A Renormalization Group for Hamiltonians, with Applications to KAM Tori

Hans Koch [1]
Department of Mathematics, University of Texas at Austin
Austin, TX 78712

Abstract. We describe a family of renormalization group transformations which apply to the problem of stability of invariant tori for analytic Hamiltonians. Each of these transformations corresponds to a fixed rotation vector, within a certain class of vectors in $\mathbb{R}^d$. As an application, we prove a KAM–type theorem that yields analytic invariant d–tori depending analytically on the Hamiltonian.

[1] Supported in Part by the National Science Foundation under Grant No. DMS–9401422, and by the Texas Advanced Research Program under Grant No. ARP–035

1. Introduction and Main Results

In this paper we describe a family of renormalization group transformations for analytic Hamiltonians in dimension $2d \geq 4$. Each of these transformations is associated with a vector $\omega = (1, \omega_2, \ldots, \omega_d)$ in $\mathbb{R}^d$ whose components span an algebraic number field of degree d, and is designed to apply to the stability problem for invariant d–tori with rotation vectors proportional to ω. Our transformations are similar in spirit to the block spin transformation in statistical mechanics, in the sense that they use a process of "elimination and rescaling". The elimination (of irrelevant frequencies) is done by means of canonical transformations, as in KAM theory [1–4], and the scaling (of the frequency lattice) involves an integral matrix T with eigenvector ω. In the case $d = 2$, with ω_2 equal to the golden mean, T is related to the composition operator used in the commuting map approach to dissipative invariant tori [17–19].

Although the concept is non-perturbative, our transformations are currently defined only near certain trivial (integrable) fixed points. Thus, the applications given in this paper fall in the domain of KAM theory. Connections between KAM theory and renormalization have been described before [6–16]. The most recent work in this area has concentrated on the problem of resumming the Lindstedt series [5], using techniques from quantum field theory to control cancelations and obtain a convergent "renormalized expansion" [13–16].

As in "ordinary" KAM theory, we apply an iterative procedure that eliminates irrelevant degrees of freedom from a given non-integrable Hamiltonian H_0. At the k-th step, H_{k-1} is transformed into a new Hamiltonian H_k which is in some sense equivalent to H_{k-1}, but closer to integrable. As usual, this step involves a canonical transformation, with a certain frequency cutoff. But there are also notable differences:

- The transformation $\Re\colon H_{k-1} \mapsto H_k$ acts on a fixed space of Hamiltonians.
- There is no loss of analyticity; in fact, $\Re$ is analyticity improving.
- There are no small denominator problems, and no major cancelations.
- On a conceptual level, the procedure is non-perturbative.

These are of course typical features of a (non-perturbative) renormalization group transformation. The absence of small denominators is e.g. due to the fact that $\Re$ includes a certain shift, which turns the "next scale" into the scale that has been considered before. The main problem here is to implement such a shift, and at the same time, improve analyticity. Before we explain in detail the action of $\Re$, let us try to sketch the main ideas of non-perturbative renormalization. After that, we will give a simple example, which in fact motivated the present work.

The typical setup for a renormalization group analysis is an open set $\mathcal{B}$ in some infinite dimensional space, and a densely defined function F on $\mathcal{B}$ that takes values in some n–dimensional manifold $\mathcal{Y}$. The function F plays the role of an observable, and the level sets of F represent objects that are equivalent with respect to the phenomenon that is being observed. Let us assume that some countable set $Y \subset \mathcal{Y}$ of possible values of F has been singled out as particularly interesting. In many cases of interest, the function F has a "critical surface" $\mathcal{W} \subset \mathcal{B}$, of the following type: There exist sequences $(y_0, y_1, \ldots)$ in Y, such that locally, the distance between $F^{-1}(y_k)$ and $\mathcal{W}$ decreases asymptotically like $|\lambda|^{-k}$, for some universal number λ of modulus > 1. In experiments on n-parameter families $V : A \subset \mathbb{R}^n \to \mathcal{B}$, the surface $\mathcal{W}$ shows up as a critical point of the function

$f = F \circ V$, and the behavior of f near this point appears to be independent of the family V. Due to this "universality", it is possible to determine all of the relevant quantities from a single function $f : A \to \mathcal{Y}$, and the system $(\mathcal{B}, F)$ need not even be known. In fact, the general picture described above is usually part of the (desired) explanation of universal critical behavior, observed in experiments on n-parameter families.

If the critical surface $\mathcal{W}$ is (believed to be) smooth, then the renormalization group method is a potentially powerful tool for analyzing such behavior. The idea is to construct a transformation $\mathcal{R} : B \subset \mathcal{B} \to \mathcal{B}$ that has a hyperbolic fixed point, whose local stable manifold coincides with the critical surface $\mathcal{W}$. Then every n–parameter family $V : A \to B$ that intersects $\mathcal{W}$ transversally near the fixed point, converges (under suitable reparametrization) to an invariant family describing the unstable manifold. But in order for this to explain universality, there has to be a certain relationship between the functions $F \circ \mathcal{R}$ and F. In all known examples, this holds by construction: There exists a (natural) map r on Y; and the transformation $\mathcal{R}$ is chosen such that $F(\mathcal{R}(h)) = y$, whenever $F(h) = r(y)$. In this case, if $(y_0, y_1, \ldots)$ is an orbit for r, then the abovementioned universal number λ should be the largest eigenvalue (in modulus) of the derivative of $\mathcal{R}$ at the fixed point. In other words: A renormalization group transformation contracts differences that are irrelevant (for the phenomenon being observed), and expands the relevant ones. It can also be viewed as a tool for introducing appropriate coordinates for the study of the observable F.

As a trivial (but instructive) example, consider the space $C^2(G)$ of twice continuously differentiable functions on a given disk $G \subset \mathbb{R}^2$, centered at the origin. To every function $h \in C^2(G)$, we can associate a Hamiltonian vector field $(q, p) \mapsto (\nabla h(p), 0)$, and thus a flow, on the product of the torus $\mathbb{T}^2$ with G. The corresponding orbit through the origin can be characterized by its slope $F(h) = w_2/w_1$, where $w = (\nabla h)(0)$. We now consider $F : \mathcal{B} \subset C^2(G) \to \mathbb{R}$ to be our observable. Here, $\mathcal{B}$ is a small open neighborhood of the following function h^ε. Let $\omega = (1, \vartheta)$ and $\omega' = (-\vartheta, 1)$, with ϑ equal to the golden mean $\frac{1}{2} + \frac{1}{2}\sqrt{5} = 1.618\ldots$, and define $h^\varepsilon(p) = \omega \cdot p + \varepsilon(\omega' \cdot p)^2$, for some fixed $\varepsilon \geq 0$. We note that ω and ω' are the two eigenvectors of the matrix $T = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, for the eigenvalues ϑ and $\vartheta' = -1/\vartheta$, respectively. As distinguished values of F, consider the best rational approximates $y_k = b_k/a_k$ for ϑ. That is, $y_0 = 1$, and $y_k = r(y_{k-1})$ if $k \geq 1$, where $r(y) = 1 + 1/y$. The corresponding integer vectors (a_k, b_k) can be obtained by iterating the map T, starting with $(a_0, b_0) = (1, 1)$. This shows that the sequence $(y_0, y_1, \ldots)$ converges to ϑ, and that $y_k - \vartheta$ is asymptotically proportional to $(-\vartheta^2)^{-k}$. Thus, if $v \mapsto h_v$ is a "typical" differentiable one-parameter family, passing near h^ε, we will find a value $v = v_\infty$ for which $F(h_v) = \vartheta$, and a sequence $(v_m, v_{m+1}, \ldots)$, accumulating at v_∞, such that $F(h_{v_k}) = y_k$, for all $k \geq m$. Furthermore, since F is actually differentiable near h^ε, the differences $v_k - v_\infty$ are asymptotically proportional to $(-\vartheta^2)^{-k}$.

Although there is not much left to explain in this example, it is still interesting to construct a useful renormalization group transformation. As a first guess, consider the map $\mathcal{R} : h \mapsto h \circ T^{-1}$. This map provides the desired relationship between $F \circ \mathcal{R}$ and F, described earlier. But unfortunately, it does not have a fixed point near h^ε. The following step is intended to correct this and other shortcomings, while keeping the essential features of our first guess. This means that $\mathcal{R}$ will be modified only in "irrelevant" directions. Here, the symmetries of F play an important role: $F(h)$ does not change if we add to h a constant function, or if we apply a scaling transformation $h \mapsto \mu^{-1}h(\mu.)$, with $\mu \neq 0$. In addition, h may be multiplied by a constant $b \neq 0$. The first two transformations do not affect the motion on the orbit passing through the origin. The third transformation (scaling of time) does not change any orbits, but the velocities get multiplied by b. Thus, let us modify our original guess for $\mathcal{R}$ by setting

$$\big(\mathcal{R}(h)\big)(p) = \frac{b}{\mu}\Big[h\big(\mu T^{-1}p\big) - h(0)\Big]\,, \tag{1.1}$$

where $b = b(h)$ and $\mu = \mu(h)$ are to be determined. We note that if we choose $b = \vartheta$ constant, then h^0 is a fixed point of $\mathcal{R}$, and if in addition $|\mu| = \vartheta^{-3}$, then each of the functions h^ε is a fixed point for $\mathcal{R}$. (The same holds for any constant multiples of these functions, if the values of b and μ are fixed.) In each of these cases, the derivative of $\mathcal{R}$, at the indicated fixed point, has an eigenvalue $-\vartheta^2$, and the corresponding eigenfunction is $p \mapsto \omega' \cdot p$. By choosing $b = b(h)$ and $\mu = \mu(h)$ appropriately, it is possible to make h^ε an isolated fixed point of $\mathcal{R}$, for any given ε, and to obtain $\lambda = -\vartheta^2$ as the only eigenvalue of $D\mathcal{R}(h^\varepsilon)$ of modulus ≥ 1. Then the stable manifold of $\mathcal{R}$ at h^ε coincides with the "critical surface" defined by the equation equation $F(h) = \vartheta'$. We will not give any proofs or details here, as this example is essentially a special case of what will be discussed later on.

Our next goal is to extend the transformation $\mathcal{R}$, and analogous transformations for higher dimensional systems, to "general" (but analytic) Hamiltonians $H = H(q,p)$ near a Hamiltonian of the form $H^0(q,p) = \omega \cdot p$. Conditions on the rotation vector $\omega \in \mathbb{R}^d$ will be specified below. In this paper, we are mainly interested in the properties of the renormalization group transformation, its local stable and unstable manifold, and in the existence of invariant tori with rotation vectors $c\omega$. The accumulation of closed orbits at these tori, and other aspects related to universality, will be discussed elsewhere [22]. Since we will no longer mention any observables, let us give (without proof) a possible characterization of the "critical" surface in this more general setting. A Hamiltonian near H^0 lies on $\mathcal{W}$ if it has an invariant d–torus whose rotation vector is a constant multiple of ω, and which is "centered at $p = 0$", in the sense that the integrals

$$A_j = \frac{1}{2\pi}\oint_{C_j} \sum_{i=1}^{d} p_i dq_i\,, \qquad j = 1, 2, \ldots, d\,, \tag{1.2}$$

vanish. Here, C_j denotes a closed path on the invariant d–torus with winding number 1 for the j-th 1-torus, and winding number 0 for the other 1-tori. We note that the quantities A_j are invariant under canonical transformations U_ϕ associated with a globally defined generating function ϕ. (For a definition of U_ϕ, see Section 2.)

In what follows, a Hamiltonian is an analytic function defined on a complex neighborhood of $\mathbb{T}^d \times \{0\}$, where $\mathbb{T}^d$ denotes the d–dimensional torus and 0 the origin in $\mathbb{R}^d$. Such a Hamiltonian may be lifted to a function on a complex neighborhood $\mathcal{D}$ of $\mathbb{R}^d \times \{0\}$, which is 2π–periodic in its first d arguments. The types of domains considered here are of the form $\mathcal{D}(\rho) = \mathcal{D}_1(\rho) \times \mathcal{D}_2(\rho)$, with

$$\mathcal{D}_1(\rho) = \big\{q \in \mathbb{C}^d\colon \|\mathrm{Im}q\| < \rho\big\}\,, \qquad \mathcal{D}_2(\rho) = \big\{p \in \mathbb{C}^d\colon \|p\| < \rho\big\}\,, \tag{1.3}$$

where $\|u\|$ denotes the ℓ^∞ norm of vector in $\mathbb{C}^d$; that is, $\|u\| = \max\{|u_1|, |u_2|, \ldots, |u_d|\}$. One way of specifying a Hamiltonian H on $\mathcal{D}(\rho)$ is by giving its Fourier-Taylor coefficients $H_{\nu,\alpha}$, defined e.g. implicitly by the equation

$$H(q,p) = \sum_{(\nu,\alpha)\in I} H_{\nu,\alpha}\, p^\alpha e^{i\nu\cdot q}\,, \qquad (q,p) \in \mathcal{D}(\rho)\,. \tag{1.4}$$

Here, $I = \mathbb{Z}^d \times \mathbb{N}^d$, and we have used the notation $p^\alpha = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_d^{\alpha_d}$ and $q \cdot \nu = q_1\nu_1 + q_2\nu_2 + \ldots + q_d\nu_d$.

Definition 1.1. *Given any positive real number ρ, define $\mathcal{A}(\rho)$ to be the Banach space of all analytic Hamiltonians H on $\mathcal{D}(\rho)$ which extend continuously to the boundary of $\mathcal{D}(\rho)$, and for which the norm*

$$\|H\|_\rho = \sum_{(\nu,\alpha)\in I} |H_{\nu,\alpha}| \rho^{|\alpha|} e^{\rho|\nu|} \tag{1.5}$$

is finite. Here, and in what follows, $|v|$ denotes the ℓ^1–norm of a vector or multi–index v; that is, $|v| = |v_1| + |v_2| + \ldots + |v_d|$. The identity operator on $\mathcal{A}(\rho)$ will be denoted by $\mathbb{I}$.

We would like to define a renormalization group transformation $\mathfrak{R}$ on $\mathcal{A}(\rho)$, which is essentially of the form (1.1), modulo some canonical change of coordinates that may depend on the Hamiltonian. The reason for such a choice is that we want to preserve the relationship, mentioned earlier, between certain orbits of the original and renormalized Hamiltonian. In the example where $T = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, this means that $\mathfrak{R}$ should include a step where $H(q,p)$ is replaced by $H(Tq,p)$. (Notice that the resulting change of coordinates is not homotopic to the identity, as in usual KAM theory.) But this step represents a highly singular operator on $\mathcal{A}(\rho)$, since T is expanding in the direction of ω. On the other hand, there are terms in the expansion (1.4) for which the substitution $q \mapsto Tq$ is harmless; namely, those that are indexed by pairs (ν,α) for which ν is approximately orthogonal to ω. These terms represent "slow modes" in the direction of the flow for $H^0(q,p) = \omega \cdot p$. They also happen to be the ones that produce small denominators in KAM theory.

One of the main ideas in our approach, motivated by an analogy (described below) with statistical mechanics, it to *completely* eliminate all non-slow modes, before applying the substitution $q \mapsto Tq$. This will be done by means of a canonical change of coordinates. An attempt at solving such an "elimination equation" to first order suggests the following definition of a slow mode. Here, and in the theorem that follows, ω can be an arbitrary vector in $\mathbb{R}^d$.

Definition 1.2. *Given real numbers $\rho, \sigma, \tau > 0$, and a vector $\omega \in \mathbb{R}^d$, denote by I^- the set of all pairs (ν,α) in $I = \mathbb{Z}^d \times \mathbb{N}^d$ which satisfy the condition*

$$|\omega \cdot \nu| > \sigma|\nu|\,, \qquad |\alpha| < \tau|\nu|\,. \tag{1.6}$$

A projection operator $\mathbb{I}^-$ is defined on $\mathcal{A}(\rho)$, by setting

$$\bigl(\mathbb{I}^- H\bigr)(q,p) = \sum_{(\nu,\alpha)\in I^-} H_{\nu,\alpha}\, p^\alpha e^{i\nu\cdot q}\,, \qquad (q,p) \in \mathcal{D}(\rho)\,. \tag{1.7}$$

In addition, define $I^+ = I \setminus I^-$ and $\mathbb{I}^+ = \mathbb{I} - \mathbb{I}^-$. The terms in the expansion (1.4), that are indexed by pairs (ν, α) in I^+, will be referred to as the "slow modes" of H in the direction of ω. (The plus sign in I^+ refers to the large time it takes for a slow mode to affect the motion in the direction of ω.)

Notice the strict inequalities in (1.6). According to this definition, a Hamiltonian $H = H(q,p)$ that only depends on p has only slow modes.

In an analogy with statistical mechanics, I plays the role of a lattice, indexing the spin variables, and I^+ represents a "coarser" sublattice, used to index local spin averages, also called block spins. The elimination of fast (non-slow) modes corresponds to the process of integrating out "irrelevant fluctuations" around given local averages. In order to push the analogy further, consider again the matrix $T = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ and its unstable eigenvector $\omega = (1, \vartheta)$. Then the shift $(\nu, \alpha) \mapsto (T\nu, \alpha)$, applied to points in I^+, corresponds to a rescaling of the coarse lattice, designed to turn a system of block spins back into a regular spin system, so that the entire process can be repeated.

The following theorem describes the elimination of fast (non-slow) modes in the direction of a fixed but arbitrary vector $\omega \in \mathbb{R}^d$, for Hamiltonians near $H^0 : (q,p) \mapsto \omega \cdot p$.

Theorem 1.3. *Let $0 < \rho' < \rho$ and $\sigma, \tau > 0$. Then there exists an open neighborhood B of H^0 in $\mathcal{A}(\rho)$, and for every $H \in B$ a canonical transformation $\mathcal{U}_H$ from $\mathcal{D}(\rho')$ into $\mathcal{D}(\rho)$, such that $\mathbb{I}^-(H \circ \mathcal{U}_H) = 0$, and such that $\mathcal{U}_H$ is the identity whenever $\mathbb{I}^- H = 0$. The map $H \mapsto H \circ \mathcal{U}_H$ is analytic from B to $\mathbb{I}^+\mathcal{A}(\rho')$, and its derivative at H^0 is $\mathbb{I}^+$.*

In the remaining part of this introduction, ρ denotes a fixed (but arbitrary) positive real number. We will also assume that $\omega = (1, \omega_2, \ldots, \omega_d)$ is a given vector in $\mathbb{R}^d$ whose components span an algebraic number field of degree d. In particular, $\omega \cdot \nu$ is nonzero for every vector $\nu \in \mathbb{Z}^d \setminus \{0\}$. Another consequence of this assumption (see Section 4) is that there exists an integral $d \times d$ matrix T with the following properties:

(T1) T has a simple eigenvalue ϑ of modulus larger than 1, and $T\omega = \vartheta\omega$.
(T2) All other eigenvalues of T are simple, nonzero, and of modulus less than 1.
(T3) $\det(T) = \pm 1$.

For the vector $\omega = (1, \omega_2)$, with ω_2 the golden mean $\frac{1}{2} + \frac{1}{2}\sqrt{5}$, we have $T = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ as an example of such a matrix.

Let now T be a fixed integral matrix satisfying (T1–2), but not necessarily (T3). Then we define a linear map T_μ on $\mathbb{C}^d \times \mathbb{C}^d$, depending on a parameter $\mu \in \mathbb{C}$, by setting

$$T_\mu(q,p) = \bigl(Tq, \mu(T^*)^{-1}p\bigr), \qquad q, p \in \mathbb{C}^d. \tag{1.8}$$

Notice that if the matrix T satisfies (T3) as well (implying that T^{-1} is integral), then T_1 represents a canonical transformation of the torus $\mathbb{T}^d$, preserving the 1-form integrated in equation (1.2).

Throughout the rest of this paper, we assume that

$$\max_{\substack{|y|=1 \\ \omega \cdot y = 0}} \bigl|T^* y\bigr| < \kappa, \qquad 0 < \max_{|y|=1} \bigl|\mu T^{-1} y\bigr| < \kappa, \tag{1.9}$$

for some positive real number $\kappa < 1$, which will be regarded as fixed. The first condition in (1.9) is purely technical and could be avoided, if necessary, by choosing an appropriate norm on $\mathbb{R}^d$, or by replacing T with a sufficiently large power of T. The second inequality in (1.9) should be viewed a condition on the parameter μ.

Consider now a Hamiltonian H that contains no slow modes. If the parameter σ has been chosen very large, then all frequencies ν for which $H_{\nu,\alpha}$ is nonzero are contracted by the matrix T^*. As a result, we have

Lemma 1.4. *Let $\rho' > \kappa\rho$ be given. If $\sigma > 0$ is sufficiently small and τ sufficiently large, then every function $H \in \mathbb{I}^+\mathcal{A}(\rho')$ extends analytically to $T_\mu\mathcal{D}(\rho)$, and $H \mapsto H \circ T_\mu$ is a compact linear map from $\mathbb{I}^+\mathcal{A}(\rho')$ to $\mathcal{A}(\rho)$.*

This lemma will be applied later for a fixed choice of ρ'. Thus, the parameters σ and τ will now be considered fixed as well; for details see Section 4.

Now we are ready to define our renormalization group transformation. Denote by Ω the eigenvector of T^* corresponding to the eigenvalue ϑ, normalized such that $\Omega \cdot \omega = 1$. If H belongs to $\mathcal{A}(\rho)$, denote by $\mathbb{E}H$ the average of H over the torus $\mathbb{T}^d$, that is,

$$(\mathbb{E}H)(p) = (2\pi)^{-d} \int_{\mathbb{T}^d} H(q,p)dq = \sum_{\alpha\in\mathbb{N}^d} H_{0,\alpha}p^\alpha . \tag{1.10}$$

Definition 1.5. *Given a Hamiltonian H sufficiently close to H^0, we define a renormalized Hamiltonian $\Re_\mu(H)$ by setting*

$$\Re_\mu(H) = \frac{1}{(\Omega\cdot\nabla\mathbb{E}\tilde{H})(0)}\bigl[\tilde{H} - (\mathbb{E}\tilde{H})(0)\bigr], \qquad \tilde{H} = \frac{\vartheta}{\mu} H\circ\mathcal{U}_H\circ T_\mu . \tag{1.11}$$

That is, $\Re_\mu$ is composed of a H–dependent canonical change of coordinates (for a global generating function that depends analytically on H), a fixed scaling of phase space, and a H–dependent scaling of time. The scaling parameter μ need not depend on H in this case, due to a symmetry of the fixed point H^0 (see below). Notice that multiplying a Hamiltonian by a constant factor does not change the orbits for the corresponding flow. The role of the factor ϑ in the above definition is to ensure that $\Re_\mu(H^0) = H^0$. To see that H^0 is indeed a fixed point, it suffices to consider $\mu = 1$, as will become clear below. In this case, we have $\widetilde{H^0}(q,p) = \vartheta\omega\cdot\bigl((T^*)^{-1}p\bigr) = \vartheta\bigl(T^{-1}\omega\bigr)\cdot p = \omega\cdot p$, and consequently $\Re_1(H^0) = H^0$.

