within the set of IEEE numbers. Notice that the "same" map F is also a bound on the sum of two functions in $\mathcal{B}_\rho$, if we replace the intervals $[c_i - r_i, c_i + r_i]$ with the corresponding balls $[c_i - r_i; c_i + r_i]$. Below we will extend the domain of this bound to other standard sets. Bounds on the difference, products, and quotient functions for real numbers are defined similarly. To be more precise, we make sure that the "same" bound also applies to balls in $\mathcal{I}$, by only using the general Banach algebra properties of $\mathbb{R}$, but not its order properties. All the other bounds used in our proof are defined in terms of these four bounds, and a few equally simple bounds, on functions like $x \mapsto |x|$ or $(x,y) \mapsto \max(x,y)$, etc.

For some non-rational functions on $\mathbb{R}$ we use either Taylor polynomials with bounds on the remainder (in the case of exp, cosh, sinh), or start with a numerical guess and then estimate the error (for the function Γ defined in (2.10), and for ln). All of our bounds in these cases are trivial, since they are used only to estimate norm factors, and thus need not be optimal.

The abovementioned bound F can be extended to standard sets (6.3) by setting $F(\mathcal{P}(\mathcal{C}, V), \mathcal{P}(\mathcal{C}', V)) = \mathcal{P}(F^{\otimes d+2}(\mathcal{C}, \mathcal{C}'), V)$, whenever $F^{\otimes d+2}(\mathcal{C}, \mathcal{C}')$ is defined. Clearly, this extended map F is still a bound on the sum of two functions in $\mathcal{B}_\rho$. A further extension, to general standard sets (6.4), is analogous. Bounds on the other vector space operations on $\mathcal{B}_\rho$ and $\mathcal{A}_\rho$ are defined similarly.

For bounds involving the product of functions, it is useful to think of $\mathcal{H}(\mathcal{C})$ in terms of a "general function"

$$h(q,z) = \sum_{\nu \in V_\ell} P_\nu(z) \cos(\nu \cdot Q) + \sum_{n=0}^{N} P_n(q,z)\,. \tag{6.5}$$

Here, P_ν represents the set $\mathcal{P}(\mathcal{C}_\nu, \{0\})$ and can be written in the form

$$P_\nu(z) = \sum_{k=0}^{d'-1} [c_k + \epsilon_k] z^k + e(z) + \sum_{k=d'}^{d} [c_k + e_k(z)] z^k \,, \tag{6.6}$$

with $c_k + \epsilon_k$, e, and $c_k + e_k$ belonging to the corresponding components of $\mathcal{C}_\nu$. Similarly, P_n represents the set $\mathcal{P}(\mathcal{C}_n, V_n^c)$ and can be written as

$$P_n(q,z) = \sum_{k=0}^{d'-1} e_k(q) z^k + e(q,z) + \sum_{k=d'}^{d} e_k(q,z) z^k \,, \tag{6.7}$$

in such a way that the functions e_k and e lie in the corresponding components of $\mathcal{C}_n$, and their frequencies ν are restricted to V_n^c. The product of two functions of the form (6.5) can be written again in the same form (canonically, if we try to preserve as much information as possible), and by determining balls that contain the resulting "coefficients" $c_k + \epsilon_k$, e, and $c_k + e_k$, we obtain a bound on the product of two functions in $\mathcal{B}_\rho$ and/or $\mathcal{A}_\rho$.

6.3. Compositions and inverses

The canonical changes of variables involved in the definition of $\mathcal{R}$ decompose into a step of the form $(h \circ_y f)(q,z) = h(q + f(q,z)\Omega, z)$, followed by a step of the form $(h \circ_z G)(q,z) = h(q, G(q,z))$. From the representation (6.6) and (6.7), we see that a bound on "$\circ_z$" is obtained essentially from the bounds on $G \mapsto G^k$ for $k \le d$. For the error terms, we use the trivial estimate $\|e_k \circ_z G\|_\rho \le \|e_k\|_{\rho'}$ from Proposition 2.3.b, after checking the necessary domain conditions.

A bound on "$\circ_y$" is obtained similarly, but instead of estimating powers of G, we need to estimate the functions $\cos(f\Omega \cdot \nu)$ and $\sin(f\Omega \cdot \nu)$, for frequencies $\nu \in V_\ell$. If the norm of $f\Omega \cdot \nu$ is sufficiently small, this is done by using a finite Taylor expansion with bound on the remainder; otherwise we first divide f by a sufficiently large factor 2^n, and at the end, use the double angle formula n times. The error terms $e_k \circ_y f$ are again estimated by using Proposition 2.3.b.

The two composition operators "$\circ_y$" and "$\circ_z$" get combined in the map L, defined by equation (3.10), which is one of the main parts of the RG transformation $\mathcal{R}$. Our bound on L is essentially a composition of the two bounds discussed above, with some changes, due to the fact that L also includes the contraction step $h \mapsto h \circ \mathcal{T}_{\mu_0}$. For the error terms P_n in the Hamiltonian (6.5), the contraction factor (for large n) is obtained from Proposition 3.2, with $\lambda = n$.