Concerning the the parameter μ in the definition (1.11), we would like to argue that, for a renormalization group analysis near H^0, almost any choice is as good as any other. Given $u \in \mathbb{C}^d$, and $z \in \mathbb{C}$ non-zero, define

$$\bigl(\mathcal{J}_u H\bigr)(q,p) = H(q-u,p), \qquad \bigl(\mathcal{S}_z H\bigr)(q,p) = zH(q,z^{-1}p), \tag{1.12}$$

whenever possible. Then $\Re_\mu$ can be written as $\mathcal{S}_\mu^{-1}\circ\Re_1$. Notice that the transformations $\mathcal{S}_z$, for $0 \neq z \in \mathbb{C}$, form a group under composition (for a suitable class of H's). They modify the dynamics in a trivial way, by scaling the momentum p. Thus, the Hamiltonians

$\mathcal{S}_z H$, for different values of z, may be regarded as equivalent. In particular, $\Re_\mu(H)$ is in the same equivalence class as $\Re_1(H)$. And since the Hamiltonian H^0 is scale invariant ($\mathcal{S}_z H^0 = H^0$), it is a fixed point of $\Re_\mu$ for any choice of $\mu \neq 0$.

Theorem 1.6. *There exists an open ball B in $\mathcal{A}(\rho)$, centered at $H^{(0)}$, such that the transformation $\Re_\mu$ is well defined, analytic, and compact, as a map from B to $\mathcal{A}(\rho)$. The same holds for the map $\Re_\mu \circ \mathcal{J}_u = \mathcal{J}_{T^{-1}u} \circ \Re_\mu$, if the imaginary part of u is sufficiently small. Furthermore, $\Re_\mu$ maps $B \cap \mathcal{S}_z^{-1} B$ to $\mathcal{A}(\rho) \cap \mathcal{S}_z^{-1}\mathcal{A}(\rho)$, for any nonzero complex number z, and $\mathcal{S}_z \circ \Re_\mu = \Re_\mu \circ \mathcal{S}_z$ on $B \cap \mathcal{S}_z^{-1} B$.*

The first part of this theorem is a straightforward consequence of Theorem 1.3 and Lemma 1.4. In the second part, we list two symmetries of $\Re_\mu$ that will play an important role later on, especially the "commutation relation" involving $\mathcal{J}_u$. The fact that $\Re_\mu$ commutes with $\mathcal{S}_z$, is used mainly to improve bounds that are not manifestly scaling invariant.

In order to describe the action of $\Re_\mu$ near the fixed point H^0, we need to analyze the derivative of $\Re_\mu$ at H^0. This part is easy, as usual, since we are dealing with a trivial fixed point. Consider the map $H \mapsto \tilde{H}$ defined by equation (1.11). ¿From Theorem 1.3 it follows that the derivative of this map at H^0 is given by the operator L_μ,

$$L_\mu h = \frac{\vartheta}{\mu}\big(\mathbb{I}^+ h\big) \circ T_\mu \,. \tag{1.13}$$

The eigenvectors of L_μ can be determined explicitly. Namely, let $\vartheta_1, \vartheta_2, \ldots, \vartheta_d$ be the eigenvalues of T, labeled such that $|\vartheta_i| \geq |\vartheta_j|$ for $i \leq j$, and let $\omega^{(1)}, \omega^{(2)}, \ldots, \omega^{(d)}$ be the corresponding eigenvectors. In particular, $\vartheta_1 = \vartheta$ and $\omega^{(1)} = \omega$. Then the nonzero eigenvalues of L_μ, and corresponding eigenvectors, are given by

$$\lambda_\beta = \frac{\vartheta}{\mu} \prod_{j=1}^{d} \left(\frac{\mu}{\vartheta_j}\right)^{\beta_j} , \qquad \phi_\beta(q,p) = \prod_{j=1}^{d} \big(\omega^{(j)} \cdot p\big)^{\beta_j} , \tag{1.14}$$

where $\beta = (\beta_1, \beta_2, \ldots, \beta_d)$ and $\beta_j \geq 0$ for all j. The second part of assumption (1.9) implies that $|\mu/\vartheta_j| < 1$ for all j. If we impose the additional condition

$$|\mu| < \frac{\vartheta_d^2}{|\vartheta_1|} \,, \tag{1.15}$$

then all eigenvalues of L_μ are of modulus less than 1, except those corresponding to $|\beta| \leq 1$. The derivative of $\Re_\mu$ at H^0 is given by the equation

$$D\Re_\mu\big(H^0\big)h = \tilde{h} - (\mathbb{E}\tilde{h})(0) - \big(\Omega \cdot \nabla \mathbb{E}\tilde{h}\big)(0) H^0 \,, \qquad \tilde{h} = L_\mu h \,. \tag{1.16}$$

It is easy to check that the eigenvectors of $D\Re_\mu(H^0)$ are exactly the same as those of L_μ. The eigenvalues are the same as well, except for two cases: The constant function $\phi_{(0,\ldots,0)} = 1$ and the function $\phi_{(1,0,\ldots,0)} = H^0$ are mapped to zero by $D\Re_\mu(H^0)$.

In what follows, we assume that μ satisfies the condition (1.15). Then the derivative of $\Re_\mu$ at H^0 has precisely $d-1$ eigenvalues of modulus ≥ 1. These eigenvalues, and the corresponding eigenvectors, are

$$\lambda'_j = \frac{\vartheta_1}{\vartheta_j}\,, \qquad h_j(q,p) = \omega^{(j)} \cdot p\,, \qquad j = 2, \ldots, d. \tag{1.17}$$

The $d-1$ dimensional affine space consisting of all functions of the form $H^0 + v_1 h_1 + \ldots + v_d h_d$ coincides in fact (near H^0) with the local unstable manifold $\mathcal{W}^u$ of the transformation $\Re_\mu$. This holds independently of the value of μ, except for the precise definition of "near". In the same sense, the (codimension $d-1$) local stable manifold $\mathcal{W}^s$ of $\Re_\mu$ at H^0 is independent of μ. This follows essentially from the last part of Theorem 1.6; see Section 4 for details.

Now we turn to some applications of the transformation $\Re_\mu$. We consider the flow defined by the vector field $\mathbb{J}\nabla H$, associated with a Hamiltonian $H \in \mathcal{A}(\rho)$. Here, $\mathbb{J}$ is the linear transformation on $\mathbb{C}^{2d}$, defined by the equation $\mathbb{J}(q,p) = (p,-q)$. The goal is to find invariant tori with rotation vectors proportional to ω.

Since H is not necessarily real-valued for real arguments, it is natural to allow time to be a complex variable. But as a result, the flow for a Hamiltonian H near H^0 is never defined for all times. Consider e.g. a domain of the form $\mathbb{T}^d(r') \times \mathcal{D}_2(r')$, where $\mathbb{T}^d(r')$ is the complexified torus $\mathcal{D}_1(r')/(2\pi\mathbb{Z})^d$. On this domain, consider the Hamiltonian $H = cH^0$, for some nonzero complex number c. For this Hamiltonian, the sets $\mathbb{T}^d(r') \times \{p'\}$, with $p' \in \mathcal{D}_2(r')$, are invariant under the time evolution $(\dot q, \dot p) = (\mathbb{J}\nabla H)(q,p) = (c\omega, 0)$. We will refer to these set as the invariant tori of H, with rotation vector $c\omega$. The flow on these tori, with initial conditions in a given compact subset K, is defined only for times $t \in \mathbb{C}$ such that $q + tc\omega$ belongs to $\mathcal{D}(r')$, whenever $q \in K$. However, this includes all times in $c^{-1}\mathbb{R}$. Along the corresponding orbits, the motion is quasiperiodic with rotation vector $c\omega$.

Theorem 1.7. *Given two positive real numbers r' and $r > r' + \rho$, there exists an open neighborhood B of H^0 in $\mathcal{A}(r)$, and for every $H \in B$ a complex number $\xi'(H)$ and an analytic function $\Gamma_H : \mathcal{D}_1(r') \to \mathcal{D}(r)$, such that the following holds. If $H \in \mathcal{W}^s \cap B$ then*

$$\bigl(\mathbb{J}\nabla H\bigr) \circ \Gamma_H = \xi'(H)\, \omega \cdot \nabla \Gamma_H \qquad \text{on } \mathcal{D}_1(r')\,. \tag{1.18}$$

If in addition, $H(q,p)$ depends on p only, then $\xi'(H) = (\Omega \cdot \nabla \mathbb{E} H)(0)$ and $\Gamma_H(q) = (q,0)$, for all $q \in \mathcal{D}_1(r')$. Furthermore, ξ' is an analytic function on B, and $H \mapsto \Gamma_H - \Gamma_{H^0}$ is an analytic map from B to some Banach space of analytic 2π-periodic functions on $\mathcal{D}_1(r')$.

In other words, if $H \in \mathcal{A}(r)$ is sufficiently close to H^0 and lies on the local stable manifold $\mathcal{W}^s$, then the corresponding time evolution on $\mathcal{D}(r)$ leaves invariant a "torus" $\Gamma_H(\mathbb{T}^d(r'))$, and the flow on this torus is conjugate to the flow of the Hamiltonian $\xi'(H)H^0$ on the torus $\mathbb{T}^d(r') \times \{0\}$. If in addition, H is real-valued for real arguments, then $\xi'(H) \in \mathbb{R}$ and $\Gamma_H(\mathbb{T}^d) \subset \mathbb{R}^d$. This fact does not follow from Theorem 1.7 directly, but from our definition (construction) of the quantities $\xi'(H)$ and Γ_H.

We note that, as usual, Γ_H is obtained as the limit of a sequence of canonical transformations (close to the identity). These transformation are extracted from the product

$\mathcal{U}_{H_0} \circ T_\mu \circ \mathcal{U}_{H_1} \circ T_\mu \circ \ldots \circ \mathcal{U}_{H_{n-1}} \circ T_\mu$ that appears in the expression for $\mathfrak{R}_\mu^n(H)$, by commuting all of the factors T_μ out to the right. In this process, analyticity is lost. Fortunately, it suffices to determine the values of Γ_H on the real torus $\mathbb{T}^d$. The result can be continued analytically, by using the values on $\mathbb{T}^d$, of the functions $\Gamma_{\mathcal{J}_u(H)}$ for complex translates of the Hamiltonian H.

Before we discuss Hamiltonians that lie off the stable manifold $\mathcal{W}^s$, let us first list some that can be shown to lie on $\mathcal{W}^s$; for details see Section 4. In the following examples, we assume that H is close to H^0 in $\mathcal{A}(\rho)$.

(H1) $H(q,p) = c\omega \cdot p + f(q)$.
(H2) $H(q,p) = c\omega \cdot p + f(q,p)$, with $f(q,p) = \mathcal{O}\bigl(|p|^2\bigr)$.
(H3) $H^s = H + v_2 h_2 + \ldots + v_d h_d$, with $v \in \mathbb{C}^{d-1}$ chosen appropriately (see below).

The Hamiltonian H^s in example (H3) is simply the projection, in the direction of $\mathcal{W}^u$, of H onto the local stable manifold $\mathcal{W}^s$. (Since $\mathcal{W}^s$ is analytic, v depends analytically on H.) In the language of quantum field theory, $v_2 h_2 + \ldots + v_d h_d$ is a "counterterm". It is of course possible to project in a different direction (use different counterterms), as long as it is transversal to $\mathcal{W}^s$ near H^0. One can also try to "project" along the orbit of translations in p. To be more specific, consider

$$H_v = H \circ R_v^{-1}\,, \qquad R_v(q,p) = (q, p+v)\,, \tag{1.19}$$

whenever defined.

Observation 1.8. *If there exists $v \in \mathbb{C}^d$ such that H_v is well defined and belongs to $\mathcal{W}^s \cap B$, where B is the set described in Theorem 1.7, then H has an invariant torus, given by the function $R_v \circ \Gamma_{H_v}$.*

Since $\mathcal{W}^s$ is of codimension $d-1$, we can expect that Observation 1.8 applies to open sets of Hamiltonians in $\mathcal{A}(r'')$, if $r'' > r$, and that most of these Hamiltonians have a one-parameter family of invariant tori, covering an invariant surface of dimension d in phase space. To give some examples, denote by $E \subset \mathbb{C}^d$ the linear span of the eigenvectors $\omega^{(2)}, \ldots, \omega^{(d)}$ of the matrix T.

Theorem 1.9. *Given $r'' > r > r' + \rho$, there exists an open neighborhood B of H^0 in $\mathcal{A}(r'')$, such that the following holds. Assume $H \in B$ is of the form $H(q,p) = \omega \cdot p + f(p)$, with $(\nabla f)(0) = 0$. If the range of $(D\nabla f)(0)$ contains E, then for every function H' in some open neighborhood $B_H \subset B$ of H, the intersection of the set $\{H'_v : v \in \mathcal{D}_2(r'' - r)\}$ with $\mathcal{W}^s$ is non-empty and contains the range of an analytic curve $\Upsilon_{H'} : A \to \mathcal{A}(r)$, where A is some open set in $\mathbb{C}$ containing 0. The map $H' \mapsto \Upsilon_{H'}$ is analytic from B_H to $C^1(A, \mathcal{A}(r))$, and $\Upsilon_H(0) = H$. Furthermore, if $(D\nabla f)(0)$ is non-singular, then the derivative of $\xi' \circ \Upsilon_{H'}$ is non-vanishing on A, for all $H' \in B_H$.*

The last statement in this theorem implies that (under the given assumptions) every $H' \in B_H$ has invariant tori with rotation vectors $c\omega$, for an open set of values $c \in \mathbb{C}$. Since $\xi'(H) = 1$, this set contains an interval of real numbers including 1, if H' is sufficiently close to H.

Remarks.

○ The last two theorems can be combined into one that contains no reference to the transformation $\mathfrak{R}_\mu$. In particular: When considering a fixed space $\mathcal{A}(r'')$, there are open sets $\mathcal{B}(r')$, for any $r' < r''$, such that every Hamiltonian in $\mathcal{B}(r')$ has a one-parameter family of invariant tori (with rotation vector $c\omega$ for varying c) that are analytic on $\mathcal{D}_1(r')$.

○ Concerning our restriction to Hamiltonians H near H^0 in Theorem 1.9, it should be mentioned that every Hamiltonian of the form $H(q,p) = \omega \cdot p + f(p)$, with f analytic in some open neighborhood of $\mathbb{T}^d \times \{0\}$, satisfying $(\nabla f)(0) = 0$, is "close" to H^0, when restricted to the appropriate domain. Namely, given any open neighborhood B of H^0 in $\mathcal{A}(r'')$, for any $r'' > 0$, it is always possible to find a nonzero complex (or real) number z of modulus ≤ 1, such that $\mathcal{S}_z^{-1}H$ belongs to B. And if $\mathcal{S}_z^{-1}H'$ has an invariant torus with rotation vector proportional to ω, then so does H'.

○ The fixed point H^0 is trivial in the sense that it has no fast modes (or fluctuations, like a mean field model in statistical mechanics), and that the relevant eigenvalues of $D\mathfrak{R}_\mu(H^0)$ are algebraic expressions in the eigenvalues of the scaling matrix T. The same is true for the fixed points H^ε (of a slightly modified transformation) corresponding to the functions h^ε discussed after equation (1.1). The behavior near these fixed points cannot really be considered "critical".

○ A more general class of renormalization group transformations is obtained if we (choose $\mu = \mu(H)$ and) replace the step $H \mapsto \vartheta H$ by a nonlinear scaling of time $H \mapsto \Lambda_H \circ H$. The obvious conjecture would be that some of these transformations have nontrivial fixed points. Of course, the map $H \mapsto \mathcal{U}_H$ first has to be defined properly, e.g. by using computer–assisted methods, in some neighborhood of the expected fixed point H^*.

○ An approximate renormalization group transformation, for $d = 2$ and Hamiltonians of the form $H(q,p) = p_1 + f(q, p_2)$, was studied in [6,8]. The results, and work on area–preserving maps [7,9], indicate that a nontrivial fixed point H^* may be related to the breakup of smooth invariant tori for the corresponding rotation vector ω.

The remaining part of this paper is organized as follows: Sections 2 and 3 deal exclusively with the canonical transformation $\mathcal{U}_H$ and its generating function. In Section 4, after a lemma concerning properties (T1–3), we describe the renormalization group transformation $\mathfrak{R}_\mu$, the spectrum of its derivative at H^0, and its local stable manifold $\mathcal{W}^s$. This includes a discussion of some symmetry properties of $\mathfrak{R}_\mu$ and $\mathcal{W}^s$, that are useful for the construction of invariant tori with rotation vectors $c\omega$. These invariant tori are the subject of Section 5.

2. The Canonical Transformation $\mathcal{U}_H$

In this section we prove Theorem 1.3. The following quantities are assumed to be fixed: Two positive real numbers $\varrho' < \varrho$, a vector $\omega \in \mathbb{R}^d$, and the two parameters $\sigma, \tau > 0$ that enter the definition of the projections $\mathbb{I}^{\pm}$; see Definition 1.2. (The proof of Theorem 1.3 uses $\varrho = \rho$ and $\varrho' = \rho'$; but in Section 4, we will choose $\rho' < \varrho' < \varrho < \rho$.)

If X and Y are Banach spaces, we denote by $\mathcal{L}(X,Y)$ the Banach space of all continuous linear operators from X to Y. We will often use the following fact (see e.g. [20]), without mentioning it explicitly.

Theorem 2.1. *Let X and Y be Banach spaces over the complex numbers, and let $\mathcal{F}$ be a locally bounded map from some open set $D \subset X$ to Y. Then the following properties are equivalent:*

(a) *$\mathcal{F}$ is Gâteaux differentiable on D.*

(b) *$\mathcal{F}$ is of class C^∞ on D.*

(c) *$\mathcal{F}$ is complex analytic on D. That is, for every $x \in D$, $h \in X$, and $\ell \in \mathcal{L}(Y,\mathbb{C})$, the function $z \mapsto \ell\bigl(\mathcal{F}(x+zh)\bigr)$ is analytic in a neighborhood of zero.*

A map $\mathcal{F}$ of the type described above, which has (one of) the three properties (a,b,c) will be referred to as a "differentiable" or "analytic" map from D to Y.

By definition, a transformation $U : (q,p) \mapsto (q+Q, p+P)$ is canonical if and only if the 1-form $(p+P)\cdot d(q+Q) - p\cdot dq$ is closed. In what follows, we will only be interested in cases where this 1-form is the differential of some function S. The function $\phi = S - Q\cdot p$, expressed in terms of $q+Q$ and p, will be referred to as the generating function for U. It satisfies the equations

$$\begin{aligned} Q(q,p) &= -\nabla_2\phi\bigl(q+Q(q,p),p\bigr),\\ P(q,p) &= \nabla_1\phi\bigl(q+Q(q,p),p\bigr), \end{aligned} \tag{2.1}$$

where $\nabla_1\phi$ and $\nabla_2\phi$ denote the gradients of the functions $q \mapsto \phi(q,.)$ and $p \mapsto \phi(.,p)$, respectively. Conversely, if ϕ is given and sufficiently small (see below), then (2.1) defines a transformation $U = U_\phi$,

$$U_\phi(q,p) = \bigl(q+Q(q,p), p+P(q,p)\bigr), \tag{2.2}$$

and this transformation is canonical.

The main goal of this section is to find a solution $\mathcal{U}_H$ of the equation $\mathbb{I}^-(H\circ\mathcal{U}_H) = 0$, given $H \in \mathcal{A}(\rho)$ sufficiently close to H^0. We will solve this equation by an iterative procedure, which is similar to a Newton algorithm; but it involves certain choices, since the solution is not unique (we could e.g. increase σ and/or decrease τ). In the first step, we determine a canonical transformation U_ϕ , where ϕ depends on H, such that $\mathbb{I}^-(H\circ U_\phi)$ vanishes to first order in $\mathbb{I}^-H$. (Then H is replaced by $H\circ U_\phi$ and the process is iterated.) We start by giving an informal description of this step. In order to simplify notation, the image under the projection $\mathbb{I}^{\pm}$ of a function f will be denoted by $f^{\pm}$.

Assume that H is of the form $H = H^0 + f$, and imagine that f depends on a (small) parameter ε in such a way that f^- is of order ε. The goal is to obtain $\mathbb{I}^-(H\circ U_\phi) = \mathcal{O}(\varepsilon^2)$.

Starting with ϕ of order ε, we expect equation (2.1) to have a solution $Q = -\nabla_2\phi + \mathcal{O}(\varepsilon^2)$ and $P = \nabla_1\phi + \mathcal{O}(\varepsilon^2)$. Thus, the following quantities should be of order ε^2.

$$\begin{aligned}
R_1(q,p) &= \omega \cdot \bigl(P(q,p) - \nabla_1\phi(q,p)\bigr)\,, \\
R_2(q,p) &= f^-\bigl(q + Q(q,p), p + P(q,p)\bigr) - f^-(q,p)\,, \\
R_3(q,p) &= f^+\bigl(q + Q(q,p), p + P(q,p)\bigr) - f^+\bigl(q - \nabla_2\phi(q,p), p + \nabla_1\phi(q,p)\bigr)\,, \\
R_4(q,p) &= f^+\bigl(q - \nabla_2\phi(q,p), p + \nabla_1\phi(q,p)\bigr) - f^+(q,p) + \bigl\{f^+,\phi\bigr\}(q,p)\,,
\end{aligned} \tag{2.3}$$

where $\{f,g\}$ denotes the Poisson bracket of two functions f and g, that is, $\{f,g\} = \nabla_1 f \cdot \nabla_2 g - \nabla_2 f \cdot \nabla_1 g$. The functions R_i defined by equation (2.3) are used in the following decomposition of $H \circ U_\phi$, where U_ϕ is defined by equation (2.2).

$$H \circ U_\phi = H^0 + f^+ - \mathbb{I}^+\bigl\{f^+,\phi\bigr\} + \Bigl[\omega \cdot \nabla_1\phi + f^- - \mathbb{I}^-\bigl\{f^+,\phi\bigr\}\Bigr] + \sum_{i=1}^{4} R_i\,. \tag{2.4}$$

We will now prove that there exists a function ϕ for which the expression $[\ldots]$ in (2.4) vanishes. Then $H \circ U_\phi$ has the desired properties, as will be shown later.