Next, we consider the map $f \mapsto 1/f$, which is used in connection with equation (3.1), and in particular, in the estimate (3.7). Let f and g_0 (an approximate inverse of f) be functions in $\mathcal{B}_\rho$, and choose

$$a \ge \|1 - f g_0\|_\rho \,, \quad b \ge \|f\|_\rho \qquad c \ge \|g_0(1 - f g_0)\|_\rho \,. \tag{6.8}$$

Proposition 6.1. *Let $\delta \geq 2a+2bc$ and $r \geq c/(1-\delta)$. Assume that $\delta \leq 1/2$ and $\delta+2bc < 1$. Then f does not vanish on D_ρ, and $1/f$ lies in $\mathcal{B}_\rho$, within a distance r of g_0.*

Proof. The function $g = 1/f$, if it exists, satisfies the equation $g = g(2-fg)$. Equivalently, $h = g - g_0$ is a fixed point of the map $\mathcal{K}$, defined by

$$\mathcal{K}(h) = h_0 + K(h)h\,, \qquad h_0 = g_0(1 - fg_0)\,, \quad K(h) = 2(1 - fg_0) - fh\,. \tag{6.9}$$

Let $k = \delta + 2bc$. By using Proposition 2.3.a, and the assumption $\delta \leq 1/2$, one easily finds that

$$\|K(h)\|_\rho \leq \delta\,, \qquad \|\mathcal{K}(h') - \mathcal{K}(h)\|_\rho \leq k\|h' - h\|_\rho\,, \tag{6.10}$$

for every $h, h' \in \mathcal{B}_\rho$ of norm r or less. Since δ and k are less than 1, these two inequalities imply that the sequence of functions

$$h_{n+1} = \mathcal{K}^{n+1}(h_0) = h_0 + K(h_n)h_0 + K(h_n)K(h_{n-1})h_0 + \ldots + K(h_n)\cdots K(h_0)h_0 \tag{6.11}$$

converges in $\mathcal{B}_\rho$ to a limit h, which is a fixed point of $\mathcal{K}$, and bounded in norm by r. The corresponding function $g = g_0 + h$ satisfies the equation $fg = fg(2 - fg)$, implying e.g. that the range of fg is a subset of $\{0, 1\}$. Since $fg \in \mathcal{B}_\rho$ is continuous, this means that $fg = 0$ or $fg = 1$. But $fg = 0$ violates the inequality

$$1 = \|1\|_\rho \leq \|1 - fg\|_\rho + \|f(g_0 - g)\|_\rho + \|fg\|_\rho \leq a + br + \|fg\|\,, \tag{6.12}$$

and thus $fg = 1$, which proves the claim. **QED**

Let now $\mathcal{H}$ be a fixed standard set in $\mathcal{B}_\rho$, centered at some Fourier-Taylor polynomial f_0. In order to determine a new standard set $\mathcal{H}'$ containing the (multiplicative) inverse of every function in $\mathcal{H}$, we first compute a representable approximate inverse g_0 for f_0. Then, by using a bound on the norm function on $\mathcal{B}_\rho$, we determine constants a, b, c that satisfy (6.8), for all $f \in \mathcal{H}$. After verifying that the assumptions of Proposition 6.1 hold, with δ and r representable, we now define $\mathcal{H}'$ to be the ball of radius r in $\mathcal{B}_\rho$, centered at g_0. The map $\mathcal{H} \mapsto \mathcal{H}'$, defined in this way, is clearly a bound on $f \mapsto 1/f$.

The same strategy is used to obtain a bound on the inverse under composition, described by equation (3.1). Here, the radius of the ball added to the approximate solution g_0 is obtained from Proposition 3.1.

6.4. The maps $\mathcal{N}_1$ and $\mathcal{N}$

Unlike the maps discussed so far, both $\mathcal{N}_1$ and $\mathcal{N}$ involve derivatives, and the projection operators $\mathbb{I}^{\pm}$.

The derivatives ∂_y and ∂_z are only needed for the Fourier-Taylor polynomials ψ_0 and h_1 , or rescaled versions of h_1 , or functions in a small ball around these polynomials. Thus, it suffices to restrict bounds to sets of polynomials, and to balls centered at the origin. (We have implemented more general bounds, but since they are not needed, we will not discuss them here.) Both are trivial to implement, using component-wise differentiation for polynomials, and and parts (e) and (f) of Proposition 2.3 for balls.

However, our standard sets (6.4) are such that in general, if $\mathcal{H}(\mathcal{C})$ contains a single polynomial, then its image under a bound on ∂_z (and other maps) also contains non-polynomial functions. The reason is that for $k \geq d'$, any uncertainty in the coefficient c_k in (6.6) (generated e.g. by rounding) is considered a bound on the norm of an unknown function $e_k(z)$. In order to avoid such loss of information, we add to $\mathrm{std}(\mathcal{B}_\rho)$ all sets of the form

$$\mathcal{H}^0(C) = \sum_{\nu \in V_\ell} \sum_{k=0}^{d} C_{\nu,k}\big(\{(0,0)\}\big) \cos(\nu \cdot Q) Z^k \,, \tag{6.13}$$

where C is a family of balls in $\mathcal{I}$, indexed by $V_\ell \times \{0, 1, \ldots d\}$. Notice that the "coefficients" $C_{\nu,k}(\{(0,0)\})$ in equation (6.13) are sets of constant functions, and thus can be identified with balls in $\mathbb{C}$. Analogous polynomial sets are added to $\mathrm{std}(\mathcal{A}_\rho)$.

Implementing bounds on the projections $\mathbb{I}^{\pm}$ is straightforward: For $\mathbb{I}^{+}$, the image of the set $\mathcal{H}(\mathcal{C})$ is obtained by setting $c_k = 0$ and $\epsilon_k = 0$ in the representation (6.7), for all $k \leq d_\nu^-$, where

$$d_\nu^- = \begin{cases} -1, & \text{if } |\omega \cdot \nu| \leq \sigma |\Omega \cdot \nu|; \\ \lfloor |\omega \cdot \nu| / \kappa \rfloor, & \text{otherwise.} \end{cases} \tag{6.14}$$

Here, $x \mapsto \lfloor x \rfloor$ denotes the floor function on $\mathbb{R}$. The radii of the interval components of $\mathcal{C}$ remain the same. Similarly for the projection $\mathbb{I}^{-}$, except that all terms (not just the centers) with $k > d_\nu^-$ in (6.7) are set to zero.

Two other operators that enter the definition of $\mathcal{N}_1$ and $\mathcal{N}$ are $\mathcal{D}_y$ and $\mathcal{D}_z$. If applied to a Fourier-Taylor polynomial (ψ_0), they can be bounded component-wise, which is trivial. In all other cases, the argument of $\mathcal{D}_y$ or $\mathcal{D}_z$ is a non-resonant function, and we can use Proposition 2.1 to estimate the non-polynomial parts in the representations (6.6) and (6.7).

For a bound on $(h_1, f_1) \mapsto \mathcal{N}_1(H_0 + h_1, f_1) - H_0 - h_1$, which involves solving the linear equation (2.5), we apply the same strategy as for other inverses: Starting with the centers of the given standard sets (for h_1 a set of polynomials), we first compute a numerical approximation ψ for the function ψ_1 defined in (2.5), then use the bounds discussed above to estimate $\mathbb{I}^{+}\mathcal{D}(h_1)\psi$, and then add to the result a ball determined by Proposition 2.2.

Finally, consider the map $(h_1, f_1) \mapsto \mathcal{N}(H_0 + h_1 + f_1) - \mathcal{N}_1(H_0 + h_1, f_1)$. Our bound on this map is restricted to polynomial standard sets (6.13) for h_1 , and to balls centered at zero for f_1 . As a first step, we determine (strict) upper bounds C_1, B_1, and A_1 (using Proposition 2.3.d), on the norms $\|f_1\|_{\rho''}$, $\|h_1\|_{\rho''}$, and $\|\mathcal{D}(h_1)\|_{\rho''}$, respectively. After that, it suffices to estimate the triples (A_n, B_n, C_n), defined by equation (4.12), for

$n = 2, 3, \ldots, m$, and to verify that they satisfy the conditions of Proposition 4.3 and Proposition 4.4. Then the desired bound is given by equation (4.29). In our implementation of this bound, we use $m = 5$ and try out 6 different choices (of relative distances) for each of the domain parameters $\rho^{(2)}, \ldots, \rho^{(5)}$, in order to optimize the final estimate. For the interpolation parameter s in (4.12), we use

$$s = \left[2K_n(0)(1 - A_n)A_n^{-k}d^{-2}C_n\right]^{1/3}, \tag{6.15}$$

which is roughly the value that minimizes ΔA_n.

6.5. The contraction $\mathcal{M}$

The only basic ingredients of $\mathcal{M}$ that we have not yet discussed are the normalization map $\mathcal{S}$, and the linear operator M. The map $\mathcal{S}$, its derivative, and the second order Taylor remainder $\widetilde{S}$, are straightforward to bound, using equation (5.2) and Proposition 5.1. This leaves us with the operator M, which appears in three of the terms in the decomposition (1.18) of $\mathcal{M}$.