First we list some basic properties of the spaces $\mathcal{A}(r)$. Given any positive integer n, denote by $\mathcal{A}^n(r)$ the direct sum of n copies of $\mathcal{A}(r)$, equipped with the norm $\|(V_1,\ldots,V_n)\|_r = \max\bigl\{\|V_1\|_r,\ldots,\|V_n\|_r\bigr\}$. Elements $V \in \mathcal{A}^n(r)$ will also be regarded as maps $x \mapsto (V_1(x),\ldots V_n(x))$ from $\mathcal{D}(r)$ to $\mathbb{C}^n$. In what follows, $\partial_j f$ denotes the partial derivative of a function f with respect to its j–th argument.

Proposition 2.2. *Let $0 < \delta < r$ and $f, g \in \mathcal{A}(r)$. Furthermore, let $P, Q \in \mathcal{A}^d(r-\delta)$ and define $V(q,p) = \bigl(q + Q(q,p), P(q,p)\bigr)$ for all $(q,p) \in \mathcal{D}(r-\delta)$. Then*
(a) $|f(x)| \le \|f\|_r$ *for all* $x \in \mathcal{D}(r)$.
(b) $\|\partial_j f\|_{r-\delta} \le \delta^{-1}\|f\|_r$ *for* $1 \le j \le 2n$, *if* $0 < \delta < r$.
(c) $fg \in \mathcal{A}(r)$ *and* $\|fg\|_r \le \|f\|_r\|g\|_r$.
(d) $f \circ V \in \mathcal{A}(r-\delta)$ *and* $\|f \circ V\|_{r-\delta} \le \|f\|_r$, *if* $\|Q\|_{r-\delta} \le \delta$ *and* $\|P\|_{r-\delta} \le r$.

The proof of these estimates is straightforward and will be omitted.

Definition 2.3. *Given $r > 0$, define $\mathcal{A}'(r)$ to be the the vector space of all analytic functions on $\mathcal{D}(r)$ whose first partial derivatives all lie in $\mathcal{A}(r)$. On this space we consider the norm $\|.\|'_r$ defined by $\|f\|'_r = \max\bigl\{\|f\|_r, \|\nabla_1 f\|_r, \|\nabla_2 f\|_r\bigr\}$.*

The following lemma indicates that the non-slow modes can indeed be eliminated. This is one of the two places where Definition 1.2 enters in a crucial way. The other place is Lemma 1.4, which shows that the slow modes can be “saved for later”.

Lemma 2.4. *Let $r > 0$, define $b = \sigma/(1+\tau/r)$, and denote by B the open ball of radius $b/2$ in $\mathcal{A}'(r)$. Then there exists a differentiable map φ from B to the Banach space of all bounded linear operators from $\mathcal{A}(r)$ to $\mathbb{I}^-\mathcal{A}'(r)$, such that for every $h \in B$ and every $f \in \mathcal{A}(r)$, the function $\phi = \varphi(h)f$ satisfies the equation*

$$\omega \cdot \nabla_1\phi + f^- - \mathbb{I}^-\bigl\{h^+,\phi\bigr\} = 0 \tag{2.5}$$

and the bounds

$$\|\phi\|_r\,,\|\nabla_1\phi\|_r \le \frac{2}{\sigma}\|f^-\|_r\,,\qquad \|\nabla_2\phi\|_r \le \frac{2\tau}{\sigma r}\|f^-\|_r\,. \tag{2.6}$$

Proof. ¿From the condition (1.6) defining the set I^-, it follows that the operator $\omega\cdot\nabla_1$ has a bounded inverse on $\mathbb{I}^-\mathcal{A}(r)$, given by the equation

$$\bigl((\omega\cdot\nabla_1)^{-1}f\bigr)(q,p) = \sum_{(\nu,\alpha)\in I^-}\frac{1}{i\omega\cdot\nu}\,f_{\nu,\alpha}\,p^\alpha e^{i\nu\cdot q}\,. \tag{2.7}$$

In this sum, the term indexed by (ν,α) is an eigenfunction for $\frac{\partial}{\partial q_j}$ and for $p_j\frac{\partial}{\partial p_j}$, with eigenvalue $i\nu_j$ and α_j , respectively. Since ν and α satisfy the condition (1.6), it follows that $\nabla_1(\omega\cdot\nabla_1)^{-1}$, and $\nabla_2(\omega\cdot\nabla_1)^{-1}$ define continuous linear operators from $\mathbb{I}^-\mathcal{A}(r)$ to $\mathbb{I}^-\mathcal{A}^d(r)$, with operator norms bounded by $1/\sigma$ and $\tau/(\sigma r)$, respectively. Similarly for the operator $(\omega\cdot\nabla_1)^{-1}$ itself. By Proposition 2.2.c we also get the following: Let $h\in\mathcal{A}'(r)$ and define $\widehat{h}=\{h,.\}$. Then $\mathbb{I}^-\widehat{h^+}(\omega\cdot\nabla_1)^{-1}\mathbb{I}^-$ is a continuous linear operator on $\mathcal{A}(r)$, and its operator norm is bounded by $\|h\|_r'/b$. Thus, for every $h\in B$, the following defines a bounded linear operator $\varphi(h)$ from $\mathcal{A}(r)$ to $\mathbb{I}^-\mathcal{A}'(r)$:

$$\varphi(h) = -(\omega\cdot\nabla_1)^{-1}\Bigl[\mathbb{I}-\mathbb{I}^-\widehat{h^+}(\omega\cdot\nabla_1)^{-1}\Bigr]^{-1}\mathbb{I}^-. \tag{2.8}$$

Clearly, $\phi=\varphi(h)f$ satisfies the equation (2.5) and the bounds (2.6). The asserted differentiability of the map $h\mapsto\varphi(h)$ follows from the fact that the inverse of the operator $[\ldots]$ in (2.8) is given by a convergent Neumann series in $\mathcal{L}\bigl(\mathbb{I}^-\mathcal{A}(r),\mathbb{I}^-\mathcal{A}(r)\bigr)$.

■

In other words, if $\phi=\varphi(f^+)f^-$, then the term $[\ldots]$ in equation (2.4) vanishes. Notice that by construction, the generating function ϕ has no slow modes.

We continue by discussing the solution of (2.1).

Proposition 2.5. *Let $r>0$ and $0<\delta<r/4$. Denote by B' the open ball in $\mathcal{A}'(r)$ of radius $\delta/(4d)$, and by B the open ball in $\mathcal{A}^d(r-\delta)$ of radius $\delta/(2d)$, both centered at the origin. Then for any given $\phi\in B'$, the equation (2.1) determines a unique pair of functions $Q,P\in B$, and these functions satisfy the bounds*

$$\|Q\|_{r-\delta}\le\|\phi\|_r'\,,\qquad \|Q+\nabla_2\phi\|_{r-\delta}\le\frac{2d}{\delta}\bigl(\|\phi\|_r'\bigr)^2\,, \tag{2.9}$$

$$\|P\|_{r-\delta}\le\|\phi\|_r'\,,\qquad \|P-\nabla_1\phi\|_{r-\delta}\le\frac{2d}{\delta}\bigl(\|\phi\|_r'\bigr)^2\,. \tag{2.10}$$

Furthermore, $\mathcal{Q}:\phi\mapsto Q$ and $\mathcal{P}:\phi\mapsto P$ are differentiable maps from B' to $\mathcal{A}^d(r-\delta)$.

Proof. To simplify notation, we will write $f(.+Q,.)$ for the composition of a function f with the map $(q,p)\mapsto\bigl(q+Q(q,p),p\bigr)$.

¿From Proposition 2.2.d we see that if ϕ is any function in B', then the equation $\mathcal{F}_\phi(Q) = -(\nabla_2\phi)(.+Q,.)$ defines a map $\mathcal{F}_\phi$ from B to some closed subset of B. The corresponding map $(\phi,Q)\mapsto\mathcal{F}_\phi(Q)$ is clearly of class C^1 on $B'\times B$. By using parts (b,d) of Proposition 2.2, with $\delta/2$ instead of δ, we obtain the following bound on the derivative of $\mathcal{F}_\phi$ at $Q\in B$, for any given $\phi\in B'$:

$$\begin{aligned}\|D\mathcal{F}_\phi(Q)h\|_{r-\delta} &= \|(h\cdot\nabla_1\nabla_2\phi)(.+Q,.)\|_{r-\delta}\\ &\leq \frac{2d}{\delta}\|\nabla_2\phi\|_r\|h\|_{r-\delta}\leq\frac{1}{2}\|h\|_{r-\delta}\,.\end{aligned}\tag{2.11}$$

Here, h is an arbitrary vector in $\mathcal{A}^d(r-\delta)$. Thus, by the contraction mapping principle, $\mathcal{F}_\phi$ has a unique fixed point $Q=\mathcal{Q}(\phi)$ in B, and by the chain rule, the map $\phi\mapsto\mathcal{Q}(\phi)$ is of class C^1 on B'. The first bound in (2.9) follows from Proposition 2.2.d. The second bound is obtained by applying the mean value theorem to the difference $\mathcal{F}_\phi\big(\mathcal{Q}(\phi)\big)-\mathcal{F}_\phi(0)$, using the bound (2.11) on the derivative of $\mathcal{F}_\phi$. Finally, define $\mathcal{P}(\phi)=\nabla_1\phi(.+\mathcal{Q}(\phi),.)$. By the chain rule, $\mathcal{P}$ is of class C^1 on B'. The proof of (2.10) is similar to that of (2.9).

■

Unless stated otherwise, we assume from now on that $Q=\mathcal{Q}(\phi)$, $P=\mathcal{P}(\phi)$, and $\phi=\varphi(f^+)f^-$. The results proved so far show, among other things, that under the appropriate regularity and smallness assumptions on f, the equation (2.4) reduces to

$$H\circ U_\phi = H^0+f^++M_1(f)+M_2(f)\,,\tag{2.12}$$

where

$$M_1(f)=\mathbb{I}^+\{f^+,\phi\},\qquad M_2(f)=\sum_{i=1}^4 R_i\,,\tag{2.13}$$

and where R_1,R_2,R_3, and R_4 are the functions defined in (2.3).

We continue by estimating the "higher order" term $M_2(f)$ in equation (2.12), by using the following fact: Let X and Y be Banach spaces over the complex numbers, let $h\in X$, and let R be a differentiable map from $D=\{x\in X\colon\ \|x\|<r\}$ to Y, for some $r>\|h\|$. If R is bounded on D and if $\|R(zh)\|=\mathcal{O}(|z|^n)$ as $z\to 0$, then

$$\|R(h)\|\leq\big(\|h\|/r\big)^n\sup_{x\in D}\|R(x)\|\,.\tag{2.14}$$

This bound is a straightforward consequence of Theorem 2.1, using the Hahn–Banach theorem and the maximum principle for analytic functions; see also the proof of Proposition 2.7.

Proposition 2.6. *There is a constant $c>0$ such that the following holds. Let $0<\delta<r/4$, $\delta<1$, and define $b=\sigma/(1+\tau/r)$. If f is a function in $\mathcal{A}'(r)$ such that $\|f^+\|'_r<b/2$ and $\|f^-\|_r<b\delta/(8d)$, then $M_2(f)\in\mathcal{A}(r-\delta)$ and*

$$\|M_2(f)\|_{r-\delta}\leq c(b\delta)^{-2}\|f^-\|_r^2\,.\tag{2.15}$$

Proof. For the function R_1 we get a bound of the form (2.15) by using the second inequality in (2.9) and (2.10). In the case of R_2 we first replace $h = (Q, P)$ by an arbitrary function $x \in \mathcal{A}^{2d}(r-\delta)$ of norm $\delta/2$ or less, and then we use Proposition 2.2.d. The resulting bound on $\|R_2\|_{r-\delta}$ can now be improved by a factor of $2\|\phi\|'_r/\delta$, using Proposition 2.5 and (2.14) with $n = 1$, since R_2 vanishes if $h = 0$. The function R_3 is estimated analogously, using that $R_3 = 0$ if $h = (Q + \nabla_2\phi, P - \nabla_1\phi)$ is zero. And similarly with R_4 and $h = (-\nabla_2\phi, \nabla_1\phi)$. Here, the bound (2.14) can be applied with $n = 2$, since both the map $h \mapsto R_4$ and its first derivative vanish at the origin.

■

In order to formulate the necessary estimates on the derivatives of the two maps M_1 and M_2, we need to introduce some notation: For any pair of nonnegative integers (j, n), define a polynomial $\Upsilon_{j,n}$ of three real variables by setting

$$\Upsilon_{j,n}(u, v, w) = \Bigl(w + v\frac{d}{du}\Bigr)^n u^j\,. \tag{2.16}$$

The polynomial $\Upsilon_{j,n}$ has the property that $\Upsilon_{j,n}(su, stv, tw) = s^j t^n \Upsilon_{j,n}(u, v, w)$.

In the following Proposition, if r and δ are given positive real numbers, then

$$B(\varepsilon) = \bigl\{f \in \mathcal{A}'(r) \colon \|f^+\|'_r + \delta^{-1}\|f^-\|'_r < \varepsilon\bigr\}\,, \qquad \varepsilon > 0. \tag{2.17}$$

Proposition 2.7. *There are constants $a, c_0, c_1, c_2 > 0$ such that the following holds, whenever $\delta > 0$ is sufficiently small and $\varrho' < r < \varrho$. The equation (2.13) defines two differentiable maps M_1 and M_2 from $B(3a)$ to $\mathcal{A}'(r - 2\delta)$, and for every $f \in B(3a)$,*

$$\|M_j(f)\|'_{r-2\delta} \le c_0\delta^{-j-1}\bigl(\|f^-\|'_r\bigr)^j\,, \qquad j = 1, 2. \tag{2.18}$$

For $j = 1, 2$ and $f \in B(2a)$, the derivative of M_j at f satisfies the bound

$$\|DM_j(f)h\|'_{r-2\delta} \le c_1\delta^{-j-2}\Upsilon_{j,1}\bigl(\|f^-\|'_r, \|h^-\|'_r, \|h\|'_r\bigr)\,, \tag{2.19}$$

where h is an arbitrary function in $\mathcal{A}'(r)$. Furthermore, if $f, h \in B(a)$ then

$$\begin{aligned}\|M_j(f + h) - M_j(f) &- DM_j(f)h\|'_{r-2\delta}\\ &\le c_2\delta^{-j-3}\Upsilon_{j,2}\bigl(\|f^-\|'_r, \|h^-\|'_r, \|h\|'_r\bigr)\,.\end{aligned} \tag{2.20}$$

Proof. The estimate (2.18) is obtained from Lemma 2.4 and Proposition 2.6, together with Proposition 2.2.b, which is used to bound the inclusion map from $\mathcal{A}(r - \delta)$ to $\mathcal{A}'(r - 2\delta)$. The differentiability of M_1 and M_2 follows from the differentiability of the maps $\mathcal{Q}$, $\mathcal{P}$, and φ, using the chain rule.

In order to estimate derivatives of the maps M_j, let us introduce a new norm $\|.\|$ on $\mathcal{A}'(r)$, defined by the equation $\|f\| = \|f^+\|'_r + \delta^{-1}\|f^-\|'_r$. Given any $f \in B(2a)$, any

$h \in \mathcal{A}'(r)$ such that $\|h\| = 1$, and any bounded linear functional ℓ: $\mathcal{A}'(r-2\delta) \to \mathbb{C}$ of norm one, we can use Cauchy's formula and (2.18) to obtain the following estimate:

$$\begin{aligned}
\big|\,\ell(D^n M_j(f)h^n)\,\big| &= \left|\,\frac{n!}{2\pi i}\oint\limits_{|z|=a} \ell\big(M_j(f+zh)\big)\frac{dz}{z^{n+1}}\,\right| \\
&\le n!a^{-n}\sup_{|z|=a}\|M_j(f+zh)\|'_{r-2\delta} \\
&\le n!a^{-n}c_0\delta^{-1}\big(\|f^-\|+a\|h^-\|\big)^j ,
\end{aligned} \tag{2.21}$$

where $j,n \in \{1,2\}$. By choosing ℓ such that $\ell(D^n M_j(f)h^n) = \|D^n M_j(f)h^n\|'_{r-2\delta}$, and using the fact that the map $h^n \mapsto D^n M_j(f)h^n$ is n–linear, this yields the bound

$$\|D^n M_j(f)h^n\|'_{r-2\delta} \le c_n\delta^{-1}\Upsilon_{j,n}\big(\|f^-\|,\|h^-\|,\|h\|\big)\,, \tag{2.22}$$

where $h \in \mathcal{A}'(r)$ can be arbitrary. In the case $n=1$, this bound implies (2.19).

In order to prove (2.20), let $f,h \in B(a)$, define $g_t = M_j(f+th) - M_j(f) - tDM_j(f)h$, and choose ℓ as in (2.21), with the additional property that $\ell(g_1) = \|g_1\|'_{r-2\delta}$. By applying a standard Taylor remainder formula to the function $t \mapsto \ell(g_t)$, we obtain the identity

$$\|M_j(f+h) - M_j(f) - DM_j(f)h\|'_{r-2\delta} = \int_0^1 \ell\big(D^2M_j(f+th)h^2\big)dt\,. \tag{2.23}$$

The desired bound is now obtained by using (2.22) with $n=2$.

■

We shall now prepare for the definition of an iterative procedure that will be used to construct a Hamiltonian $\widetilde{H} = H \circ \mathcal{U}_H$ that satisfies $\mathbb{I}^-\widetilde{H} = 0$, if $f = H - H^0$ is sufficiently small. To give a rough idea of what follows: For every $m \ge 0$, denote by H_m the m–th iterate of H under the transformation $H \mapsto H \circ U_\phi$ described above, and write $H_m = H^0 + f_m$. Then the sequence of functions $\widetilde{F}_m = f_m^- + f_m^+ - f_{m-1}^+$, defined for $m \ge 1$, is expected to tend to zero (almost quadratically) as $m \to \infty$. In order to show that $\widetilde{H}$ depends differentiably on H, it is convenient to characterize the sequence $\widetilde{F} = (0, \widetilde{F}_1, \widetilde{F}_2, \ldots)$ as the fixed point of a certain contraction on a space of sequences. The map $H \mapsto \widetilde{H}$ can then be written as the composition of $H \mapsto \widetilde{F}$ with the map $\widetilde{F} \mapsto \widetilde{H}$, given by the equation

$$\widetilde{H} = H^0 + f^+ + \sum_{m=1}^{\infty}\widetilde{F}_m^+\,. \tag{2.24}$$

We start by choosing a number r_∞ in the interval (ϱ', ϱ), and a positive $\delta_0 < (\varrho - r_\infty)/6$, such that the hypothesis of Proposition 2.7 is satisfied for $\delta \le \delta_0$. Additional smallness conditions on δ_0 will be specified later. The values of r and δ that will be used to estimate functions like $\widetilde{F}_m$ are chosen as follows:

$$r_m = r_\infty + 6\delta_m\,, \qquad \delta_m = (2/3)^m\delta_0\,, \qquad m = 0,1,2,\ldots\,, \tag{2.25}$$

so that $r_{m+1} = r_m - 2\delta_m$, for all $m \geq 0$.

Definition 2.8. *On each of the spaces $\mathcal{A}'(r_m)$, for $m \geq 0$, consider the two norms defined by the equation*

$$\|f\|_{k,m} = e^{\frac{k}{3}\delta_m^{-1}} \|f^+\|'_{r_m} + e^{\delta_m^{-1}} \|f^-\|'_{r_m} , \qquad k = 0, 1. \tag{2.26}$$

The vector space $\mathcal{A}'(r_m)$, when equipped with the norm $\|.\|_{k,m}$, will be referred to as $\mathcal{A}_{k,m}$. The open ball in $\mathcal{A}_{k,m}$ centered at the origin, and with radius $a > 0$, is denoted by $B_{k,m}(a)$.

Definition 2.9. *Given $k \in \{0, 1\}$, define $\mathcal{A}_k$ to be the vector space of all sequences $F = (F_0, F_1, F_2, \ldots)$ of functions $F_m \in \mathcal{A}_{k,m}$ such that $\|F_m\|_{k,m} \to 0$ as $m \to \infty$. We equip $\mathcal{A}_k$ with the norm*

$$\|F\|_k = \sup_{m \geq 0} \|F_m\|_{k,m} , \tag{2.27}$$

which makes $\mathcal{A}_k$ a Banach space. The open ball in $\mathcal{A}_k$ centered at the origin, and with radius $a > 0$, is denoted by $B_k(a)$. Given $F \in \mathcal{A}_k$, we define F^- to be the sequence $(F_0^-, F_1^-, \ldots)$.