Our approximate inverse M of $\mathbb{I} - D\mathcal{R}(H_1)$ is of the form $\mathbb{I} - W$, for an operator W whose matrix elements $W_{j,i}$ (with respect to the canonical basis for $\mathcal{B}_\rho$) are non-zero only if j and i belong to both I^+ and $V_\ell \times \{0, 1, \ldots, d-1\}$. Although W is just a finite matrix, there is no good general bound on such an operator, within the collection of standard sets (6.4). This is the reason why we apply M before any other step (in the definition of $\mathcal{M}$), when we still have maximal control of its input. We also use the maps $\mathcal{S}$ and $\mathcal{T}_{\mu_0}$ as early as possible, for the following reasons. Any error in the $(0, 2)$ coefficient of a function propagates to almost all other coefficients under the action of $\mathcal{S}$. And the composition with $\mathcal{T}_{\mu_0}$ contracts resonant terms $\cos(\nu \cdot Q)Z^k$ with large ν or k, so that if we loose track of some of these terms, we also loose contraction.

Ideally, a bound on a contraction like $\mathcal{M}$ needs to be defined only on balls. In our case, this can be done for the last two terms in the decomposition (1.18), since they involve no significant cancellations. In fact, a partial bound on M suffices for these terms, in the form of a bound on the map $h \mapsto (\|Mh\|_\rho, (Mh)_{0,2})$. This requires only an estimate of the partial norms

$$\|\mathbb{J}M\|_\rho = \sup_{i \in I} \sum_{j \in J} |M_{j,i}| w_j / w_i, \qquad w_{(\nu,k)} = \cosh(\rho_1 \Omega \cdot \nu)\rho_2^k, \tag{6.16}$$

for $J = I$, and for $J = \{(0, 2)\}$. For simplicity, we integrate these estimates directly into a bound on the combined map $h \mapsto \mathcal{S}(H_0 + Mh) - H_0$ and its derivative.

Bounds on the maps represented by the last two terms in (1.18), are now obtained by composing bounds that have already been discussed, using the representations (1.17) and (5.8). Similarly for the first term in equation (1.18).

What we still need is an estimate on the norm of

$$M' = \mathbb{I} - \big(\mathbb{I} - D\mathcal{R}_1(H_1)\big)M = W + D\mathcal{R}_1(H_1)M, \tag{6.17}$$

as a linear operator on $\mathcal{B}_\rho$. First, we note that if this norm is less than 1, then M is invertible, as claimed in Theorem 5.2. This can be seen as follows. Let J be a finite subset of I such that $\mathbb{J}W\mathbb{J} = W$. If the norm of M' is less than 1, then the kernel of M has to be trivial. This implies that $\mathbb{J}M\mathbb{J}$ is invertible on $\mathbb{J}\mathcal{B}_\rho$. Thus, M is invertible on $\mathcal{B}_\rho$.

The operator norm of M' is given by the equation

$$\|M'\|_\rho = \sup_{i\in I^+} \|M'G_i\|_\rho\,, \qquad G_{(\nu,k)}(q,z) = w_{(\nu,k)}^{-1} \cos(\nu\cdot q) z^k\,. \tag{6.18}$$

In order to estimate this norm, we first choose a finite number of standard sets $\mathcal{H}_n$ whose union contains all of the basis vectors G_i for $\mathcal{B}_\rho$. Then we apply a bound on $h \mapsto \|M'h\|_\rho$ to each of them, and take the maximum. Among the standard sets used are all singletons $\{G_i\}$, with i belonging to both I^+ and $V_\ell \times \{0,1,\ldots,d-1\}$. This part of our bound relies heavily on cancellations, which are due to the fact that M is an approximate inverse of $\mathbb{I} - D\mathcal{R}_1(H_1)$. In addition, we consider the standard sets $C(V_\ell^c \times \{k\})$, $C(V_\ell^c \times \{d', d'+1,\ldots\})$, and $C(\{\nu\}\times\{d,d+1,\ldots\})$, for all $k \in \{0,1,\ldots d'-1\}$ and $\nu \in V_\ell$, where $C = [-1,1]$. When restricted to these sets, M coincides with $D\mathcal{R}_1(H_1)$. Here, the contraction does not come from cancellations, but from the composition with $\mathcal{T}_{\mu_0}$.

6.6. Organization of programs: Packages and data types

The bounds described in the previous four subsections have been implemented in the programming language Ada95 [33]. Most of our code is contained in a collection Ada95 "packages", which encapsulate groups of related procedures and data, and make their functionality available through a well defined interface. The main packages and data types are organized in layers, with each layer providing the main building blocks for the next.

At the lowest level we have a package named `Reps`, defining the set of representable real numbers, in the form of a data type `Rep`. Depending on `Reps` are two children packages, `Reps.Ops` and `Reps.IO`, which provide some elementary functions (not used for bounds) and input/output procedures, respectively, for the type `Rep`.

The next level contains a package named `Intervals`, which defines the data type `Interval`, and a few simple operations that do not involve rounding. This data type represents real intervals, as well as "generalized intervals" (6.2). An `Interval` has two components of type `Rep`, specifying a center and a radius. However, this information is "private" to the package `Intervals` and its children `Intervals.Ops` and `Intervals.IO`. These three packages provide the basic bounds (and input/output) on the level of intervals and balls, and specifications on how to use them, but no information about how data is represented. The same strategy is followed at higher levels. We note that our bound on the sum of two (generalized) intervals is implemented as a binary operator "`+`", with left and right argument of type `Interval`. Similarly for other elementary arithmetic operations.