The following four maps χ, $\mathcal{V}$, $\mathcal{M}_1$, and $\mathcal{M}_2$ will be used to define a contraction on $\mathcal{A}_1$ whose fixed point is the desired sequence $\widetilde{F}$. Define χ: $\mathcal{A}_{1,0} \to \mathcal{A}_1$ by setting

$$\big(\chi(f)\big)_0 = f , \qquad \big(\chi(f)\big)_m = 0 , \qquad m = 1, 2, \ldots , \tag{2.28}$$

for every function f in $\mathcal{A}_{1,0}$, and define $\mathcal{V}$: $\mathcal{A}_1 \to \mathcal{A}_0$ by the equation

$$\big(\mathcal{V}(F)\big)_0 = 0 , \qquad \big(\mathcal{V}(F)\big)_m = F_m^- + \sum_{k=0}^{m} F_k^+ , \qquad m = 1, 2, \ldots , \tag{2.29}$$

where F is an arbitrary sequence in $\mathcal{A}_1$. Clearly, χ and $\mathcal{V}$ are linear and bounded. χ is of norm one, and if δ_0 is sufficiently small (which we are assuming) then $\mathcal{V}$ is of norm less than 2. Finally, using the same constant $a > 0$ as in Proposition 2.7, we can define, for every $m \geq 0$, $j \in \{1, 2\}$, and $F \in B_0(a)$, a function $\big(\mathcal{M}_j(F)\big)_m$ by setting

$$\big(\mathcal{M}_j(F)\big)_0 = 0 , \qquad \big(\mathcal{M}_j(F)\big)_m = M_j(F_{m-1}) , \qquad m = 1, 2, \ldots , \tag{2.30}$$

Proposition 2.10. *There are constants $a, c > 0$, with c independent of the choice of δ_0 , such that the following holds: The equation (2.30) defines two differentiable maps $\mathcal{M}_1$ and $\mathcal{M}_2$ from $B_0(a)$ to $\mathcal{A}_1$, and these maps satisfy the bounds*

$$\|\mathcal{M}_j(F)\|_1 \leq ce^{-\frac{1}{4}\delta_0^{-1}} \|F\|_0^j , \qquad j \in \{1, 2\} , \tag{2.31}$$

and

$$\|D\mathcal{M}_j(F)G\|_1 \leq ce^{-\frac{1}{4}\delta_0^{-1}} \Upsilon_{j,1}\big(\|F^-\|_0, \|G^-\|_0, \|G\|_0\big) , \qquad j \in \{1, 2\} , \tag{2.32}$$

for every $F \in B_0(a)$ *and* $G \in \mathcal{A}_0$ *.*

Proof. Let $j \in \{1,2\}$ be fixed. By Proposition 2.7, if $a > 0$ is sufficiently small then for every $m > 0$ and every $n \in \{0,1,2\}$, we can define a map $\mathcal{M}_{j,n,m}$ from $B_0(a) \times B_0(a)$ to $\mathcal{A}_{1,m}$, by setting

$$\begin{aligned}
&\mathcal{M}_{j,0,m}(F,G) = M_j(F_{m-1})\,, \\
&\mathcal{M}_{j,1,m}(F,G) = DM_j(F_{m-1})G_{m-1}\,, \\
&\mathcal{M}_{j,2,m}(F,G) = M_j(F_{m-1}+G_{m-1}) - M_j(F_{m-1}) - DM_j(F_{m-1})G_{m-1}\,.
\end{aligned} \tag{2.33}$$

Here, and in what follows, we assume $F, G \in B_0(a)$. The estimates given in Proposition 2.7 imply that

$$\|\mathcal{M}_{j,n,m}(F,G)\|'_{r_m} \le c_n \delta^{-j-n-1} \Upsilon_{j,n}\bigl(\|F^-_{m-1}\|'_{r_{m-1}}, \|G^-_{m-1}\|'_{r_{m-1}}, \|G_{m-1}\|'_{r_{m-1}}\bigr)\,, \tag{2.34}$$

for the values of j, n considered here. We will now rewrite this bound in terms of the norms introduced in in Definition 2.8 and Definition 2.9, by using that the polynomials $\Upsilon_{j,n}$ have nonnegative coefficients and satisfy $\Upsilon_{j,n}(su, sv, w) = s^j \Upsilon_{j,n}(u,v,w)$. Since $\mathbb{I}^- M_1 = 0$, we obtain for $j = 1$

$$\|\mathcal{M}_{1,n,m}(F,G)\|_{1,m} \le c_n \delta_{m-1}^{-n-2} e^{\frac{1}{3}\delta_m^{-1} - \delta_{m-1}^{-1}} \Upsilon_{1,n}\bigl(\|F^-\|_0, \|G^-\|_0, \|G\|_0\bigr)\,, \tag{2.35}$$

and in the case $j = 2$, the bound (2.34) implies that

$$\|\mathcal{M}_{2,n,m}(F,G)\|_{1,m} \le c_n \delta_{m-1}^{-n-3} e^{\delta_m^{-1} - 2\delta_{m-1}^{-1}} \Upsilon_{2,n}\bigl(\|F^-\|_0, \|G^-\|_0, \|G\|_0\bigr)\,. \tag{2.36}$$

In both of these estimates, the difference in the exponential is equal to $-\frac{1}{2}\delta_{m-1}^{-1}$. Thus, we can define six maps $\mathcal{M}_{j,n}$ from $B_0(a) \times B_0(a)$ to $\mathcal{A}_1$, by setting $\bigl(\mathcal{M}_{j,n}(F,G)\bigr)_m$ equal to $\mathcal{M}_{j,n,m}(F,G)$ if $m > 0$, and equal to zero if $m = 0$. Furthermore, we have the bounds

$$\|\mathcal{M}_{j,n}(F,G)\|_1 \le c_n \delta_0^{-n-3} e^{-\frac{1}{2}\delta_0^{-1}} \Upsilon_{j,n}\bigl(\|F^-\|_0, \|G^-\|_0, \|G\|_0\bigr)\,, \tag{2.37}$$

for $j \in \{1,2\}$ and $n \in \{0,1,2\}$. This shows that $\mathcal{M}_j$ is differentiable as a map from $B_0(a)$ to $\mathcal{A}_1$, and that $D\mathcal{M}_j(F) = \mathcal{M}_{j,1}(F,.)$ for all $F \in B_0(a)$. The bounds (2.31) and (2.32) follow from (2.37) with $n = 1$.

■

Define $\mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2$, and assume that δ_0 has been chosen sufficiently small such that $c \exp -\frac{1}{4}\delta_0^{-1} < \frac{1}{4}$, where c is the second constant mentioned in Proposition 2.10.

Proposition 2.11. *If* $s > 0$ *is sufficiently small then for every* $f \in B_{1,0}(s)$*, the equation* $\mathcal{K}_f(F) = F$*, where*

$$\mathcal{K}_f(F) = \mathcal{M}\bigl(\mathcal{V}(\chi(f) + F)\bigr)\,, \qquad \mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2\,, \tag{2.38}$$

has a unique solution $F = \widetilde{F}$ in $B_1(s)$, and the map $\mathcal{F}: f \mapsto \widetilde{F}$ is differentiable from $B_{1,0}(s)$ to $B_1(s)$.

Proof. Let $s < \max(a,1)/4$, where a is the first constant mentioned in Proposition 2.10. Then, by Proposition 2.10 and the chain rule, $(f,F) \mapsto \mathcal{K}_f(F)$ is a differentiable map from $B_{1,0}(s) \times B_1(s)$ to $B_1(s/2)$, and the derivative $D\mathcal{K}_f(F)$ is bounded in norm by $1/2$, for all $f \in B_{1,0}(s)$ and $F \in B_1(s)$. Thus, $\mathcal{K}_f$ has a unique fixed point $\tilde{F} = \mathcal{F}(f)$, and by the implicit function theorem, $\mathcal{F}$ is differentiable.

■

Now we are ready to prove Theorem 1.3. In order to avoid confusion later on (when $\rho' < \varrho'$ and $\rho > \varrho$), let us restate the claim:

Theorem 2.12. *There exists an open neighborhood B of H^0 in $\mathcal{A}(\varrho)$, and for every $H \in B$ a canonical transformation $\mathcal{U}_H$ from $\mathcal{D}(\varrho')$ to $\mathcal{D}(\varrho)$, such that $\mathbb{I}^-(H \circ \mathcal{U}_H) = 0$, and such that $\mathcal{U}_H$ is the identity whenever $\mathbb{I}^- H = 0$. The map $H \mapsto H \circ \mathcal{U}_H$ is analytic from B to $\mathbb{I}^+\mathcal{A}(\varrho')$, and its derivative at H^0 is $\mathbb{I}^+$.*

Proof. Given $0 < \varrho' < \varrho$, choose the sequences $\{r_m\}$ and $\{\delta_m\}$ as specified in (2.25). Let $s > 0$ be sufficiently small for the conclusions of Proposition 2.11 to hold. Since $r_0 < \varrho$, the set $B_{1,0}(s)$ contains some open neighborhood of zero in the space $\mathcal{A}(\varrho)$. Thus, we can assume now that $H = H^0 + f$ with $f \in B_{1,0}(s)$.

Let $\widetilde{F} = \mathcal{F}(f)$ be the function described in Proposition 2.11, and consider the sequence $V(f)$ defined by the equation

$$\big(V(f)\big)_m = \big(\mathcal{V}(\chi(f) + \widetilde{F})\big)_m\,, \qquad m = 0, 1, 2, \ldots . \tag{2.39}$$

The functions $f_m = (V(f))_m$ satisfy the recursion relation

$$\begin{aligned} f_m &= f_{m-1}^+ + \widetilde{F}_m = f_{m-1}^+ + \big(\mathcal{M}(\mathcal{V}(\chi(f) + \widetilde{F}))\big)_m \\ &= f_{m-1}^+ + M_1(f_{m-1}) + M_2(f_{m-1})\,, \end{aligned} \tag{2.40}$$

for all $m > 0$. This shows that the Hamiltonian $H_m = H^0 + f_m$ is the m–th iterate, under the map $H \mapsto H \circ U_\phi$ defined by equation (2.4), of the Hamiltonian $H = H^0 + f$. In particular, if we set $\phi_m = \varphi(f_{m-1}^+) f_{m-1}^-$ and

$$U'_m = U_{\phi_1} \circ U_{\phi_2} \circ \ldots \circ U_{\phi_m}\,, \qquad m = 1, 2, \ldots\,, \tag{2.41}$$

then $H_m = H \circ U'_m$ for all $m > 0$. If $f^- = 0$ then we see from (2.40) and (2.18) that $f_m^- = 0$, for all $m > 0$. In this case, $\phi_1 = \phi_2 = ... = 0$, so that U'_m is the identity map for all m. Similarly, if $f_1^- = 0$ then $f_m = f_1$ for all $m > 1$. This fact will be used below.

Before discussing the limit as $m \to \infty$, we note that $V(f) \in B_0(3s)$, and in particular,

$$\|f_m^-\|_{r_m} < 3s e^{-\delta_m^{-1}}\,, \qquad m = 0, 1, 2, \ldots . \tag{2.42}$$

This follows from the properties of the maps χ and $\mathcal{V}$, mentioned before Proposition 2.10. For later reference, we also note that by Proposition 2.11, the map $f \mapsto V(f)$ is differentiable on $B_{1,0}(s)$.

By Lemma 2.4 and Proposition 2.5, the transformation $U_{\phi_{m+1}} - U_0$ is an element of $\mathcal{A}^{2d}(r_m - \delta_m)$ whose norm is bounded by a constant times $\|f_m\|_{r_m}$. Thus, there are constants $c, c' > 0$ such that

$$\begin{aligned} \|U'_{m+1} - U'_m\|_{r_\infty} &= \|U'_m \circ U_{\phi_{m+1}} - U'_m \circ U_0\|_{r_\infty} \\ &\leq c\delta_m^{-1} \|U'_m\|_{r_m} \|f_m^-\|_{r_m} \leq c'\delta_m^{-1} e^{-\delta_m^{-1}} , \end{aligned} \tag{2.43}$$

for all $m \geq 0$. Here, we have used the mean value theorem, Proposition 2.2.b, and the bound (2.42). This last estimate shows that the sequence of transformations U'_m converges in $\mathcal{A}(r_\infty)$ to a limit

$$\mathcal{U}_H = U'_\infty = \lim_{m \to \infty} U'_m , \tag{2.44}$$

and from Proposition 2.2.a we see that $H_n \to H \circ \mathcal{U}_H$ pointwise on $\mathcal{D}(r_\infty)$.

Next, we note that the sequence $(f_0, f_1, \ldots)$ converges in $\mathcal{A}'(r_\infty)$ to the function

$$\tilde{f} = f^+ + \Sigma^+ \mathcal{F}(f) , \tag{2.45}$$

where

$$\Sigma^+ F = \sum_{m=1}^{\infty} F_m^+ , \qquad F \in \mathcal{A}_1 . \tag{2.46}$$

The operator Σ^+ given by equation (2.46) is clearly continuous from $\mathcal{A}_1$ to $\mathbb{I}^+\mathcal{A}'(r_\infty)$. Thus, by Proposition 2.11, the equation (2.45) defines a differentiable map $f \mapsto \tilde{f}$ from $B_{1,0}(s)$ to $\mathbb{I}^+\mathcal{A}'(r_\infty)$. The differentiability of the map $H \mapsto H \circ \mathcal{U}_H$ now follows from the identity

$$H \circ \mathcal{U}_H = H^0 + \tilde{f} , \tag{2.47}$$

and from the fact that the inclusion maps $\mathcal{A}(\varrho) \to \mathcal{A}'(r_0)$ and $\mathcal{A}'(r_\infty) \to \mathcal{A}(\varrho')$ are continuous.

In order to see why $\|\tilde{f} - f^+\|_{\varrho'} = \mathcal{O}(\|f\|_\varrho)^2$, as claimed in Theorem 2.12, we can write the map $f \mapsto \tilde{f}$ as the composition of the first step $f \mapsto f_1$ with the map $f_1 \mapsto \tilde{f}$, where $f_1 \in B_{0,1}(3s)$. Then we can use that $\|\tilde{f} - f_1\|_{\varrho'} = \mathcal{O}(\|f_1^-\|_{0,1})$ and $\|f_1 - f^+\|_{0,1} = \mathcal{O}(\|f\|_\varrho)^2$. The first bound follows from the fact that $f_1^- = 0$ implies $\tilde{f} = f_1$, and the second bound is obtained from the identity $f_1 - f^+ = \mathbb{I}^+\{f^+, \varphi(f^+) f^-\} + M_2(f)$, together with Proposition 2.6.

■

3. Properties of the Transformations $\mathcal{U}_H$

One of the goals in this section is to describe the connection between $\mathcal{U}_H$ and $\mathcal{U}_{H'}$, if H and H' are related by a translation (in the torus variable q) or scaling. But first, we prove the following theorem.

Theorem 3.1. *Let $0 < \varrho' < \varrho$. Then there exists an open neighborhood B of $H^{(0)}$ in $\mathcal{A}(\varrho)$, a number $\varrho'' \in (\varrho', \varrho)$, and a differentiable map ψ from B to $\mathcal{A}(\varrho'')$, such that $U_{\psi(H)} = \mathcal{U}_H$ for every $H \in B$, where $\mathcal{U}_H : \mathcal{D}(\varrho') \to \mathcal{D}(\varrho)$ is the canonical transformation described in Theorem 2.12. Furthermore, $\psi(H) = 0$ whenever $\mathbb{I}^- H = 0$.*

Here, and in what follows, σ, τ, and $\varrho' < \varrho$ are assumed to be the same (arbitrary) positive real numbers that were used in the last section. Similarly with $\omega \in \mathbb{R}^d$.

In order to prove Theorem 3.1, it turns out to be useful to determine at once all of the generating functions $\Psi_m = \psi(H_m)$, for $m = 0, 1, 2, \ldots$. This will be done by solving an implicit equation for the sequence of differences $\tilde{G}_m = \Psi_m - \Psi_{m+1} - \phi_{m+1}$, where $\phi_1, \phi_2, \ldots$ are the functions defined in the proof of Theorem 2.12. The resulting expression for the generating function of $\mathcal{U}_H$ is

$$\psi(H) = \Psi_0 = \tilde{G}_0 + \sum_{m=1}^{\infty} \bigl(\phi_m + \tilde{G}_m\bigr) . \tag{3.1}$$

In order to estimate the generating function of a composite transformation $U_{\phi_0} \circ U_{\phi_1}$, we use the following

Proposition 3.2. *Let $r > 0$ and $0 < \delta < r/2$, and denote by B' the open ball in $\mathcal{A}'(r + 2\delta)$ of radius $\delta/(4d)$ centered at the origin. Then there exists a differentiable map N from $B' \times B'$ to $\mathcal{A}'(r)$, satisfying the bound*

$$\|N(\phi_0, \phi_1)\|'_r \leq \frac{6d}{\delta} \|\phi_0\|'_{r+2\delta} \|\phi_1\|'_{r+2\delta} , \qquad \phi_0 , \phi_1 \in B' , \tag{3.2}$$

such that the following holds. If ϕ_0 and ϕ_1 are any functions in B' then the restriction of $U_{\phi_0} \circ U_{\phi_1}$ to $\mathcal{D}(r - \delta)$ satisfies

$$U_{\phi_0} \circ U_{\phi_1} = U_\phi , \qquad \phi = \phi_0 + \phi_1 + N(\phi_0, \phi_1) . \tag{3.3}$$

Proof. A straightforward computation shows that if

$$U_{\phi_1}(q'', p) = (q', p') , \qquad U_{\phi_0}(q', p') = (q, p'') , \tag{3.4}$$

and $\phi(q, p) = \phi_0(q, p') + \phi_1(q', p) + (q - q') \cdot (p' - p)$, then $U_\phi(q'', p) = (q, p'')$. The ranges of all these variables are determined a posteriori, after estimating the function ϕ.

In order to give a proper definition of ϕ, we first express $Q = q - q'$ and $P = p' - p$ in terms of the variables (q, p), by solving the equations

$$\begin{aligned} Q(q, p) &= -\bigl(\nabla_2 \phi_0\bigr)\bigl(q, p + P(q, p)\bigr) , \\ P(q, p) &= \bigl(\nabla_1 \phi_1\bigr)\bigl(q - Q(q, p), p\bigr) . \end{aligned} \tag{3.5}$$

The solution is similar to that of equation (2.1), described in the proof of Proposition 2.5. It yields a differentiable map $(\phi_0, \phi_1) \mapsto (Q, P)$ from $B' \times B'$ to $B \times B$, where B is the open ball in $\mathcal{A}^d(r + \delta)$ of radius $\delta/(2d)$, centered at the origin. The norms $\|Q\|_{r+\delta}$ and $\|P\|_{r+\delta}$ are bounded by $\|\phi_0\|'_{r+2\delta}$ and $\|\phi_1\|'_{r+2\delta}$, respectively. Now we can write $\phi = \phi_0 + \phi_1 + N(\phi_0, \phi_1)$, where

$$\begin{aligned}\big(N(\phi_0, \phi_1)\big)(q, p) = &\big[\phi_0\big(q, p + P(q, p)\big) - \phi_0(q, p)\big] \\ &+ \big[\phi_1\big(q - Q(q, p), p\big) - \phi_1(q, p)\big] + Q(q, p) \cdot P(q, p)\,,\end{aligned} \tag{3.6}$$

for (q, p) in $\mathcal{D}(r)$. The bound (3.2), and the remaining assertions, now follow from Proposition 2.5 and Proposition 2.2.

■

Our next goal is to compose the canonical transformations associated with certain sequences Φ of generating functions. Given $0 < \varrho' < r_\infty < \varrho$, we choose a positive $\delta_0 < (r_\infty - \varrho')/7$ that is sufficiently small to meet the conditions mentioned earlier. An additional smallness condition will be specified below. Let

$$r'_m = r_\infty - 6\delta_m\,, \qquad \delta_m = (2/3)^m \delta_0\,, \qquad m = 0, 1, 2, \ldots\,, \tag{3.7}$$

so that $r'_{m+1} = r'_m + 2\delta_m$, for all $m \geq 0$.

Definition 3.3. *On each of the the spaces $\mathcal{A}'(r'_m)$, for $m \geq 0$, consider the norm*

$$\|\phi\|'_{1,m} = e^{\frac{1}{3}\delta_m^{-1}} \|\phi\|'_{r'_m}\,. \tag{3.8}$$

The vector space $\mathcal{A}'(r'_m)$, when equipped with this new norm, will be referred to as $\mathcal{A}'_{1,m}$. The open ball in $\mathcal{A}'_{1,m}$ centered at the origin, and with radius $a > 0$, will be denoted by $B'_{1,m}(a)$.

Definition 3.4. *Let $\mathcal{A}'_1$ be the vector space of all sequences $\Phi = (\Phi_0, \Phi_1, \Phi_2, \ldots)$ of functions $\Phi_m \in \mathcal{A}'_{1,m}$ such that $\|\Phi_m\|'_{1,m} \to 0$ as $m \to \infty$. We equip $\mathcal{A}'_1$ with the norm*

$$\|\Phi\|'_1 = \sup_{m \geq 0} \|\Phi_m\|_{1,m}\,, \tag{3.9}$$

which makes $\mathcal{A}'_1$ a Banach space. The open ball in $\mathcal{A}'_1$ centered at the origin, and with radius $a > 0$, will be denoted by $B'_1(a)$.

Corollary 3.5. *There are constants $a, c > 0$, such that the equation*

$$\big(\mathcal{N}(\Phi, \Psi)\big)_m = N(\Phi_{m+1}, \Psi_{m+1})\,, \qquad m = 0, 1, 2, \ldots \tag{3.10}$$

defines a map $\mathcal{N}$ from $B'_1(a) \times B'_1(a)$ to $\mathcal{A}'_1$. Furthermore, this map is differentiable and satisfies

$$\|\mathcal{N}(\Phi, \Psi)\|'_1 \leq c \|\Phi\|'_1 \|\Psi\|'_1\,, \qquad \Phi, \Psi \in B'_1(a)\,. \tag{3.11}$$

Proof. The bound (3.11) follows from Proposition 3.2, which implies that

$$\|N(\Phi_{m+1},\Psi_{m+1})\|'_{r'_m} \le \frac{6d}{\delta_m} e^{-\delta_m^{-1}} \|\Phi'\|'_1 \|\Phi\|'_1 , \tag{3.12}$$

for all $\Phi, \Psi \in B'_1(a)$, if $a > 0$ is sufficiently small.

Consider the the maps $\mathcal{N}_n$, for $n = 1, 2, \ldots$, where $(\mathcal{N}_n(\cdot))_m$ is either $(\mathcal{N}(\cdot))_m$ or zero, depending on whether $m \le n$ or $m > n$. These maps are clearly differentiable on $B'_1(a) \times B'_1(a)$. Furthermore, as (3.12) shows, we have $\mathcal{N}_n(\Phi,\Psi) \to \mathcal{N}(\Phi,\Psi)$ as $n \to \infty$, uniformly on $B'_1(a) \times B'_1(a)$. Thus, $\mathcal{N}$ is differentiable (complex analytic).