For purely numerical computations (recall that several of our bounds require approximate solutions), we use three packages named `Numerics`, `Numerics.Ops` and `Numerics.IO`. Their public part is the same as that of `Intervals` and its children, except that the private type `Interval` is replaced by a public type `Numeric`, derived from `Rep`. This allows e.g. other packages to work with a "generic type" `Scalar`, which can be instantiated later to either Interval (for rigorous bounds) or Numeric (for numerical computations). In fact, all of our higher level packages are generic in this sense.

The third level packages Taylors, Taylors.Ops and Taylors.IO, implement bounds involving the standard sets $\mathcal{P}(\mathcal{C}) = \mathcal{P}(\mathcal{C}, \mathcal{V})$, represented by a private type Taylor. They take as argument (package parameter) a data type named Scalar, to be used for the components P.C(K) and P.E of a record P of type Taylor, which represent the coefficients C_k and E of the standard set $\mathcal{P}(\mathcal{C})$. Unless specified otherwise, we will assume now that Scalar is chosen to be Interval. For efficiency, some of our bounds return standard sets that have $C_k = \{0\}$ for all k larger than a specified cutoff D, and this "cutoff" is stored in a designated component P.D of the result P. We also allow the special value P.D=-1, but only if P is identically zero. It should be noted that the type Taylor does not carry any information about domains. This information has to be provided by the user, as arguments to the bounds (procedures or functions) that depend on domains.

We also note that standard set $\mathcal{P}(\mathcal{C}, V)$ can be associated with pairs $(\mathcal{P}(\mathcal{C}), V)$. Similarly, the balls $C(J)$ can be associated with pairs (C, J). At the levels discussed so far, we only keep track of, and estimate, the first component of these pairs.

Level four consists of the packages Fouriers, Fouriers.Ops, and Fouriers.IO. The package parameters include the length Lmax and width Wmax of a rectangle in frequency space. For the purpose of the work discussed here, this rectangle is always the square V_ℓ. The remaining parameters correspond to those used in Taylors. The bounds at this level are still mostly general purpose, involving the standard sets for $\mathcal{B}_\rho$ and $\mathcal{A}_\rho$, represented by a (private) data type Fourier. A variable F of type Fourier has several components, including two arrays F.C and F.E, whose elements represent the coefficients $\mathcal{P}(\mathcal{C}_\nu, \{0\})$ and $\mathcal{P}(\mathcal{C}_n, V_n^c)$, respectively, in the equation (6.4), as well as a parity component F.J, whose value (0 or 1) indicates whether F is even or odd. The domain parameter ρ is stored in a component F.R. Finally, a Boolean component F.P specifies whether or not F represents a set (6.13) of Fourier-Taylor polynomials, as opposed to a general standard set (6.4). Since this representation is "private", the packages at this level also provide facilities for accessing this information. A few types and functions, that are independent of the choice of Scalar, are defined in separate package called Fouriers_Init.

We note that Fouriers.Ops includes two procedures (versions of Comp3 and CompOmegaTmu0) that process simultaneously an entire array of standard sets. These procedures implement bounds on the composition operator "$\circ_z$", and on the combination of "$\circ_y$" with the map $h \mapsto h \circ \mathcal{T}_{\mu_0}$. The bound on "$\circ_y$" is extremely computation intensive, but most of the estimates are independent of the input data. Thus, we carry out several bounds in parallel, whenever possible. This is used e.g. for the standard sets $\mathcal{H}_n$ mentioned after equation (6.18).

In addition, some purely numeric procedures (generating error messages, if used with type Interval), that are more renormalization group specific, have been included in Fouriers.Ops, for reasons of efficiency. In order to facilitate the use of numeric procedures inside an otherwise generic package, we added a generic child Numerics.Ops.Fou to the package Numerics.Ops, which uses a Numeric instantiation of the packages Fouriers and Fouriers.Ops, and makes some of their entities available under different names. In particular, the Numeric version of the type Fourier is renamed to Polynom.

The next level consists of a single generic package called `RG_Ops`. It takes the same package parameters as `Fouriers`, and two additional parameters `Sigma` and `Kappa`, which specify the values of σ and κ defining the projections $\mathbb{I}^{\pm}$. For "historical" reasons, the generic version of `Fourier` is renamed to `Hamilton`. The bounds given in `RG_Ops` are mostly the ones described in Subsections 6.4 and 6.5. But we also include bounds on inverses, described in Subsection 6.3, which are based on approximate solutions. These approximate solutions are computed by using the `Polynom` procedures from `Numerics.Ops.Fou`.

There are four auxiliary packages that have not yet been mentioned: `Lin`, `Lin.Ops`, `Lin.IO`, and `Messages`. The first three of them define a data type `Sparse`, which is used to represent the matrix W, and they implement some trivial procedures for handling such matrices. The package `Messages` is used by various program parts to display messages or prompt for input.

6.7. Main programs

In order to prove Theorem 5.2, we need to estimate the constants $K_1, K_2, K_1', K_2', \varepsilon, K$, and then verify that the inequalities (5.11) hold. We have split this task into several main programs (procedures), some of which can be run independently from others, e.g. on different machines, in order to reduce computation time. Since all the necessary bounds have been implemented in the abovementioned packages, the main programs do little else than defining parameters, reading data, calling procedures from `RG_Ops`, and displaying results or writing intermediate results to disk (in hexadecimal ASCII format).

The input to these programs are the two Fourier-Taylor polynomials h_1 and ψ_0 , and the matrix W. The function ψ_0 defines the canonical transformation U_{ϕ_0} , via the solution g of equation (3.1). Since the function g is needed repeatedly, our first step is to determine a standard set containing g, and to save it for all other programs to use. This is done in the program `Run_UpsiAux`. We also save standard sets containing $\partial_y\psi_0, \partial_z\psi_0$, and $\mathcal{D}_z\psi_0$, for reasons that are explained below.

The next four programs can be run independently of each other. `Run_RG1` estimates the norm $\|\mathcal{R}_1(H_1) - H_1\|_{\rho^*}$. Two other programs, `Run_TildeR` and `Run_RminusR1`, are used to obtain the upper bounds K_n and K_n' , respectively, in the inequalities (5.7) and (5.9). These bounds are needed for $n = 1$ and $n = 2$. `Run_TildeR` prompts for the value of n, and has to be run separately (possibly in parallel) for each value, while `Run_RminusR1` automatically covers both cases. All of the programs mentioned so far, as well as the program `RGAux` described below, use the "degrees" $\ell = 40$ and $d = 16$.

Much of our strategy of proof has been motivated by the fact that bounding the derivative of our RG transformation in the "traditional" way (for computer-assisted proofs) would be computationally prohibitive with current technology. This includes the use of the approximate RG transformation $\mathcal{R}_1$ (which is much easier to differentiate than $\mathcal{R}$), and measures for achieving the desired accuracy with the relatively low degrees $\ell = 40$ and $d = 16$. (This influenced many choices, including the order of the maps in the definition of $\mathcal{R}$). Still, it was necessary to lower the degrees even further, in the programs that estimate the operator norm of M', despite the parallelization of our bounds for compositions. Thus, we need to convert (enlarge) the standard sets (6.13) for ψ_0 and other polynomials, to standards sets (6.4) with a lower value of ℓ and d, before they can be used by these programs. However, a set of polynomials is no longer recognized as such after the conversion, and even if this information were preserved, it would not help to get accurate bounds on derivatives or coefficients. In order to minimize such effects, we carry out some of these bounds beforehand, at $\ell = 40$ and $d = 16$, and convert the results afterwards. To be more precise, we save the results (as mentioned in the description of `Run_UpsiAux`) and convert them later with a program called `LowerQDegs`.

Such an estimate is carried out by running the program `RGAux`, which determines (and then saves) a standard set containing the image of the function $f(q,z) = z^2$, under the derivative of $R = \mathcal{L} \circ \mathcal{S}$ at H_1. The result is used later to estimate functions $DR(H_1)MG_i$, as part of our bound on the norm (6.18) of the operator M'. More specifically, we decompose MG_i into a sum $f_i + c_i f$, such that the $(0,2)$ coefficient of f_i is zero. Given a standard set containing $DR(H_1)f$, it now suffices to estimate $DR(H_1)f_i$. This can be done accurately at low degrees, since $D\mathcal{L}(H_1)$ acts trivially on f_i.

Our estimate on the norm of M' is organized in the two programs `MM1Norm_W` and `MM1Norm_noW`. Among the standard sets $\mathcal{H}_n$ described after equation (6.18), `MM1Norm_W` considers those with frequencies ν outside V_ℓ, as well as all singletons $\{G_i\}$ for which at least one matrix element $W_{j,i}$ is non-zero. This makes a total of 759 standard sets. Whenever possible, `MM1Norm_W` processes 20 of them in parallel. In addition, `MM1Norm_W` prompts for, and can be restricted to, a subrange of standard sets to be processed, so that the task can be distributed among several machines. The program `MM1Norm_noW` covers those standard sets $\mathcal{H}_n$ that are not considered by `MM1Norm_W`.

The parameters used in the proof of Theorem 5.2 are $\sigma \approx 0.85001$, $\kappa \approx \sigma/0.4$, $\rho \approx (0.85, 0.15)$, $\rho^* = (\mathtt{Up}(\rho_1), \mathtt{Up}(\rho_2))$, $\rho'' \approx (0.9, 0.165)$, and $r \approx 3 * 10^{-12}$, where $\mathtt{Up}(\rho_i)$ is the successor of ρ_i in the set of representable numbers. The precise value of r, in standard base 16 notation, is 3.4C6CC8389BE334C8*$10^{-\mathrm{A}}$. The precise value of the other parameters can be found in Section.7 [31].