■

Consider now the linear operator $\mathcal{V}'$ on $\mathcal{A}'_1$, defined by the equation

$$\left(\mathcal{V}'G\right)_m = \sum_{k=m+1}^{\infty} G_k \,, \qquad m = 0, 1, \ldots . \tag{3.13}$$

Clearly, $\mathcal{V}'$ is continuous, and of norm less than one, if δ_0 has been chosen sufficiently small (which we assume).

Proposition 3.6. *If $s > 0$ is sufficiently small then for every $\Phi \in B'_1(s)$, the equation $\mathcal{K}'_\Phi(G) = G$, where*

$$\mathcal{K}'_\Phi(G) = \mathcal{N}\big(\Phi, G + \mathcal{V}'(\Phi + G)\big) , \tag{3.14}$$

has a unique solution $G = \tilde{G}$ in $B'_1(s)$, and the map $\Phi \mapsto \tilde{G}$ is differentiable on $B'_1(s)$.

Proof. Let a, c be the constants mentioned in Corollary 3.5. By using the bound (3.11) and Cauchy's formula (as e.g. in the proof of Proposition 2.7), the partial derivatives of $\mathcal{N}$ can be estimated as follows:

$$\|D_1\mathcal{N}(\Phi,\Psi)\| \le 2c\|\Psi\|'_1 , \qquad \|D_2\mathcal{N}(\Phi,\Psi)\| \le 2c\|\Phi\|'_1 , \tag{3.15}$$

for every $\Phi, \Psi \in B'_1(a/2)$. Thus, if $s > 0$ is sufficiently small, $\mathcal{K}'_\Phi$ is a contraction from $B'_1(s)$ to $B'_1(s/2)$, for every $\Phi \in B'_1(s)$. And in particular, $\mathcal{K}'_\Phi$ has a unique fixed point $\tilde{G}$ in $B'_1(s)$. By the implicit function theorem, this fixed point depends differentiably on the parameter Φ.

■

The sequence $\Phi = \hat{\varphi}_n(f)$ described in the proposition below will be used to construct the generating function for the canonical transformation U'_n considered in the proof of Theorem 2.12. The maps φ and V that enter the definition of $\hat{\varphi}_n(f)$ are defined in Lemma 2.4 and equation (2.39), respectively.

Proposition 3.7. *There exists a constant $c > 0$, such that the following holds if $s > 0$ is sufficiently small: For $1 \le n \le \infty$, the equation*

$$\left(\hat{\varphi}_n(f)\right)_m = \begin{cases} \varphi(f^+_{m-1}) f^-_{m-1} , & \textit{if } 0 \le m-1 < n; \\ 0, & \textit{otherwise;} \end{cases} \tag{3.16}$$

where $f_k = (V(f))_k$ for all $k \geq 0$, defines a differentiable map $\hat{\varphi}_n$ from $B_{1,0}(s)$ to $B'_1(cs)$. Furthermore, $\hat{\varphi}_n(f) \to \hat{\varphi}_\infty(f)$ as $n \to \infty$, uniformly on $B_{1,0}(s)$.

Proof. As was mentioned in the proof of Theorem 2.12, the map V defined by (2.39) is differentiable from $B_{1,0}(s)$ to $B_0(3s)$; and by Lemma 2.4, there exists a constant $c > 0$ such that $g \mapsto \varphi(g^+)g^-$ defines a differentiable map from $B_{0,m-1}(3s)$ to $B'_{1,m}(c_m s)$, for every $m > 0$, where $c_m = c\exp -\frac{1}{3}\delta_m^{-1}$. Here, and in what follows, we need to assume that $s > 0$ is sufficiently small. By composing these maps according to equation (3.16), we obtain a differentiable map $\hat{\varphi}_n$ from $B_{1,0}(s)$ to $B'_1(cs)$, for every positive integer n. ¿From the fact that the sequence $\{c_m\}$ tends to zero, we see that $\hat{\varphi}_n(f) \to \hat{\varphi}_\infty(f)$ uniformly on $B_{1,0}(s)$. This in turn implies that $\hat{\varphi}_\infty(f)$ is differentiable (complex analytic).

■

Proof of Theorem 3.1. Given $0 < \varrho' < \varrho$, choose the sequences $\{r_m\}$, $\{r'_m\}$, and $\{\delta_m\}$ as specified in (2.25) and (3.7), and let $\varrho'' = r'_0$.

We define ψ as the composition of three maps $H \mapsto f \mapsto \Psi \mapsto \Psi_0$, where $f \in \mathcal{A}_{1,0}$ is given by the equation $f = H - H^0$, and where $\Psi_0 \in \mathcal{A}'_{1,0}$ is the first component of the sequence Ψ that will be defined below. Since $\varrho' < r'_0 < r_0 < \varrho$, the inclusion maps $\mathcal{A}(\varrho) \to \mathcal{A}_{1,0}$ and $\mathcal{A}'_{1,0} \to \mathcal{A}(\varrho')$ are continuous. The map $\Psi \mapsto \Psi_0$ from $\mathcal{A}'_1$ to $\mathcal{A}(\varrho'')$ is clearly continuous as well. Thus, it suffices to prove that the map $f \mapsto \Psi$ is differentiable on $B_{1,0}(s)$, for some $s > 0$, and that Ψ_0 is the generating function for $\mathcal{U}_H$.

The first part is an immediate consequence of Proposition 3.6 and Proposition 3.7. Namely, if $s > 0$ is sufficiently small, then the equation

$$\Psi = \tilde{G} + \mathcal{V}'(\Phi + \tilde{G}), \qquad \Phi = \hat{\varphi}_n(f), \tag{3.17}$$

where $\tilde{G}$ is the sequence described in Proposition 3.6, defines a differentiable map $f \mapsto \Psi$ from $B_{1,0}(s)$ to $\mathcal{B}'_1(3cs)$. Here, n can be an arbitrary positive integer or the symbol ∞, and $c > 0$ is some constant that is independent of n. The abovementioned definition $\psi(H) = \Psi_0$ refers to the case $n = \infty$.

Let now $f \in B_{1,0}(s)$ be fixed, with $s > 0$ sufficiently small for the above to hold. We also assume that $3cs < \delta_0/(4d)$, so that Proposition 2.5 can be used to define the canonical transformation U_{Ψ_0} on the domain $\mathcal{D}(\varrho'' - \delta_0)$. This domain contains $\mathcal{D}(\varrho')$, due to our choice of δ_0. By definition, the sequence Ψ satisfies the equation

$$\begin{aligned} \Psi_{m-1} &= \tilde{G}_{m-1} + \sum_{k=m}^{\infty} \bigl(\Phi_k + \tilde{G}_k\bigr) = \tilde{G}_{m-1} + \Phi_m + \Psi_m \\ &= N(\Phi_m, \Psi_m) + \Phi_m + \Psi_m \,, \end{aligned} \tag{3.18}$$

for all $m > 0$. In what follows, let $0 < n < \infty$. Then $\Phi_m = 0$ for all $m > n$. ¿From the definition (3.14) of $\mathcal{K}'_\Phi(G)$ one readily sees that the set of all sequences $G \in B'_1(s)$, that satisfy $G_m = 0$ whenever $m \geq n - 1$, is invariant under the map $\mathcal{K}'_\Phi$. Thus, the fixed point $\tilde{G}$ of $\mathcal{K}'_\Phi$ also belongs to this invariant set. As a consequence, we have $\Psi_m = 0$

for all $m \geq n$. The remaining components of Ψ are determined by the recursion relation (3.18), with "initial" condition $\Psi_n = 0$. By using Proposition 3.2, we can identify Ψ_0 as the generating function for the canonical transformation $U_n' = U_{\Phi_1} \circ U_{\Phi_2} \circ \ldots \circ U_{\Phi_n}$. As was shown in the proof of Theorem 2.12, $U_n' \to \mathcal{U}_H$ in $\mathcal{A}(\varrho')$, as $n \to \infty$. On the other hand, Proposition 3.7 and Proposition 3.6 imply that the sequence $n \mapsto \Psi_0$, as defined by equation (3.17), converges in $B_1'(3cs)$ to $\psi(H)$. Thus, since the map $\phi \mapsto U_\phi$ is continuous (by Proposition 2.5), we conclude that $U_{\psi(H)} = \mathcal{U}_H$.

Finally, if we assume that $\mathbb{I}^- f = 0$, then $\Phi = 0$, and thus $\tilde{G} = 0$ (since $\mathcal{K}_0'(0) = 0$), implying that $\psi(H) = \Psi_0 = 0$.

■

In the remaining part of this section, we consider the composition of certain changes of coordinates with the maps $\phi \mapsto U_\phi$ and $H \mapsto \mathcal{U}_H$. The results will be used in subsequent sections.

In order to estimate the composition of a function $H \in \mathcal{A}(r)$ with a linear change of coordinates, it is convenient to split H into homogeneous polynomials $H_\nu^{(n)}$ as follows:

$$H(q,p) = \sum_{\substack{\nu \in \mathbb{Z}^d \\ n \geq 0}} H_\nu^{(n)}(p) e^{i\nu \cdot q}, \qquad H_\nu^{(n)}(p) = \sum_{|\alpha| = n} H_{\nu,\alpha} p^\alpha, \tag{3.19}$$

where $H_{\nu,\alpha}$ are the Fourier–Taylor coefficients of H, as defined by the equation (1.4). In this representation, the norm of H can be written as

$$\|H\|_r = \sum_{\substack{\nu \in \mathbb{Z}^d \\ n \geq 0}} \|H_\nu^{(n)}\| r^n e^{r|\nu|}, \qquad \|H_\nu^{(n)}\| = \sum_{|\alpha| = n} |H_{\nu,\alpha}|. \tag{3.20}$$

If T is an arbitrary $d \times d$ matrix, and if x is a vector in $\mathbb{R}^d$, define

$$\|T\|_x = \max_{\substack{|y|=1 \\ x \cdot y = 0}} |Ty|. \tag{3.21}$$

Let now T be an integral $d \times d$ matrix with determinant ± 1, and let μ be a nonzero complex number. Then the equation (1.8) defines a linear map T_μ on $\mathbb{C}^d \times \mathbb{C}^d$. In the following proposition, we assume that $r, r' > 0$ satisfy

$$\|T^*\|_0 \leq \frac{r}{r'}, \quad \left\|\mu T^{-1}\right\|_0 \leq \frac{r}{r'}, \quad \|T\|_0 \leq \frac{r}{r'}, \quad \left\|\mu (T^*)^{-1}\right\|_0 \leq \frac{r}{r'}. \tag{3.22}$$

The first two of these inequalities guarantee that T_μ maps $\mathcal{D}(r')$ into $\mathcal{D}(r)$. The last two are added for convenience.

Proposition 3.8. *The linear operator $\phi \mapsto \phi \circ T_\mu$ is continuous from $\mathcal{A}(r)$ to $\mathcal{A}(r')$, and of norm 1. It is also continuous from $\mathcal{A}'(r)$ to $\mathcal{A}'(r')$, and in this case, its operator norm is $\leq r/r'$. Let now δ be a positive number less than $r/4$, and define $\delta' = \delta r'/r$. Assume* *that $\phi \in \mathcal{A}'(r)$ is of norm less than* $\min(|\mu|(r'/r)^2, 1)\delta/(4d)$. *Then, on* $\mathcal{D}(r'-\delta')$, *we have the identity*

$$T_\mu^{-1} \circ U_\phi \circ T_\mu = U_{\phi'} , \qquad \phi' = \mu^{-1}\phi \circ T_\mu . \tag{3.23}$$

Proof. ¿From the bound

$$\begin{aligned} \big\|\phi_\nu^{(n)} \circ \big[\mu(T^*)^{-1}\big]\big\|(r')^n e^{r'|T^*\nu|} &\leq \big\|\phi_\nu^{(n)}\big\| \, \big\|\mu T^{-1}\big\|_0^n (r')^n e^{r'\|T^*\|_0|\nu|} \\ &\leq \big\|\phi_\nu^{(n)}\big\| r^n e^{r|\nu|} , \end{aligned} \tag{3.24}$$

we see that $\phi \circ T_\mu \in \mathcal{A}(r')$ whenever $\phi \in \mathcal{A}(r)$, and that $\|\phi \circ T_\mu\|_{r'} \leq \|\phi\|_r$. Consider now $\phi \in \mathcal{A}'(r)$ of norm less than $\min(|\mu|(r'/r)^2, 1)\delta/(4d)$, and define $\phi' = \mu^{-1}\phi \circ T_\mu$. Then the bound (3.24), with ϕ replaced by $\nabla\phi$, shows that $\|\phi'\|'_{r'} \leq |\mu|^{-1}(r/r')\|\phi\|'_r$. And this in turn implies $\|\phi'\|_{r'} < \delta'/(4d)$. Now we can apply Proposition 2.5 to obtain two canonical transformations $U_\phi \colon \mathcal{D}(r-\delta) \to \mathcal{D}(r)$ and $U_{\phi'} \colon \mathcal{D}(r'-\delta') \to \mathcal{D}(r')$, whose generating functions are ϕ and ϕ', respectively. Notice that T_μ maps the domain $\mathcal{D}(r'-\delta')$ of $U_{\phi'}$ into the domain $\mathcal{D}(r-\delta)$ of U_ϕ.

In order to prove (3.23), let (Q, P) be the increments of the map U_ϕ. By this we mean that $U_\phi(q,p) = (q + Q(q,p), p + P(q,p))$, for all (q,p) in the domain of U_ϕ. Recall that U_ϕ is uniquely defined by the equation

$$Q = -(\nabla_2\phi) \circ V(Q) , \qquad P = (\nabla_1\phi) \circ V(Q) , \tag{3.25}$$

where $V(Q)$ denotes the map with increments $(Q, 0)$. The following computation shows that the increments (Q', P') of the map $T_\mu^{-1} \circ U_\phi \circ T_\mu$ agree with those of $U_{\phi'}$:

$$\begin{aligned} Q' &= T^{-1}Q \circ T_\mu = -T^{-1}(\nabla_2\phi) \circ V(Q) \circ T_\mu \\ &= -T^{-1}(\nabla_2\phi) \circ T_\mu \circ V(T^{-1}Q \circ T_\mu) = -T^{-1}(\nabla_2\phi) \circ T_\mu \circ V(Q') \\ &= -\big(\nabla_2(\mu^{-1}\phi \circ T_\mu)\big) \circ V(Q') = -(\nabla_2\phi') \circ V(Q') , \\ \\ P' &= \mu^{-1}T^*P \circ T_\mu = \mu^{-1}T^*(\nabla_1\phi) \circ V(Q) \circ T_\mu \\ &= \mu^{-1}T^*(\nabla_1\phi) \circ T_\mu \circ V(Q') \\ &= \big(\nabla_1(\mu^{-1}\phi \circ T_\mu)\big) \circ V(Q') = (\nabla_1\phi') \circ V(Q') . \end{aligned} \tag{3.26}$$

Thus, the proposition is proved.

■

Given any vector $u \in \mathbb{C}^d$, and any complex number z, define two linear transformations J_u and S_z on $\mathbb{C}^d$, by setting $J_u(q,p) = (q+u, p)$ and $S_z(q,p) = (q, zp)$, respectively. Assuming that $z \neq 0$, we can rewrite the definition (1.12) in the form $\mathcal{J}_u H = H \circ J_u^{-1}$ and $\mathcal{S}_z H = zH \circ S_z^{-1}$. Below it will be shown that, under suitable assumptions,

$$\mathcal{U}_{\mathcal{J}_u H'} = J_u \circ \mathcal{U}_{H'} \circ J_u^{-1} , \qquad \mathcal{U}_{\mathcal{S}_z H'} = S_z \circ \mathcal{U}_{H'} \circ S_z^{-1} . \tag{3.27}$$

But first, we need the following

Proposition 3.9. *Consider $u \in \mathbb{R}^d$, and $z \in \mathbb{C}$ satisfying $|z| = 1$. Under the same assumptions (and using the same notation) as in Proposition 2.5, the following holds for all $\phi \in B'$. If $\phi' = \mathcal{J}_u^{-1}\phi$ then $J_u^{-1} \circ U_\phi \circ J_u = U_{\phi'}$, and if $\phi' = \mathcal{S}_z^{-1}\phi$ then $S_z^{-1} \circ U_\phi \circ S_z = U_{\phi'}$.*

The proof of this proposition is straightforward and will be omitted. Notice that if $u \in \mathbb{R}^d$ and $|z| = 1$ then $\mathcal{J}_u$ and $\mathcal{S}_z$ are isometries on $\mathcal{A}(r)$, for any given $r > 0$.

For every $r, b > 0$, denote by $B(r, b)$ the open ball in $\mathcal{A}(r)$ of radius b centered at H^0.

Lemma 3.10. *Let $\delta > 0$ and $0 < \rho' + \delta \leq \varrho' < \varrho \leq \rho - \delta$. There exists $b > 0$, such that the following holds, for all $u \in \mathcal{D}_1(\delta)$, and for all $z \in \mathbb{C}$ satisfying $0 < |z| \leq 1$. If $H \in B(\rho, b)$ and $H' = \mathcal{J}_u^{-1}H$, then $J_u \circ \mathcal{U}_{H'} = \mathcal{U}_H \circ J_u$ on $\mathcal{D}(\rho')$. Furthermore, if $H \in B(\varrho, |z|b)$ and $H' = \mathcal{S}_z^{-1}H$, then $S_z \circ \mathcal{U}_{H'} = \mathcal{U}_H \circ S_z$ on $\mathcal{D}(\varrho')$.*

Proof. Choose $b > 0$ such that $H \mapsto \mathcal{U}_H - \mathrm{I}$ is well defined and analytic as a map from $\hat{B} = B(\varrho, 2b)$ to $\mathcal{A}^{2d}(\varrho')$. Such a choice is possible by Theorem 3.1 and Proposition 2.5.

Consider first the scalings S_z, in the case $|z| = 1$. Let $H_0 \in \hat{B}$ be fixed, but arbitrary. Define V_0 and V_0' to be the identity maps on $\mathbb{C}^{2d}$. We recall that $\mathcal{U}_{H_0}$ is a composition of canonical transformations $U_{\mathcal{F}(H_m)}$, where $\mathcal{F}(H^0 + f) = \varphi(f^+)f^-$, with φ defined by equation (2.8). To be more precise, define $V_{m+1} = V_m \circ U_{\mathcal{F}(H_m)}$ and $H_{m+1} = H_0 \circ V_{m+1}$, for $m = 0, 1, 2, \ldots$. Then, as was shown in the proof of Theorem 2.12, $V_m \to \mathcal{U}_{H_0}$ pointwise on $\mathcal{D}(\varrho')$. Similarly, if we set $H_0' = \mathcal{S}_z^{-1}H_0$ and define $V_{m+1}' = V_m' \circ U_{\mathcal{F}(H_m')}$ and $H_{m+1}' = H_0' \circ V_{m+1}'$, for $m \geq 0$, then $V_m' \to \mathcal{U}_{H_0'}$ pointwise on $\mathcal{D}(\varrho')$. Thus, in order to prove the assertion, it suffices to show that $S_z \circ V_k' = V_k \circ S_z$ and $H_k' = \mathcal{S}_z^{-1}H_k$, for all k. Assume now that these identities holds for $k = m \geq 0$, and consider $k = m + 1$. By using that $\mathcal{S}_z^{-1}$ commutes with $\mathbb{I}^{\pm}$ and ∇_1, and that $\nabla_2 \circ \mathcal{S}_z^{-1} = z\mathcal{S}_z^{-1} \circ \nabla_2$, one easily verifies that $\mathcal{F}(\mathcal{S}_z^{-1}H_k) = \mathcal{S}_z^{-1}\mathcal{F}(H_k)$. Now we can apply Proposition 3.9 to obtain

$$S_z \circ V_{m+1}' = S_z V_m' \circ U_{\mathcal{F}(H_m')} = V_m \circ U_{\mathcal{F}(H_m)} \circ S_z = V_{m+1} \circ S_z\,, \tag{3.28}$$

and

$$H_{m+1}' = z^{-1}H_0 \circ S_z \circ V_{m+1}' = z^{-1}H_0 \circ V_{m+1} \circ S_z = \mathcal{S}_z^{-1}H_{m+1}\,, \tag{3.29}$$

which completes the induction step. Thus, $S_z \circ \mathcal{U}_{H_0'} = \mathcal{U}_{H_0} \circ S_z$ on $\mathcal{D}(\varrho')$, whenever $|z| = 1$ and $H_0 \in \hat{B}$.

Our next goal is to extend this identity to z's in sets of the form $G_\varepsilon = \{z \in \mathbb{C}\colon \varepsilon < |z| < 1 + \varepsilon\}$. To this end, let (q, p) be a fixed but arbitrary point in $\mathcal{D}(\varrho')$, and choose $r < \varrho'$ such that (q, p) belongs to $\mathcal{D}(r)$. Next, choose $\varepsilon > 0$ sufficiently small such that for each $z \in G_\varepsilon$, S_z maps $\mathcal{D}(r)$ into $\mathcal{D}(\varrho')$, and $\mathcal{S}_z^{-1}$ maps $B(\varrho, \varepsilon b)$ into $\hat{B}$. Let $H \in B(\varrho, \varepsilon b)$, and define a function $g_H : G_\varepsilon \to \mathbb{C}^{2d}$ by setting

$$g_H(z) = \big(S_z \circ \mathcal{U}_{\mathcal{S}_z^{-1}H} - \mathcal{U}_H \circ S_z\big)(q, p)\,, \qquad z \in G_\varepsilon\,. \tag{3.30}$$

All of the maps used in this definition are analytic (on the appropriate domains, due to our choice of parameters), and thus g_H itself is analytic. But as was shown above, g_H

vanishes on the circle $|z| = 1$. Thus, g_H is identically zero on G_ε. Let now $z \in G_\varepsilon$ be fixed. Since $\mathcal{S}_z^{-1}$ maps $B(\varrho, |z|b) \subset \hat{B}$ into the ball $\hat{B}$, where $H \mapsto \mathcal{U}_H - \mathrm{I}$ is analytic, the function $H \mapsto g_H(z)$ defined by equation (3.30) is analytic on $B(\varrho, |z|b)$. But this function vanishes on $B(\varrho, \varepsilon b)$, and thus $g_H(z) = 0$, for all H in $B(\varrho, |z|b)$. This proves the assertion regarding scalings, since ε and (q,p) were arbitrary.