We developed our programs with the public version 3.13p of the GNAT compiler [34]. As our type `Rep` of "representable numbers", we chose GNAT's floating point type `Long_Long_Float`. On the architecture used (Intel Pentium class PC), this type maps directly to the 80-bit "extended real" format [35] used by the floating point unit (FPU), which implements the IEEE "double extended" precision format, and complies with the IEEE Standard for Binary Floating-Point Arithmetic [36]. In particular, it is possible to specify that the four elementary operations never return a number that is less than the true result. As described in Subsection 6.2, this suffices for our purpose. Thus, we put the FPU in "round up" mode, for the duration of all of our programs: When the package `Intervals.Ops` is elaborated, a procedure `Round_Up` sets the binary FPU control word to `0000101101100000`. This one (assembly) instruction is machine dependent, and needs to be changed before running our programs on different hardware. It also un-masks all exceptions, besides rounding, and thus excludes de-normalized numbers (and infinite values) from std($\mathbb{R}$).

There are certainly other good choices for std($\mathbb{R}$). It should be noted that the quality of our bounds is not determined on the level of representable numbers or intervals. What is far more important (besides the general strategy of proof) is the control of the discretization errors that arise from estimating objects in infinite dimensions by a finite number of inequalities.

Our programs were compiled (at optimization level `O3`) and run on various PC's, using the GNU/Linux operating system. The estimated running time for a single 1GHz Pentium class PC is around 4 months. All programs terminated without signaling any errors, implying e.g. that the ball $B(2r)$ is contained in the domain of $\mathcal{M}$. The resulting upper bounds on the norms in (5.7), (5.9), and (5.10) are $K_1 < 7.0*10^{-18}$, $K_2 < 2.8*10^{-17}$, $K_1' < 4.2 * 10^{-14}$, $K_2' < 1.7 * 10^{-13}$, $\varepsilon < 9.0 * 10^{-15}$, $K < 0.84$, and it is straightforward to check that these constants satisfy the inequalities (5.11). Furthermore, we find that $c_H = 1.024332969\ldots$ for every Hamiltonian H in $H_0 + MB(r)$. This completes the proof of Theorem 5.2.

Acknowledgments
The author would like to thank P. Wittwer and M. Zou for many helpful discussions and suggestions, and the Department of Theoretical Physics at the University of Geneva for its hospitality, while part of this work was carried out. The visits to the University of Geneva were supported by the Swiss National Science Foundation under Grants No. 20-56931.99 and 20-64831.01.

References

[1] D.F. Escande and F. Doveil, *Renormalisation Method for Computing the Threshold of the Large Scale Stochastic Instability in Two Degree of Freedom Hamiltonian Systems,* J. Stat. Phys. **26** (1981), 257–284.
[2] S.J. Shenker and L.P. Kadanoff, *Critical Behavior of a KAM Surface. I. Empirical Results,* J. Stat. Phys. **27** (1982), 631–656.
[3] R.S. MacKay, *Renormalisation in Area Preserving Maps,* Thesis, Princeton (1982). World Scientific, London (1993).
[4] R.S. MacKay, *Renormalisation Approach to Invariant Circles in Area–Preserving Maps,* Physica D **7** (1983), 283–300.
[5] A. Mehr and D.F. Escande, *Destruction of KAM Tori in Hamiltonian Systems: Link with the Destabilization of Nearby Cycles and Calculation of Residues,* Physica D **13** (1984), 302–338.
[6] K. Khanin, Ya.G. Sinai, *The Renormalization Group Method and KAM Theory.* In "Nonlinear Phenomena in Plasma Physics and Hydrodynamics", R.Z. Sagdeev (ed), Mir (1986), 93–118.
[7] J.M. Greene, H. Johannesson, B. Schaub and H. Suhl, *Scaling Anomaly at the Critical Transition of an Incommensurate Structure,* Phys. Rev. A **36** (1987), 5858–5861.
[8] J. Wilbrink, *New Fixed Point of the Renormalization Operator for Invariant Circles,* Phys. Lett. A **131** (1988), 251–255.
[9] K. Khanin, Ya.G. Sinai, *Renormalization Group Methods in the Theory of Dynamical Systems.* Int. J. Mod. Phys. B **2** (1988), 147–165.
[10] J.A. Ketoja and R.S. MacKay, *Fractal Boundary for the Existence of Invariant Circles for Area–Preserving Maps: Observations and Renormalisation Explanation,* Physica D **35** (1989), 318–334.
[11] J. M. Mao and R. H. G. Helleman, *Existence of KAM Tori for Four–Dimensional Volume Preserving Maps.* Nuovo Cimento B **104** (1989), 177–182.
[12] J.M. Greene and J. Mao, *Higher–Order Fixed Points of the Renormalization Operator for Invariant Circles,* Nonlinearity **3** (1990), 69–78.
[13] J. Wilbrink, *New Fixed Point of the Renormalisation Operator Associated with the Recurrence of Invariant Circles in Generic Hamiltonian Maps,* Nonlinearity **3** (1990), 567–584.

[14] R. Artuso, G. Casati, D.L. Shepelyansky, *Breakdown of Universality in Renormalization Dynamics for Critical Invariant Torus.* Europhys. Lett. **15** (1991), 381–386.

[15] R.S. MacKay, J.D. Meiss, and J. Stark, *An Approximate Renormalization for the Break-up of Invariant Tori with Three Frequencies,* Phys. Lett. A **190** (1994), 417–424.

[16] R.S. MacKay, *Three Topics in Hamiltonian Dynamics.* Preprint U. Warwick (1994). Also in: "Dynamical Systems and Chaos", Vol.2, Y. Aizawa, S. Saito, K. Shiraiwa (eds), World Scientific, London (1995).

[17] J. del Castillo-Negrete, J. M. Greene, P. Morrison, *Area Preserving Non–Twist Maps: Periodic Orbits and Transition to Chaos.* Physica D **91** (1996), 1–23.

[18] A. Stirnemann, *Towards an Existence Proof of MacKay's Fixed Point,* Comm. Math. Phys. **188** (1997), 723–735.

[19] M. Govin, C. Chandre, H.R. Jauslin, *KAM–Renormalization Group Approach to the Breakup of Invariant Tori in Hamiltonian Systems.* Phys. Rev. E **57**, 1536–1543 (1998).

[22] C. Chandre, M. Govin, H.R. Jauslin, and H. Koch, *Universality for the Breakup of Invariant Tori in Hamiltonian Flows,* Phys. Rev. E **57** (1998), 6612–6617.

[21] J.J. Abad, H. Koch, and P. Wittwer, *A Renormalization Group for Hamiltonians: Numerical Results,* Nonlinearity **11** (1998), 1185–1194.

[22] H. Koch, *A Renormalization Group for Hamiltonians, with Applications to KAM Tori,* Erg. Theor. Dyn. Syst. **19** (1999), 1–47.

[23] C. Chandre, H. R. Jauslin, G. Benfatto, A. Celletti, *An Approximate Renormalization–Group Transformation for Hamiltonian Systems with Three Degrees of Freedom,* Phys. Rev. E **60** (1999), 5412–5421

[24 C. Chandre and R.S. MacKay, *Approximate Renormalization with Codimension–One Fixed Point for The Break-up of some Three–Frequency Tori,* Phys. Lett. A **275** (2000), 394–400.

[25] J.J. Abad and H. Koch, *Renormalization and Periodic Orbits for Hamiltonian Flows,* Commun. Math. Phys. **212** (2000), 371–394.

[26] J. Lopes Dias, *Renormalisation of Flows on the Multidimensional Torus Close to a KT Frequency Vector,* Nonlinearity **15** (2002), 647–664.

[27] J. Lopes Dias, *Renormalisation Scheme for Vector Fields on T^2 with a Diophantine Frequency,* Nonlinearity **15** (2002), 665–679.

[28] C. Chandre and H.R. Jauslin, *Renormalization–Group Analysis for the Transition to Chaos in Hamiltonian Systems,* Physics Reports **365** (2002), 1–64.

[29] C. Chandre, *Renormalization for Cubic Frequency Invariant Tori in Hamiltonian Systems with Two Degrees of Freedom,* Discrete Contin. Dynam. Systems B **2** (2002), 457–465.

[30] H. Koch, *On the Renormalization of Hamiltonian Flows, and Critical Invariant Tori,* Discrete Contin. Dynam. Systems A, **8**, 633–646 (2002).

[31] The preprint `mp_arc 02-175` of this paper includes the source code for the programs used in our proof of Theorem 5.2. See also `ftp://ftp.ma.utexas.edu/pub/papers/koch/kam-fixpt/`.

[32] H. Koch, A. Schenkel and P. Wittwer, *Computer–Assisted Proofs in Analysis and Programming in Logic: A Case Study,* SIAM Rev. **38** (1996), 565–604.

[33] S.T. Taft and R.A. Duff (eds), *Ada 95 Reference Manual: Language and Standard Libraries, International Standard ISO/IEC 8652:1995(E)*, Lecture Notes in Computer Science **1246** (1997). See also `http://www.adahome.com/rm95/`.

[34] Ada Core Technologies, 73 Fifth Avenue, New York, NY 10003, USA. See also `ftp://cs.nyu.edu/pub/gnat`.

[35] Intel Corp., *Intel Architecture Software Developer's Manual, Volume 1: Basic Architecture,* `http://developer.intel.com/design/pentiumii/manuals/243190.htm`.

[36] The Institute of Electrical and Electronics Engineers, Inc., *IEEE Standard for Binary Floating–Point Arithmetic,* ANSI/IEEE Std 754–1985.