The translations J_u can be handled similarly. Assume first that $u \in \mathbb{R}^d$. Then we can follow the same steps as in the case of $|z| = 1$ scalings, replacing S_z and $\mathcal{S}_z$ by J_u and $\mathcal{J}_u$, respectively. Again, the function $\mathcal{F}$ has a symmetry: $\mathcal{F}(\mathcal{J}_u^{-1}H_k) = \mathcal{J}_u^{-1}\mathcal{F}(H_k)$, which follows from the fact that $\mathcal{J}_u$ commutes with $\mathbb{I}^\pm$ and differentiation. The conclusion is that $J_u \circ \mathcal{U}_{H_0'} = \mathcal{U}_{H_0} \circ J_u$ on $\mathcal{D}(\varrho')$, for all $u \in \mathbb{R}^d$ and $H_0 \in \hat{B}$.

If u belongs to the set $\mathcal{D}_1(\delta)$, then J_u maps $\mathcal{D}(\rho')$ into $\mathcal{D}(\varrho')$, and $\mathcal{J}_u$ maps $B(\rho, b)$ into $B(\varrho, b)$. Thus, for any given point (q,p) in $\mathcal{D}(\rho')$, and any given $H \in B(\rho, b)$, the function $f_H : \mathcal{D}_1(\delta) \to \mathbb{C}^{2d}$,

$$f_H(u) = \bigl(J_u \circ \mathcal{U}_{\mathcal{J}_u^{-1}H} - \mathcal{U}_H \circ J_u\bigr)(q,p)\,, \tag{3.31}$$

is analytic on $\mathcal{D}_1(\delta)$. But f_H vanishes on $\mathbb{R}^d$, as was shown above, which implies that f_H is identically zero. Since (q,p) and H were arbitrary, this proves the assertion regarding translations.

$\blacksquare$

4. Properties of the Transformations $\Re_\mu$

After proving a claim made in the introduction concerning the existence of matrices satisfying (T1–3), and a corollary on the diophantine property of ω, we will give a proof of Lemma 1.4 and Theorem 1.6. Then some symmetry properties of the transformation $\Re_\mu$ and its local stable manifold $\mathcal{W}^s$ will be discussed. Here, and in the remaining part of this paper, ρ is considered a fixed (but arbitrary) positive real number.

Let $\omega = (1, \omega_2, \ldots, \omega_d)$ be a vector whose components are real algebraic numbers.

Lemma 4.1. *The algebraic extension of $\mathbb{Q}$ by the numbers $\omega_2, \ldots, \omega_d$ is of degree d if and only if there exists an integral $d \times d$ matrix T with the properties (T1–2). Furthermore, some of the (integral) matrices satisfying (T1–2) have determinant 1.*

Proof. Let K be the smallest field containing the numbers $1, \omega_2, \ldots, \omega_d$. For any $\alpha \in K$, define $\varphi_1(\alpha) = \alpha$, and denote by $\varphi_2(\alpha), \ldots, \varphi_d(\alpha)$ the field conjugates of α.

In order to prove the "if" part, assume that T is an integral $d \times d$ matrix satisfying (T1) and (T2). Then the eigenvalue ϑ of T is an algebraic integer of degree d. ¿From the equation $(\vartheta\mathrm{I} - T)\omega = 0$ we see that the numbers $\omega_2, \ldots, \omega_d$ belong to the field $\mathbb{Q}[\vartheta]$. But $T^n\omega = (\vartheta^n, \ldots)$ for all $n \geq 0$, which shows that $\mathbb{Q}[\vartheta]$ is contained in K. Thus, K is equal to $\mathbb{Q}[\vartheta]$ which is of degree d.

Conversely, assume now that K is of degree d. Let $\gamma \in K$ be an algebraic integer of degree d. Then there exists a $d \times d$ matrix C over $\mathbb{Q}$ such that $C\omega = \gamma\omega$. Let $s \in \mathbb{Z}$ be such that $B = sC$ is a matrix over $\mathbb{Z}$, and define $\beta = s\gamma$. Clearly, β is an algebraic integer in K of degree d, and $B\omega = \beta\omega$. Denote by D the discriminant of the basis $1, \beta, \ldots, \beta^{d-1}$.

The remaining part of the proof is based on the construction, given in [21, Theorem 5.2.2], of a Pisot–Vijayaraghavan number (in an arbitrary real algebraic field) that is a unit. The construction involves a sequence of real numbers $\delta_1 > \delta_2 > \ldots > 0$ which will be defined later.

Let now δ_j be an arbitrary positive real number less than 1. By Minkowski's theorem on linear forms, we can find $c_0, \ldots, c_{d-1} \in \mathbb{Z}$, not all zero, such that

$$\left| \sum_{n=0}^{d-1} c_n \beta^n \right| \leq |D|^{1/2} \delta_j^{1-d} , \qquad \left| \sum_{n=0}^{d-1} c_n \varphi_k(\beta^n) \right| \leq \delta_j \, , \qquad k = 2, \ldots, d \, . \tag{4.1}$$

In other words, if p_j is the polynomial defined by the equation $p_j(x) = c_0 + c_1 x + \ldots + c_{d-1} x^{d-1}$, then $\alpha_j = p_j(\beta)$ is an algebraic integer in K whose conjugates are all of modulus less that δ_1. In addition, we have $|\alpha_j| > 1$ since $N(\alpha_j) = \varphi_1(\alpha_j) \cdots \varphi_d(\alpha_j)$ is a nonzero element of $\mathbb{Z}$. Thus, given that $|\varphi_k(\alpha_j)| > 1$ only for $k = 1$, the degree of α_j has to be d, and in particular, the numbers $\varphi_1(\alpha_j), \ldots, \varphi_d(\alpha_j)$ are pairwise distinct and nonzero. Now define $A_j = p_j(B)$. Then $A_j \omega = \alpha_j \omega$. Since the minimal polynomial for α_j is unique, and thus equal to $t \mapsto \det(t\mathrm{I} - A_j)$, the eigenvalues of A_j are precisely the number α_j and its conjugates $\varphi_k(\alpha_j)$. This shows that if we set $T = A_j$, then T satisfies (T1) and (T2).

In order to find a matrix T that satisfies (T3) as well, consider the above construction for $j = 1, 2, \ldots$, where $\delta_{j+1} > 0$ is chosen to be less than the minimum of the numbers $|\varphi_2(\alpha_j)|, \ldots, |\varphi_d(\alpha_j)|$. Since $|N(\alpha_j)| \leq |D|^{1/2}$ for all j, we can find $m \in \mathbb{Z}$, and an infinite set $J \subset \mathbb{N}$, such that $N(\alpha_j) = m$ for all $j \in J$. Then there exists a pair of indices $i > j$ in J, such that each coefficient of the polynomial p_i is congruent modulo m to the corresponding coefficient of p_j. In other words, $p_i = p_j + mq$ for some polynomial q with coefficients in $\mathbb{Z}$. Let now $T = \mathrm{I} + q(B) A_j^{\mathrm{adj}}$, where A_j^{adj} is the matrix over $\mathbb{Z}$ such that $A_j A_j^{\mathrm{adj}} = \det(A_j)\mathrm{I}$. Since $\det(A_j) = m$, we have $TA_j = A_j + mq(B) = A_i$. Thus, $T = A_i A_j^{-1}$, and in particular, $\det(T) = 1$. The eigenvalues of T are $\vartheta_k = \varphi_k(\alpha_i)/\varphi_k(\alpha_j)$, for $k = 1, 2, \ldots, d$, and the corresponding eigenvectors are the same as those of B. All eigenvalues except $\vartheta_1 = \alpha_i/\alpha_j$ are of modulus less than one, since $|\varphi_k(\alpha_i)| \leq \delta_i < |\varphi_k(\alpha_j)|$ for all $k > 1$. Thus, as in the case of the matrices A_j, we conclude that T satisfies (T1) and (T2).

■

In what follows, T is assumed to be an integral $d \times d$ matrix satisfying (T1) and (T2). Property (T3) will not be needed in this section. Let $\vartheta_1, \vartheta_2, \ldots, \vartheta_d$ be the eigenvalues of T, labeled such that $|\vartheta_i| \geq |\vartheta_j|$ for $i \leq j$, and let $\omega^{(1)}, \omega^{(2)}, \ldots, \omega^{(d)}$ be the corresponding eigenvectors. In particular, $\vartheta_1 = \vartheta$ and $\omega^{(1)} = \omega$.

Corollary 4.2. *There exists a constant $c > 0$ such that every nonzero $\nu \in \mathbb{Z}^d$ satisfies $|\omega \cdot \nu| \geq c|\nu|^{-\gamma}$, where $\gamma = -\ln|\vartheta_1| / \ln|\vartheta_2|$.*

Proof. Consider the norm on $\mathbb{R}^d$ defined by the equation

$$\|x\| = \max_{1 \leq j \leq d} \left| \, \omega^{(j)} \cdot x \, \right| , \qquad x \in \mathbb{R}^d \, . \tag{4.2}$$

Since all norms on $\mathbb{R}^d$ are equivalent, there exists a constant $b > 0$ such that $\|\nu\| \geq b$ for every nonzero $\nu \in \mathbb{Z}^d$. Let now $\nu \in \mathbb{Z}^d$ be fixed, and assume that $|\omega^{(1)} \cdot \nu| < \|\nu\|$. Let k be the smallest non-negative integer such that

$$\|(T^*)^{k+1}\nu\| = \left|\omega^{(1)} \cdot \left[(T^*)^{k+1}\nu\right]\right| . \tag{4.3}$$

Then, for some $j > 1$, we have

$$b \leq \|(T^*)^k\nu\| = \left|\omega^{(j)} \cdot \left[(T^*)^k\nu\right]\right| = |\vartheta_j|^k \left|\omega^{(j)} \cdot \nu\right| \leq |\vartheta_2|^k \|\nu\| , \tag{4.4}$$

which implies that

$$|\vartheta_1|^{-k} = |\vartheta_2|^{\gamma k} \geq b^\gamma \|\nu\|^{-\gamma} . \tag{4.5}$$

By combining (4.3) with this last inequality, we obtain

$$\begin{aligned} \left|\omega^{(1)} \cdot \nu\right| &= |\vartheta_1|^{-k-1} \left|\omega^{(1)} \cdot \left[(T^*)^{k+1}\nu\right]\right| = |\vartheta_1|^{-k-1} \|(T^*)^{k+1}\nu\| \\ &\geq b|\vartheta_1|^{-k-1} \geq b^{\gamma+1}|\vartheta_1|^{-1}\|\nu\|^{-\gamma} , \end{aligned} \tag{4.6}$$

and the assertion follows.

■

Proof of Lemma 1.4. Let $r > \rho$ be such that $\rho' > \kappa r$. Define $M = \mu(T^*)^{-1}$. Then, by assumption (1.9), we have $\|T^*\|_\omega < \kappa$ and $\|M^*\|_0 < \kappa < \rho'/r$, where $\|.\|_x$ is the seminorm introduced in (3.21).

Choose $\sigma > 0$ such that $|T^* x| \leq \kappa$ whenever $|x| = 1$ and $|\omega \cdot x| \leq \sigma$. In addition, assume that $|\sigma| < 1/|\Omega|$. This condition will not be used until later on. Furthermore, choose $\tau > 0$ such that

$$\left(\kappa r/\rho'\right)^\tau e^{r\|T^*\| - \rho'} < 1 . \tag{4.7}$$

The two constants σ and τ define the projections $\mathbb{I}^\pm$, as described in Definition 1.2.

Let now H be a function in $\mathcal{A}(\rho')$, such that $\mathbb{I}^- H = 0$. Consider the representation (3.19) for H. Then $H_\nu^{(n)} = 0$, unless either $|\omega \cdot \nu| \leq \sigma|\nu|$ or $n \geq \tau|\nu|$. In the first case, when $|\omega \cdot \nu| \leq \sigma|\nu|$, we have the bound

$$\begin{aligned} \|H_\nu^{(n)} \circ M\| r^n e^{r|T^*\nu|} &\leq \|M^*\|_0^n \|H_\nu^{(n)}\| r^n e^{r|T^*\nu|} \\ &\leq \left(\kappa r/\rho'\right)^n e^{(r\kappa - \rho')|\nu|} \|H_\nu^{(n)}\| (\rho')^n e^{\rho'|\nu|} \\ &\leq \|H_\nu^{(n)}\| (\rho')^n e^{\rho'|\nu|} , \end{aligned} \tag{4.8}$$

and in the case $n \geq \tau|\nu|$ we obtain

$$\begin{aligned} \|H_\nu^{(n)} \circ M\| r^n e^{r|T^*\nu|} &\leq \|M^*\|_0^n \|H_\nu^{(n)}\| r^n e^{r|T^*\nu|} \\ &\leq \left[\left(\kappa r/\rho'\right)^\tau e^{r\|T^*\| - \rho'}\right]^{|\nu|} \|H_\nu^{(n)}\| (\rho')^n e^{\rho'|\nu|} \\ &\leq \|H_\nu^{(n)}\| (\rho')^n e^{\rho'|\nu|} , \end{aligned} \tag{4.9}$$

by using the inequality (4.7). These two estimates imply that $H \circ T_\mu$ extends analytically to $\mathcal{D}(r)$ and continuously to the boundary of $\mathcal{D}(r)$, and that

$$\|H \circ T_\mu\|_r = \sum_{\substack{\nu \in \mathbb{Z}^d \\ n \geq 0}} \|H_\nu^{(n)} \circ M\| r^n e^{r|T^*\nu|} \leq \|H\|_{\rho'} . \tag{4.10}$$

The assertion now follows from the fact that the inclusion map from $\mathcal{A}(r)$ to $\mathcal{A}(\rho)$ is compact.

■

In the remaining part of this paper, the parameters σ and τ that define the projection operators $\mathbb{I}^\pm$ are considered fixed. To be more precise, we pick some arbitrary number ρ' in $(\kappa\rho, \rho)$ and choose σ, τ as described in the proof of Lemma 1.4.

Proof of Theorem 1.6. Choose two numbers $\varrho' < \varrho$ in the interval (ρ', ρ). By Theorem 2.12, the map $\mathcal{R}: H \mapsto H \circ \mathcal{U}_H$ is analytic, from some open ball $\hat{B} \subset \mathcal{A}(\varrho)$ centered at H^0, to $\mathcal{A}(\rho')$. And by Lemma 1.4, the linear operator $L_\mu: \mathcal{A}(\rho') \to \mathcal{A}(\rho)$, defined by equation (1.13), is compact. Consider now the two linear functionals $f_0: H \mapsto (\mathbb{E}H)(0)$ and $f_1: H \mapsto (\Omega \cdot \nabla \mathbb{E}H)(0)$ that enter the definition (1.11) of $\Re_\mu$. These functionals are clearly continuous on $\mathcal{A}(\varrho)$. Furthermore, if f_1 is restricted to a sufficiently small neighborhood B' of H^0 in $\mathcal{A}(\varrho)$, then its range is contained in the disk $|z-1| < 1/2$. Thus, the equation $\mathcal{F}(H) = [H - f_0(H)]/f_1(H)$ defines a bounded and analytic map $\mathcal{F}$ from B' to $\mathcal{A}(\varrho)$. By combining all this, we see that if the radius of $\hat{B}$ is chosen sufficiently small, then $\Re_\mu = \mathcal{F} \circ L_\mu \circ \mathcal{R}$ is well defined, analytic, and compact, as a map from $\hat{B}$ to $\mathcal{A}(\rho)$.

Consider now $u \in \mathbb{C}^d$ such that $|u| < \rho - \varrho$ and $|u| < \varrho' - \rho'$. Then $\mathcal{J}_u H \in \mathcal{A}(\varrho)$ whenever $H \in \mathcal{A}(\rho)$, and $\mathcal{J}_u$ is bounded as a linear operator from $\mathcal{A}(\rho)$ to $\mathcal{A}(\varrho)$. In addition, $\mathcal{J}_u H^0 = H^0$. Thus, some open ball $B \subset \mathcal{A}(\rho)$, centered at H^0, is mapped into $\hat{B}$ by $\mathcal{J}_u$; and the map $\Re_\mu \circ \mathcal{J}_u$ is analytic and compact, when restricted to B. Let now $H \in B$ be fixed. Define $H' = \mathcal{R}(H)$ and $H'' = L_\mu H'$. If B has been chosen sufficiently small (which we assume), then Lemma 3.10 implies that $\mathcal{R}(\mathcal{J}_u H) = \mathcal{J}_u \mathcal{R}(H)$. Next, by using that $\mathcal{J}_u$ commutes with $\mathbb{I}^+$, and that $T_\mu^{-1} \circ J_u \circ T_\mu = J_{T^{-1}u}$, we obtain

$$\begin{aligned} L_\mu \mathcal{J}_u H' &= \frac{\vartheta}{\mu}\bigl(\mathbb{I}^+ H'\bigr) \circ J_u^{-1} \circ T_\mu = \frac{\vartheta}{\mu}\bigl(\mathbb{I}^+ H'\bigr) \circ T_\mu \circ \bigl(T_\mu^{-1} \circ J_u \circ T_\mu\bigr)^{-1} \\ &= \mathcal{J}_{T^{-1}u} L_\mu H' . \end{aligned} \tag{4.11}$$

In addition, $\mathcal{F}(\mathcal{J}_{T^{-1}u} H'') = \mathcal{J}_{T^{-1}u} \mathcal{F}(H'')$, since $\mathbb{E}\mathcal{J}_v H'' = \mathbb{E}H''$ for any v. By combining these "commutation relations", the claim $\Re_\mu(\mathcal{J}_u H) = \mathcal{J}_{T^{-1}u} \Re_\mu(H)$ follows.

Next, let $z \in \mathbb{C} \setminus \{0\}$ and $H \in B \cap \mathcal{A}(|z|\rho)$ be given, such that $\mathcal{S}_z^{-1} H \in B$. Assume first that $|z| \leq 1$. Define $H' = \mathcal{R}(H)$ and $H'' = L_\mu H'$, as before. By Lemma 3.10 we have $\mathcal{R}(\mathcal{S}_z^{-1} H) = \mathcal{S}_z^{-1} \mathcal{R}(H)$. Two analogous identities for L_μ and $\mathcal{F}$ follow directly from the definitions: $L_\mu \mathcal{S}_z^{-1} H' = \mathcal{S}_z^{-1} L_\mu H'$ and $\mathcal{F}(\mathcal{S}_z^{-1} H'') = \mathcal{S}_z^{-1} \mathcal{F}(H'')$. By combining these three identities, we obtain $\Re_\mu(\mathcal{S}_z^{-1} H) = \mathcal{S}_z^{-1} \Re_\mu(H)$, as claimed in Theorem 1.6. In particular, $\Re_\mu(H) \in \mathcal{S}_z \mathcal{A}(\rho)$.

In the case $|z| > 1$, consider the function $H_z = \mathcal{S}_z^{-1}H$. Then the argument above shows that $\mathfrak{R}_\mu(\mathcal{S}_z H_z) = \mathcal{S}_z \mathfrak{R}_\mu(H_z)$. By substituting back $H_z = \mathcal{S}_z^{-1}H$, we obtain the same conclusion as in the case $|z| \le 1$.

■

In order to study the spectral properties of $D\mathfrak{R}_\mu(H^0)$, it is convenient to introduce the following projection operators $\mathbb{P}_n$, $\mathbb{P}_{\ge n}$, and $\mathbb{P}_{1,1}$. If $n \in \mathbb{N}$ and $H \in \mathcal{A}(\rho)$, define

$$\bigl(\mathbb{P}_n H\bigr)(q,p) = H_0^{(n)}(p), \qquad \bigl(\mathbb{P}_{\ge n} H\bigr)(q,p) = \sum_{m\ge n} H_0^{(m)}(p)\,, \tag{4.12}$$

and in particular, $(\mathbb{P}_{\ge 0}H)(q,p) = (\mathbb{E}H)(p)$, for all $(q,p) \in \mathcal{D}(\rho)$. Here, $H_0^{(m)}$ denotes the part of H that is homogeneous of order m in the second variable, averaged over the torus $\mathbb{T}^d$; see the definition (3.19). Furthermore, let $\mathbb{P}_{1,1}H = \bigl(\Omega\cdot\nabla H_0^{(1)}\bigr)H^0$, where Ω is the eigenvector of T^* corresponding to the eigenvalue ϑ, normalized such that $\Omega\cdot\omega = 1$. Notice that all these projection operators commute with $L_\mu\colon h \mapsto \frac{\vartheta}{\mu}(\mathbb{I}^+ h)\circ T_\mu$.

The following is a spectral theorem for the operator L_μ .

Lemma 4.3. *The eigenvectors ϕ_β of L_μ that are listed in (1.14) span a dense subspace of $\mathbb{P}_{\ge 0}\mathcal{A}(\rho)$. The spectral projection corresponding to the eigenvalue zero of L_μ is $\mathbb{I} - \mathbb{P}_{\ge 0}$. More precisely, there are constants $a > 0$ and $c > 1$, such that for all positive integers m,*

$$\bigl\|L_\mu^{m+1}\bigl(\mathbb{I} - \mathbb{P}_{\ge 0}\bigr)\bigr\| \le \left|\frac{\vartheta}{\mu}\right| e^{-ac^m}\bigl\|L_\mu^{m}\bigl(\mathbb{I} - \mathbb{P}_{\ge 0}\bigr)\bigr\|\,. \tag{4.13}$$

Proof. The first part, including the fact that $L_\mu\phi_\beta = \lambda_\beta\phi_\beta$ for all $\beta \in \mathbb{N}^d$, is obvious.

What remains to be proved is the bound (4.13). Let $k = m-1$. For every integer $n \ge 0$, denote by V_n^+ the set of all vectors $\nu \in \mathbb{Z}^d$ that satisfy $|\omega\cdot\nu| \le \sigma|\nu|$ or $\tau|\nu| \le n$. By using the representation (3.19) for $h \in \mathcal{A}(\rho)$, we can write

$$\bigl(L_\mu^{k+1}h\bigr)(q,p) = \sum_{n\ge 0}\bigl(\vartheta\mu^{n-1}\bigr)^{k+1}\sum_\nu h_\nu^{(n)}\bigl((T^*)^{-k-1}p\bigr)\exp\Bigl(i\bigl[(T^*)^{k+1}\nu\bigr]\cdot q\Bigr)\,, \tag{4.14}$$

for all $(q,p) \in \mathcal{D}(\rho)$, where the sum $\sum_\nu$ ranges over all $\nu \in \mathbb{Z}^d$ such that $(T^*)^j\nu \in V_n^+$, for $j = 0,1,\ldots,k$. In particular, if $\mathbb{P}_{\ge 0}h = 0$, then $L_\mu^{k+1}h$ only depends on those components $h_\nu^{(n)}$ of h that are indexed by pairs (ν,n) with the property that $(T^*)^k\nu \in V_n^+\setminus\{0\}$. Below we will show that these indices satisfy one of the bounds

$$n \ge b_1|\vartheta_1|^{k/2}\,, \quad \text{or} \quad |\nu| \ge b_2|\vartheta_2|^{-k/2}\,, \tag{4.15}$$

where b_1 and b_2 are positive constants independent of k.

Consider now the projection operator $\mathbb{I}_k^+$, which acts on a function $H \in \mathcal{A}(\rho)$ by restricting the sum in the expansion (3.19) to index pairs (ν,n) satisfying $(T^*)^k\nu \in V_n^+ \setminus \{0\}$. In order to prove (4.13), we write $L_\mu^{m+1}(\mathbb{I} - \mathbb{P}_{\ge 0})$ as a product $A\mathbb{I}_k^+L_\mu$, where $A =$

$L_\mu^{k+1}(\mathbb{I}-\mathbb{P}_{\geq 0})$, and then we bound the operator norm of this product by $\|A\|\cdot\|\mathbb{I}_k^+ L_\mu\|$. As was shown in the proof of Lemma 1.4, there exists $r>\rho$, such that L_μ is continuous as a linear operator from $\mathcal{A}(\rho)$ to $\mathcal{A}(r)$, and bounded in norm by $|\vartheta/\mu|$. Now regard $\mathbb{I}_k^+$ as a linear operator from $\mathcal{A}(r)$ to $\mathcal{A}(\rho)$. By using the inequalities (4.15), the norm of this operator can be bounded from above by $e^{-a'c^k}$, with $a'>0$ and $c>1$ independent of k. This proves the bound (4.13).

In order to prove that (4.15) holds whenever $(T^*)^k\nu\in V_n^+\setminus\{0\}$, consider first the case where $\bigl|\omega\cdot(T^*)^k\nu\bigr|\leq\sigma\bigl|(T^*)^k\nu\bigr|$. Then

$$|\omega\cdot\nu|=|\vartheta_1|^{-k}\bigl|\omega\cdot(T^*)^k\nu\bigr|\leq\sigma\bigl|(\vartheta_1^{-1}T^*)^k\nu\bigr|\leq\sigma\bigl|(\omega\cdot\nu)\Omega\bigr|+\sigma\bigl|(\vartheta_1^{-1}T^*)^kP\nu\bigr|\,, \tag{4.16}$$

where $P\nu=\nu-(\omega\cdot\nu)\Omega$ is the component of ν in the spectral subspace of T^* corresponding to the eigenvalues $\vartheta_2,\ldots,\vartheta_d$ of modulus less than one. Thus, since $|\sigma|<1/|\Omega|$, we can bound $|\omega\cdot\nu|$ from above by a constant times $|\vartheta_2/\vartheta_1|^k|\nu|$. Combining this with the lower bound from Corollary 4.2, we find that

$$b_3|\nu|^{-\gamma}\leq|\vartheta_2/\vartheta_1|^k\,|\nu|\,, \tag{4.17}$$

for some constant $b_3>0$. By using the definition of γ, this inequality can be rewritten as $|\nu|\geq b_2|\vartheta_2|^{-k}$, which implies (4.15).

Next, consider the case where $\tau\bigl|(T^*)^k\nu\bigr|\leq n$. Then there exists $b_4>0$ such that

$$|\omega\cdot\nu|=|\vartheta_1|^{-k}\bigl|\omega\cdot(T^*)^k\nu\bigr|\leq b_4|\vartheta_1|^{-k}\bigl|(T^*)^k\nu\bigr|\leq b_4|\vartheta_1|^{-k}n/\tau\,. \tag{4.18}$$

This bound, together with Corollary 4.2, implies that

$$n|\nu|^\gamma\geq b_5|\vartheta_1|^k\,, \tag{4.19}$$

for some constant $b_5>0$. The two alternatives in (4.15) are now obtained by distinguishing the cases $n\geq|\nu|^\gamma|$ and $|\nu|^\gamma\geq n$.

■

This lemma translates easily into a spectral theorem for $D\Re_\mu(H^0)$, since $D\Re_\mu(H^0)$ differs from L_μ only on the two-dimensional invariant subspace spanned by the constant functions and the function $(q,p)\mapsto\omega\cdot p$. More precisely, we have

$$D\Re_\mu\bigl(H^0\bigr)=L_\mu\bigl(\mathbb{I}-\mathbb{P}_0-\mathbb{P}_{1,1}\bigr)\,. \tag{4.20}$$

Notice also that, by Theorem 1.3, the intersection of $\mathbb{P}_{\geq 0}\mathcal{A}(\rho)$ with the domain of $\Re_\mu$ is mapped into $\mathbb{P}_{\geq 0}\mathcal{A}(\rho)$ by the transformation $\Re_\mu$.

Throughout the rest of this paper, we assume that μ satisfies the bound (1.15). Then the spectrum of $D\Re_\mu\bigl(H^0\bigr)$ splits into an unstable part, lying in the complex domain $|z|\geq\theta_u>1$, and a stable part, lying in $|z|\leq\theta_s(\mu)<1$, where

$$\theta_u=|\lambda_{(0,1,0,\ldots,0)}|=|\vartheta_1/\vartheta_2|\,,\qquad\theta_s(\mu)=|\lambda_{(0,\ldots,0,2)}|=|\mu\vartheta_1/\vartheta_d^2|\,. \tag{4.21}$$

The corresponding spectral projections are $\mathbb{P}^u = \mathbb{P}_1 - \mathbb{P}_{1,1}$ and $\mathbb{P}^s = \mathbb{I} - \mathbb{P}^u$, respectively. The space $\mathcal{A}^u = \mathbb{P}^u\mathcal{A}(\rho)$ is of dimension $d-1$, spanned by the functions $h_2, \ldots, h_d$ listed in (1.17). More explicitly, we have

$$\mathbb{P}^u h = \sum_{j=2} \Omega^{(j)} \cdot \bigl(\nabla \mathbb{E} h\bigr)(0) h_j\,, \qquad h \in \mathcal{A}(\rho)\,, \tag{4.22}$$

where $\Omega^{(j)}$, for $1 \le j \le d$, denotes the eigenvector of T^* for the eigenvalue ϑ_j, normalized such that $\Omega^{(j)} \cdot \omega^{(j)} = 1$.

Let $\mathcal{A}^s(\rho)$ be the affine spaces consisting of all functions of the form $H^0 + h$, with $h \in \mathbb{P}^s\mathcal{A}(\rho)$. Given an two sets A, B, contained in the domain (as described in Theorem 1.6) of $\Re_\mu$, define $\mathcal{W}^s(A,B)$ to be the set of all functions $H \in A$ such that $\Re_\mu^n(H) \in B$ for all $n \ge 1$.

By the local stable manifold theorem [HSP], there exists an open neighborhood B of H^0 in $\mathcal{A}(\rho)$, such that $\mathcal{W}^s(B,B)$ is the graph of a differentiable map $W\colon B \cap \mathcal{A}^s(\rho) \to \mathcal{A}^u$, whose derivative vanishes at H^0 and is bounded in norm by 1 at every point in its domain. Furthermore, if $B' \subset B$ then $\mathcal{W}^s(A,B') = \mathcal{W}^s(A,B)$, for every sufficiently small open neighborhood A of H^0 that is contained in B'.

$\mathcal{W}^s = \mathcal{W}^s(B,B)$ will be referred to as the local stable manifold of $\Re_\mu$ at H^0. A similar construction, using inverse images, yields $\mathcal{W}^u = \{H \in B\colon H - H^0 \in \mathcal{A}^u\}$ as the local unstable manifold of $\Re_\mu$ at H^0. Having already introduced the necessary notation, let us also define (for use later on) the projection $\wp$, in the direction of $\mathcal{A}^u$, from B onto $\mathcal{W}^s$, by setting

$$\wp(H) = H^s = \mathbb{P}^s H + W(\mathbb{P}^s H)\,, \qquad H \in B\,. \tag{4.23}$$

Trivial examples of Hamiltonians on $\mathcal{W}^s$ are functions $H = cH^0 + f$ in B, with $f(q,p)$ depending on p only, satisfying $(\nabla_2 f)(0,0) = 0$. In this case, $\mathbb{P}^u H = 0$, as can be seen from (4.22), and $\Re_\mu(H) = H^0 + c^{-1} D\Re_\mu(H^0) f$.

The following lemma describes a consequence of the fact that the renormalization group transformation $\Re_\mu$ "commutes" with scalings $\mathcal{S}_z$; see Theorem 1.6.

Lemma 4.4. *Let $0 < a < 1$. Then there exists an open neighborhood B of H^0 in $\mathcal{A}(\rho)$ such that the following holds whenever $a < |z| \le 1$.*
(1) *$\mathcal{S}_z^{-1}$ maps $\mathcal{A}^s(\rho)$ into itself, and $W \circ \mathcal{S}_z^{-1} = W$ on $\mathcal{A}^s(\rho) \cap B$.*
(2) *The local stable manifolds of $\Re_\mu$ and $\Re_{z\mu} = \mathcal{S}_z^{-1} \circ \Re_\mu$ coincide in B.*

Proof. Given any $\varepsilon > 0$, define $B(\varepsilon)$ to be the pointwise sum $B_1(\varepsilon) \oplus B_2(\varepsilon)$ of the two sets

$$B_1(\varepsilon) = \bigl\{H \in \mathcal{A}^s(\rho)\colon \|H - H^0\|_\rho < \varepsilon\bigr\}\,, \quad B_2(\varepsilon) = \bigl\{h \in \mathcal{A}^u\colon \|h\|_\rho < \varepsilon\bigr\}\,. \tag{4.24}$$

The neighborhood B mentioned in Lemma 4.4 will be $B(\delta)$, with $\delta > 0$ specified below.

For the proof of (1), choose $0 < \delta < \varepsilon$ such that $B(\varepsilon/a) \subset B$ and $\mathcal{W}^s(B(\delta), B(\varepsilon)) = \mathcal{W}^s(B(\delta), B)$. Let z be a complex number satisfying $a \le |z| \le 1$, and let $h \in B_1(\delta)$. Then the functions $H = h + W(h)$ and $\Re_\mu^n(H)$ belong to $\mathcal{W}^s \cap B(\delta)$ and $B(\varepsilon)$, respectively, for all $n \ge 1$. Consequently, the functions $\mathcal{S}_z^{-1} H$ and $\mathcal{S}_z^{-1}\Re_\mu^n(H)$ belong to $B(\delta/a)$ and $B(\varepsilon/a)$, respectively, for all $n \geq 1$. This in turn implies that $\mathcal{S}_z^{-1}H \in \mathcal{W}^s$, since $\mathcal{S}_z^{-1}\Re_\mu^n(H) = \Re_\mu^n(\mathcal{S}_z^{-1}H)$ by Theorem 1.6. But $\mathcal{S}_z^{-1}H = \mathcal{S}_z^{-1}h + \mathcal{S}_z^{-1}W(h) = \mathcal{S}_z^{-1}h + W(h)$, and $\mathcal{S}_z^{-1}h \in B_1(\delta/a)$, which shows that $W(\mathcal{S}_z^{-1}h) = W(h)$. Thus, $W \circ \mathcal{S}_z^{-1} = W$ on $B_1(\delta)$, as claimed.

In order to prove (2), using analyticity arguments, we first extend the function $z \mapsto \Re_{z\mu}$ to an open region containing the circle $|z| = 1$. To this end, it suffices to note that in the proof of Theorem 1.6, the value of μ enters only through the condition $0 < \|\mu T^{-1}\|_0 < \kappa$. Consequently, we can find constants $\varepsilon > 0$ and $c > 1$, such that $\Re_{z\mu} = \mathcal{S}_z^{-1} \circ \Re_\mu$ is well defined, analytic, and compact, as a map from some open neighborhood of $\overline{B(\varepsilon)}$, to $\mathcal{A}(\rho)$, for all z in the region $U = \{z \in \mathbb{C} \colon 0 < |z| < c\}$. Here, and in what follows, $\overline{V}$ denotes the closure of a set V in a given topological space.

We claim that $z \mapsto \Re_{z\mu}$ is differentiable, as a map from U to $C^1\bigl(\overline{B(\varepsilon)}, \mathcal{A}(\rho)\bigr)$. This can be seen as follows. Let $G \subset \mathbb{C}$ be a fixed but arbitrary open disk whose closure is contained in U, and define

$$\mathcal{F}_z = \frac{1}{2\pi i} \oint_{\zeta \in \partial G} \frac{d\zeta}{\zeta - z} \Re_{\zeta\mu}\,, \qquad z \in G\,, \tag{4.25}$$

where ∂G denotes the (counterclockwise oriented) boundary of G. This integral clearly converges in $C^1\bigl(\overline{B(\varepsilon)}, \mathcal{A}(\rho)\bigr)$. Now we evaluate both sides of (4.25), first at $H \in B(\varepsilon)$, and then at $(q,p) \in \mathcal{D}(\rho)$. Since convergence in C^1 and in $\mathcal{A}(\rho)$ implies pointwise convergence, and since the function $\zeta \mapsto (\Re_{\zeta\mu}(H))(q,p) = \zeta^{-1}(\Re_\mu(H))(q,\zeta p)$ is analytic on U, we find that $(\mathcal{F}_z(H))(q,p) = (\Re_{z\mu}(H))(q,p)$, for all $z \in G$. But since (q,p) and H were arbitrary, this implies that $\mathcal{F}_z = \Re_{z\mu}$ for all $z \in G$. The identity (4.25) now shows that $z \mapsto \Re_{z\mu}$ is analytic on G. Analyticity on U follows since G was arbitrary.

Let now a and b be fixed but arbitrary constants such that $0 < a < 1 < b < c$. Define $V = \{z \in \mathbb{C} \colon a < |z| < b\}$. By the (local) stable manifold theorem, there exists, for every $z \in \overline{V}$, an open neighborhood $B(\delta_z)$ of H^0, where the local stable manifold $\mathcal{W}_z^s$ of the transformation $\Re_{z\mu}$ is the graph of a differentiable map $W_z \colon B_1(\delta_z) \to B_2(\delta_z)$, whose derivative is bounded in norm by 1. The proof [HPS] of this theorem also yields differentiable dependence of the local stable manifold on the transformation, or in our case (since $z \mapsto \Re_{z\mu}$ is differentiable), on the parameter z. That is, the abovementioned constants δ_z can all be chosen larger than some positive $\delta < \varepsilon$, and $z \mapsto W_z$ is differentiable (analytic) as a map from V to $C^1\bigl(\overline{B_1(\delta)}, \mathcal{A}^u\bigr)$. Now restrict W_z and $\mathcal{W}_z^s$ to the set $B_1(\delta)$ and $B(\delta)$, respectively, for each $z \in V$. According to Theorem 1.6, if $|z| = 1$ then $\Re_\mu$ commutes with $\mathcal{S}_z^{-1}$. This implies that $\mathcal{W}_z^s = \mathcal{W}_1^s$, and thus $W_z = W_1$, whenever $|z| = 1$. Since $z \mapsto W_z$ is analytic on V, we conclude that $W_z = W_1$ for all $z \in V$.

■

One of the consequences of this lemma is the following

Corollary 4.5. *If H is sufficiently close (but not necessarily equal) to H^0 in $\mathcal{A}(\rho)$, and either of the form $H(q,p) = c\omega \cdot p + f(q)$, or $H(q,p) = c\omega \cdot p + f(q,p)$ with $f(q,p) = \mathcal{O}(|p|^2)$, then H lies on $\mathcal{W}^s$.*

Proof. First, we note that if H is of the form indicated above, then H belongs to $\mathcal{A}^s(\rho)$. Let now $0 < a < b < 1$. Let B' be an open ball in $\mathcal{A}(\rho)$, centered at H^0, such that a ball with twice the radius of B' is contained in the neighborhood B mentioned in Lemma 4.4.

Consider $H_z \in \mathcal{A}(\rho)$ of the form $H_z(q,p) = c\omega \cdot p + zf(q)$. Assume that H_1 belongs to B'. Then $H_z \in B$, for all z in the disk $G_2 = \{z \in \mathbb{C}\colon |z| < 2\}$. Furthermore, $H_z = \mathcal{S}_z H_1$ if $z \neq 0$. Thus, by Lemma 4.4, the function $z \mapsto W(H_z)$ is constant on the unit circle $|z| = 1$. On the other hand, this function is analytic on G_2, and therefore constant on G_2. Consequently, $W(H_1) = W(H_0) = 0$, as claimed.

Next, consider the case where f is a function in $\mathcal{A}(\rho)$ whose Fourier–Taylor coefficients $f_{\nu,\alpha}$ are zero unless $|\alpha| \geq 2$. Let $H = H_0 + f$, where H_0 is a constant multiple of H^0, and assume that $H \in B'$. Define $H_z = \mathcal{S}_z^{-1} H$, whenever $0 < |z| \leq 1$. Then $z \mapsto H_z$ is analytic as a map from the disk $G_1 = \{z \in \mathbb{C}\colon |z| < 1\}$ to B'. Thus, $z \mapsto W(H_z)$ is an analytic function on G_1. But by Lemma 4.4, this function is constant on the circle $|z| = b$, and $W(H_b) = W(H)$. Thus, $W(H) = W(H_b) = W(H_0) = 0$, and the assertion follows.

■

Now the analogue of Lemma 4.4 for translations on the torus:

Lemma 4.6. *Let $r - \rho > \varepsilon > 0$. Then there exists an open neighborhood B of H^0 in $\mathcal{A}(r)$, such that that the following holds, whenever $u \in \mathcal{D}_1(\varepsilon)$. The operator $\mathcal{J}_u : \mathcal{A}(r) \to \mathcal{A}(\rho)$ maps $\mathcal{A}^s(r)$ into $\mathcal{A}^s(\rho)$ and $\mathcal{W}^s \cap B$ into $\mathcal{W}^s$. Furthermore, $W \circ \mathcal{J}_u = W$ on $\mathcal{A}^s(\rho) \cap B$.*

Proof. Under the given assumption, $\mathcal{J}_u$ is a bounded linear operator from $\mathcal{A}(r)$ to $\mathcal{A}(\rho)$, and it commutes with $\mathbb{P}^s$, in the sense that $\mathcal{J}_u \mathbb{P}^s H = \mathbb{P}^s \mathcal{J}_u H$ for all $H \in \mathcal{A}(r)$.

Consider an open neighborhood $B(\delta') = B_1(\delta') \oplus B_2(\delta')$ of H^0 in $\mathcal{A}(\rho)$, as defined in (4.24), where $\mathcal{W}^s$ is the graph of an analytic function $W : B_1(\delta') \to B_2(\delta')$. Choose a positive $\delta < \delta'$ such that $\Re_\mu^n(H)$ belongs to $B(\delta')$, for all $H \in \mathcal{W}^s \cap B(\delta)$ and $n \geq 0$. Then choose an open ball B in $\mathcal{A}(r)$, centered at H^0, such that $\mathcal{J}_u$ maps B into $B(\delta)$, whenever u belongs to $\mathcal{D}_1(\varepsilon)$. Let $H \in \mathcal{W}^s \cap B$ be fixed. Then

$$\mathcal{J}_u H = \mathcal{J}_u \mathbb{P}^s H + \mathcal{J}_u W\bigl(\mathbb{P}^s H\bigr) = \mathbb{P}^s \mathcal{J}_u H + W\bigl(\mathbb{P}^s H\bigr)\,. \tag{4.26}$$

This equation shows that $\mathcal{J}_u H \in \mathcal{W}^s$ if and only if $W(\mathbb{P}^s H) = W(\mathcal{J}_u \mathbb{P}^s H)$. Assume now that $u \in \mathbb{R}^d$. Since $\mathcal{J}_{T^{-n}u}$ is an isometry, it follows from Theorem 1.6, that $\Re_\mu^n(\mathcal{J}_u H)$ belongs to $B(\delta')$, for all $n \geq 0$. Thus $\mathcal{J}_u H \in \mathcal{W}^s \cap B$, which implies that $W(\mathbb{P}^s H) = W(\mathcal{J}_u \mathbb{P}^s H)$. Next, consider the function $f\colon \mathcal{D}_1(\varepsilon) \mapsto \mathcal{A}^u$, defined by the equation $f(u) = W(\mathbb{P}^s H) - W(\mathcal{J}_u \mathbb{P}^s H)$. This function is analytic, since $u \mapsto \mathcal{J}_u H$ is analytic as a map from $\mathcal{D}_1(\varepsilon)$ to $B(\delta)$. But f vanishes on $\mathbb{R}^d$, and thus $W(\mathbb{P}^s H) - W(\mathcal{J}_u \mathbb{P}^s H) = 0$ for all $z \in \mathcal{D}_1(\varepsilon)$. This in turn implies that $\mathcal{J}_u H \in \mathcal{W}^s \cap B$ for all $z \in \mathcal{D}_1(\varepsilon)$. Since $H \in \mathcal{W}^s \cap B$ was arbitrary, the assertion is proved.

■

Finally, we estimate how fast a point on $\mathcal{W}^s$ approaches H^0 and its torus average, respectively, under the iteration of $\Re_\mu$.

Lemma 4.7. *Let $\mu \in \mathbb{C}$ and $\theta \in \mathbb{R}$ be given, satisfying $0 < \theta_s(\mu) < \theta < 1$. Then there exists an open neighborhood B of H^0 in $\mathcal{A}(\rho)$, three constants $K, a > 0$ and $c > 1$, and a positive integer N, such that*

$$\left\|\Re_\mu^m(H) - H^0\right\|_\rho \leq K\theta^m \left\|H - H^0\right\|_\rho, \tag{4.27}$$

$$\left\|\left(\mathbb{I} - \mathbb{P}_{\geq 0}\right)\Re_\mu^n(H)\right\|_\rho \leq e^{-ac^n}\left\|H - H^0\right\|_\rho + \theta^{2n}\left\|H - H^0\right\|_\rho^2, \tag{4.28}$$

whenever $m \geq 0$, $n \geq N$, and $H \in \mathcal{W}^s \cap B$.

Proof. Fix two numbers $t' < t$ in the interval $(\theta_s(\mu), \theta)$. Since $D\Re_\mu(H^0)$ is compact, we can choose a norm $\|.\|$ on $\mathcal{A}(\rho)$, which is equivalent to $\|.\|_\rho$, such that $\|D\Re_\mu(H^0)f\| \leq t'\|f\|$ for all $f \in \mathbb{P}^s\mathcal{A}(\rho)$.

Given a function $H' \in \mathcal{W}^s$, we can write $h' = H' - H^0$ as the sum of $f' = \mathbb{P}^s h'$ and $g' = \mathbb{P}^u h'$. Since $DW(H^0) = 0$, there exists a constant $b_1 > 0$ such that $\|g'\| \leq b_1\|h'\|^2$. Consider now the function $H'' = \Re_\mu(H')$. Then $h'' = H'' - H^0$ may be written as $h'' = f'' + g''$, where

$$\begin{aligned} f'' &= D\Re_\mu(H^0)f', \\ g'' &= D\Re_\mu(H^0)g' + \left[\Re_\mu(H^0 + f' + g') - D\Re_\mu(H^0)(f' + g')\right]. \end{aligned} \tag{4.29}$$

Since $\Re_\mu$ is differentiable (complex analytic), there are fixed constants b_2 and b_3 such that

$$\|f''\| \leq t'\|f'\|, \qquad \|g''\| \leq b_2\|g'\| + b_3\|f' + g'\|^2. \tag{4.30}$$

By combining this with the inequalities $\|g'\| \leq b_1\|h'\|^2$ and $\|f'\| \leq \|h'\| + b_1\|h'\|^2$, we see that there exists $r > 0$, such that $\|h''\| \leq t\|h'\|$, whenever $\|h'\| < r$. This inequality can now be iterated an arbitrary number of times. Our claim concerning the bound in (4.27) follows, if we take $B = \{H \in \mathcal{A}(\rho) \colon \|H - H^0\| < r\}$ and K sufficiently large.

To prepare for the proof of (4.28), choose $\varepsilon > 0$ such that

$$t^{-2\varepsilon}(b_2 + 9b_3)^\varepsilon < (\theta/t)^2. \tag{4.31}$$

By using Lemma 4.3, we can find constants $a > 0$ and $c > 1$, and a sequence of positive integers $k_n \leq \varepsilon n$, such that $n - k_n \geq M$ and

$$\left\|(\mathbb{I} - \mathbb{P}_{\geq 0})[D\Re_\mu(H^0)]^{k_n}\right\| \leq e^{-ac^n}, \tag{4.32}$$

whenever n is sufficiently large, say $n \geq N'$.

Given $n \geq N'$, and an arbitrary function $H = H^0 + h$ in $\mathcal{W}^s \cap B$, consider $H' = \Re_\mu^m(H)$ with $m = n - k_n$. As was shown in the first part of this proof, the norm of $h' = H' - H^0$ is bounded by $\delta = t^m\|h\| \leq t^m r$. Starting with the pair $(f', g') = (\mathbb{P}^s h', \mathbb{P}^u h')$, we now iterate the map $(f', g') \mapsto (f'', g'')$ defined by equation (4.29). The initial pair satisfies $\|f'\| \leq 2\delta$ and $\|g'\| \leq b_1\delta^2$. In order to estimate the pair $(\hat{f}, \hat{g})$ obtained after k_n steps, we can successively apply the inequalities (4.30): If $\|g'\| \leq b\delta^2$ with $\delta \leq b\delta \leq 1$, then

$\|g''\| \leq b_2 b\delta^2 + b_3(3\delta)^2 \leq (b_2 + 9b_3)b\delta^2$. Here, we have used that f' remains bounded in norm by 2δ throughout the iteration procedure. Thus, since $(b_2 + 9b_3)^{k_n} b_1 \delta \leq 1$ by our choice of ε, we find that

$$\hat{f} = \big[D\Re_\mu(H^0)\big]^{k_n} f' , \qquad \|\hat{g}\| \leq b_1 \delta^2 (b_2 + 9b_3)^{k_n} . \tag{4.33}$$

By using (4.31), (4.32), (4.33), and the fact that $H^0 + \hat{f} + \hat{g} = \Re_\mu^n(H)$, it is now straightforward to check that the function $(\mathbb{I} - \mathbb{P}_{\geq 0})\Re_\mu^n(H) = (\mathbb{I} - \mathbb{P}_{\geq 0})\hat{f} + (\mathbb{I} - \mathbb{P}_{\geq 0})\hat{g}$ satisfies the bound (4.28), if n is sufficiently large.

■

We note that, although the bound (4.28) is used in Section 5, it is not really needed to prove our main results (as it turned out).

Remark. Our definition of $\Re_\mu : B_\mu \subset \mathcal{A}(\rho) \to \mathcal{A}(\rho)$ involves many choices, such as the positive real numbers $\rho' < \varrho' < \varrho < \rho$, and the two parameters σ and μ (restricted by the choice of ρ'). All these numbers are now considered fixed. This does not, however, restrict the possible choices for the parameter μ, as long as the condition (1.9) is satisfied. In particular, if we are only interested in $|\mu| \leq s$ for some (allowed) constant s, then $\Re_\mu$ can be defined as $\mathcal{S}_{\mu/s}^{-1} \circ \Re_s$ on an open ball of radius $|\mu| b$ in $\mathcal{A}(\rho)$ centered at H^0, where $b > 0$ is determined by B_s.

5. Invariant Tori

Our construction of invariant tori involves iterates of the transformation $\Re_\mu : B \subset \mathcal{A}(\rho) \to \mathcal{A}(\rho)$. Here, $\rho > 0$ is fixed, and the parameter μ is assumed to satisfy the condition (1.15), which guarantees that the local unstable manifold $\mathcal{W}^u$ of $\Re_\mu$ at H^0 is of dimension $d - 1$. We start by introducing some notation. Recall from Definition 1.5 that $\Re_\mu(H) = [\tilde{H} - \Xi(H)]/\xi(H)$, where

$$\xi(H) = \big(\Omega \cdot \nabla \mathbb{E}\tilde{H}\big)(0) , \qquad \Xi(H) = \big(\mathbb{E}\tilde{H}\big)(0) , \tag{5.1}$$

and $\tilde{H} = (\vartheta/\mu) H \circ \mathcal{U}_H \circ T_\mu$. For any $n \geq 0$, define $H_n = \Re_\mu^n(H)$ whenever H belongs to the domain of $\Re_\mu^n$. An explicit expression for H_n is given by the equation

$$H_n = \frac{1}{\xi_n'(H)} \left(\frac{\vartheta}{\mu}\right)^n \Big[H_0 \circ \mathcal{U}_{H_0} \circ T_\mu \circ \mathcal{U}_{H_1} \circ T_\mu \circ \ldots \circ \mathcal{U}_{H_{n-1}} \circ T_\mu - \Xi_n'(H)\Big] , \tag{5.2}$$

where

$$\xi_n'(H) = \xi(H_0) \cdots \xi(H_{n-1}) , \qquad \Xi_n'(H) = \sum_{k=0}^{n-1} \xi_k'(H) \left(\frac{\vartheta}{\mu}\right)^k \Xi(H_k) . \tag{5.3}$$

If we define $V_k(H) = T_\mu^k \circ \mathcal{U}_{H_k} \circ T_\mu^{-k}$, then the relation (5.2) between H and H_n can be rewritten as follows:

$$H \circ V_0(H) \circ V_1(H) \circ \ldots \circ V_{n-1}(H) = \xi_n'(H) \left(\frac{\mu}{\vartheta}\right)^n H_n \circ T_\mu^{-n} + \Xi_n'(H) . \tag{5.4}$$

Here, and in what follows, the matrix T is assumed to satisfy all three conditions (T1-3) given in the introduction.

As we will show later, the transformations $V_k(H)$ are canonical. Thus, equation (5.4) shows that a suitable restriction of H is canonically equivalent to another Hamiltonian (the right hand side of (5.4)) that is generally much closer to H^0. The assumption here is that H belongs to the domain of $\mathfrak{R}_\mu^n$. Notice that if H_n happens to be equal to H^0, then the image of $\mathbb{T}^d \times \{0\}$ under $V_0(H) \circ \ldots \circ V_{n-1}(H)$ is an invariant torus for H.

Another important property of $V_k(H)$ is that it is (roughly speaking) independent of the parameter μ. We will prove this below, in the case where H lies on the stable manifold $\mathcal{W}^s$ of $\mathfrak{R}_\mu$. Or equivalently, we replace H by H^s, where H^s denotes the projection (in the direction of $\mathcal{A}^u$) of H onto $\mathcal{W}^s$, as defined by equation (4.23).

The positive number $\rho' < \rho$ that appears in the proposition below, and later on as well, is the one that was used to define the projections $\mathbb{I}^\pm$ and the transformations $\mathfrak{R}_\mu$; see also the remark at the end of the last section.

Proposition 5.1. *Given $z \in \mathbb{C}$ satisfying $0 < |z| \le 1$, there exists an open neighborhood B of H^0 in $\mathcal{A}(\rho)$, such that for each $k \ge 1$,*

$$T_{z\mu}^k \circ \mathcal{U}_{\mathfrak{R}_{z\mu}^k(H^s)} \circ T_{z\mu}^{-k} = T_\mu^k \circ \mathcal{U}_{\mathfrak{R}_\mu^k(H^s)} \circ T_\mu^{-k} \qquad \text{on } T_{z\mu}^k \mathcal{D}(\rho'), \tag{5.5}$$

and $\xi(\mathfrak{R}_{z\mu}^k(H^s)) = \xi(\mathfrak{R}_\mu^k(H^s))$, whenever $H \in B$.

Proof. Let $k \ge 1$ be fixed but arbitrary. If $H' \in \mathcal{A}(\rho)$ is sufficiently close (depending on k) to H^0 then, by Theorem 1.6, we have $\mathfrak{R}_{z\mu}^k(H') = \mathcal{S}_z^{-k}\mathfrak{R}_\mu(H')$, and by Lemma 3.10, this implies that $F_k(H') = 0$ on $S_z^k\mathcal{D}(\rho')$, where

$$F_k(H') = S_z^k \circ \mathcal{U}_{\mathfrak{R}_{z\mu}^k(H')} \circ S_z^{-k} - \mathcal{U}_{\mathfrak{R}_\mu^k(H')} \,. \tag{5.6}$$

Notice that the equation $F(H^s) = 0$ on $S_z^k\mathcal{D}(\rho')$ is equivalent to (5.5). The goal is to extend this identity to a neighborhood B of H^0 that is independent of k.

To this end, consider an open set $B' \ni H^0$ with the following properties: The map $H \mapsto \mathcal{U}_H$ is analytic on B', and the the local stable manifolds of $\mathfrak{R}_{z\mu}$ and $\mathfrak{R}_\mu$ coincide in B'. By Theorem 3.1 and Lemma 4.4, these conditions are satisfied if B' is sufficiently small. In addition, the map $\wp: H \mapsto H^s$ is analytic on B'. Next, we choose an open ball $B \subset B'$, centered at H^0, such that $\mathfrak{R}_\mu^n(H^s)$ and $\mathfrak{R}_{z\mu}^n(H^s)$ belong to B', for all $H \in B$, and for all $n \ge 1$. The existence of such a ball follows e.g. from the bound (4.27) in Lemma 4.7. As a consequence, $\mathfrak{R}_\mu^n \circ \wp$ and $\mathfrak{R}_{z\mu}^n \circ \wp$ are analytic as maps from B to B'.

Let now (q, p) be a fixed but arbitrary point in $S_z^k\mathcal{D}(\rho')$. Define $f_k: B \to \mathbb{C}^{2d}$ by setting $f_k(H) = (F_k(H^s))(q, p)$, for all $H \in B$. This function f, being the (difference and) composition of analytic maps, is analytic on B. But as was shown above, f is identically zero on an open neighborhood of H^0. Thus, $f = 0$ on B, and since (p, q) and k were arbitrary, the assertion regarding $V_k(H^s)$is proved.

The identity $\xi(\mathfrak{R}_{z\mu}^k(H^s)) = \xi(\mathfrak{R}_\mu^k(H^s))$ can be proved similarly, using the fact that $\xi \circ \mathcal{S}_z = \xi$ on B; see e.g. the proof of Theorem 1.6.

■

We note that both sides of equation (5.4) are defined (naturally) as functions on $T_\mu^n\mathcal{D}(\rho)$. But for simplicity, we will now restrict these functions to domains of the form $S_\mu^n\mathcal{D}(\rho_n)$, which are necessarily much smaller. Here, S_z denotes (for any $z\in\mathbb{C}$) the linear transformation on $\mathbb{C}^{2d}$, defined by the equation $S_z(q,p)=(q,zp)$.

Denote by $m(T)$ the largest of the norms $\|T\|_0\,,\|T^*\|_0\,,\|T^{-1}\|_0\,$, and $\|(T^*)^{-1}\|_0\,$. Let $\eta<1/m(T)$ and $\kappa'<\kappa$ be fixed positive real numbers, where κ is the constant that appears in the condition (1.9), and set

$$\rho_k=\eta^k\kappa'\rho\,,\qquad k=0,1,2,\ldots\,. \tag{5.7}$$

Definition 5.2. *Assume $\mu\in\mathbb{C}\backslash\{0\}$ is given. For every integer $n\ge 0$ define $A_n=S_\mu^n\mathcal{D}(\rho_n)$, and denote by $\mathcal{A}_n$ the vector space of all analytic maps G: $A_n\to\mathbb{C}^{2d}$, such that $G\circ S_\mu^n$ belongs to $\mathcal{A}^{2d}(\rho_n)$. On the space $\mathcal{A}_n$ we consider the following (equivalent) norms:*

$$\|G\|_{k,n}=\big\|S_\mu^{-k}G\circ S_\mu^n\big\|_{\rho_n}\,,\qquad G\in\mathcal{A}_n\,. \tag{5.8}$$

The next two propositions will be used to estimate the composite transformation $V_0(H^s)\circ\ldots\circ V_{n-1}(H^s)$, and some of its derivatives, on the domain $\mathcal{A}_n$.

Proposition 5.3. *If $|\mu|$ is chosen sufficiently small (but positive) then there exists a real number $a>0$, and an open neighborhood B of H^0 in $\mathcal{A}(\rho)$, such that the following holds, for every $k\ge 0$. If $H\in B$ then $V_k(H^s)$ is a canonical transformation from A_{k+1} into A_k . The map $H\mapsto V_k(H^s)-\mathrm{I}$ is analytic from B to $\mathcal{A}_{k+1}$ and satisfies the bound*

$$\|V_k(H^s)-\mathrm{I}\|_{k,k+1}\le a|b\mu|^{2k}\|H^s-H^0\|_\rho\,,\qquad H\in B\,, \tag{5.9}$$

where $b>0$ is some universal constant depending only on the matrix T.

Proof. Choose η' in $(\eta,1/m(T))$, and define $r_k=\eta'\rho_k$, for every $k\ge 0$. By Lemma 4.7, there are positive constants $b'<|\mu|^{-1}$ and $c_1(\mu)$, such that

$$\|\mathbb{I}^-H_k^s\|_\rho\le c_1(\mu)|b'\mu|^{2k}\|H^s-H^0\|_\rho\,,\qquad k=0,1,\ldots\,, \tag{5.10}$$

whenever H (and thus H^s) is sufficiently close to H^0. Here, $H_k^s=\mathfrak{R}_\mu^k(H^s)$. Furthermore, since $H\mapsto H^s$ and $\mathfrak{R}_\mu$ are analytic near H^0, the same is true for the map $H\mapsto H_k^s$. The latter is now composed with the (analytic) map ψ described in Theorem 3.1, and we get the bound

$$\|\psi(H_k^s)\|_{r_0}'\le c_2\|\mathbb{I}^-H_k^s\|_\rho\le c_1(\mu)c_2|b'\mu|^{2k}\|H^s-H^0\|_\rho\,,\qquad k\ge 0\,, \tag{5.11}$$

for H sufficiently close to H^0, where c_2 is some fixed constant. The identity

$$S_\mu^{-k}\big[V_k(H^s)-\mathrm{I}\big]\circ S_\mu^{k+1}=T_1^k\big[U_{\psi(H_k^s)}-\mathrm{I}\big]\circ T_1^{-k}S_\mu \tag{5.12}$$

and Proposition 2.5 now show that if $|\mu|>0$ is sufficiently small, then $H\mapsto V_k(H^s)-\mathrm{I}$ is well defined and analytic, as a map from some open neighborhood of H^0 in $\mathcal{A}(\rho)$ to

$\mathcal{A}_{k+1}$. Here, we have also used that $\phi \mapsto \phi \circ T_1^{-k}$ is bounded from $\mathcal{A}(\rho_1)$ to $\mathcal{A}(\rho_{k+1})$, as was shown in Proposition 3.8.

By Proposition 3.8, $S_\mu^{-k} V_k(H^s) \circ S_\mu^k$ is a canonical transformation defined on $\mathcal{D}(\rho_k)$, with generating function $\psi_k(H^s) = \psi(H_k^s) \circ T_1^{-k}$ belonging to $\mathcal{A}(r_k)$ and satisfying a bound of the form

$$\left\|\psi_k(H^s)\right\|'_{r_k} \leq c_3 |\eta|^{-k} \|\psi(H_k^s)\|'_{r_0}\,, \qquad k \geq 0\,, \tag{5.13}$$

for some constant $c_3 > 0$. Again, this holds under the assumption that $|\mu|$ is sufficiently small and H sufficiently close (depending on μ) to H^0. The bound Proposition 5.3 is now obtained by combining (5.13) with (5.11), and using Proposition 2.5.

■

We note that by keeping track of the constants in the above proof, one finds that any $b > \vartheta_1 \vartheta_d^{-2}$ will do, and that the assertion of Proposition 5.3 holds whenever $|\mu| < \eta/b$.

Proposition 5.4. *Let $0 < m < n$ and $0 < \varepsilon < \delta = \rho_m - \rho_n$. Define $B = \{G \in \mathcal{A}_n \colon \|G\|_{m,n} < \varepsilon\}$. Then $(F, G) \mapsto F \circ (\mathrm{I} + G)$ is well defined and analytic as a map from $\mathcal{A}_m \oplus B$ to $\mathcal{A}_n$, and $\|F \circ (\mathrm{I} + G)\|_{k,n} \leq \|F\|_{k,m}$, for all $F \in \mathcal{A}_m$, $G \in B$, and $k \geq 0$.*

Proof. Let $G \in B$, and define $V = S_\mu^{n-m} + S_\mu^{-m} G \circ S_\mu^n$. Then $V(q,p) = (q + Q(q,p), P(q,p))$ for all $(q,p) \in \mathcal{D}(\rho_n)$, with $Q, P \in \mathcal{A}^d(\rho_n)$ satisfying $\|Q\|_{\rho_n} < \varepsilon$ and $\|P\|_{\rho_n} < |\mu|^{n-m}\rho_n + \varepsilon$. Thus, by Proposition 2.2.d, if $F' \in \mathcal{A}(\rho_m)$, then $F' \circ V$ belongs to $\mathcal{A}(\rho_n)$, and $\|F' \circ V\|_{\rho_n} \leq \|F'\|_{\rho_m}$. In particular, if $F' = S_\mu^{-k} F \circ S_\mu^m$ with $F \in \mathcal{A}_m$, then we obtain $\|F \circ (\mathrm{I} + G)\|_{k,n} \leq \|F\|_{k,m}$, as claimed.

The above shows that the map $F' \mapsto F' \circ (S_\mu^{n-m} + G')$ is linear and bounded, and hence differentiable, from $\mathcal{A}(\rho_m)$ to $\mathcal{A}(\rho_n)$, for every G' in $B' = \{G' \in \mathcal{A}(\rho_n) \colon \|G'\|_{\rho_n} < \varepsilon\}$. And for every $F' \in \mathcal{A}(\rho_m)$, the map $G' \mapsto F' \circ (S_\mu^{n-m} + G')$ is differentiable on B' since F' is analytic. Thus, $(F', G') \mapsto F' \circ (S_\mu^{n-m} + G')$ is differentiable (complex analytic) on $\mathcal{A}(\rho_m) \oplus B'$. The assertion now follows from the fact that $F \mapsto S_\mu^{-k} F \circ S_\mu^m$ is an isometry of $\mathcal{A}_m$ (with the norm $\|.\|_{k,m}$) onto $\mathcal{A}(\rho_m)$, for every $k, m \geq 0$.

■

If H belongs to the domain of V_k, for all $k < n$, define

$$V_n'(H) = V_0(H) \circ V_1(H) \circ \ldots \circ V_{n-1}(H)\,. \tag{5.14}$$

Lemma 5.5. *If $|\mu|$ is sufficiently small (but positive) then there are real numbers $a, a' > 0$, and an open neighborhood B of H^0 in $\mathcal{A}(\rho)$, such that the following holds. If $H \in B$ and $n \geq 1$ then $V_n'(H^s) - \mathrm{I}$ belongs to $\mathcal{A}_n$, and the map $H \mapsto V_n'(H^s) - \mathrm{I}$ is analytic on B. Similarly, the functions $H \mapsto \xi_n'(H^s)$ defined by (5.3) and (4.23), are analytic on B. Furthermore,*

$$\begin{aligned} \left\|V_{n+1}'(H^s) - V_n'(H^s)\right\|_{0,n+1} &\leq a|b\mu|^{2n} \|H^s - H^0\|_\rho\,, \\ \left\|V_{n+1}'(H^s) - \mathrm{I}\right\|_{0,n+1} &\leq a' \|H^s - H^0\|_\rho\,, \end{aligned} \tag{5.15}$$

and

$$\begin{aligned} \bigl|\xi'_{n+1}(H^s) - \xi'_n(H^s)\bigr| &\le a|b\mu|^n \|H^s - H^0\|_\rho\,, \\ \bigl|\xi'_{n+1}(H^s) - 1\bigr| &\le a'\|H^s - H^0\|_\rho\,, \end{aligned} \tag{5.16}$$

for all $H \in B$ and $n \ge 1$, where $b > 0$ is some fixed constant.

Proof. By Proposition 5.3, if $|\mu| > 0$ is sufficiently small then there exists an open neighborhood B' of H^0 in $\mathcal{A}(\rho)$, and constants $a, b, c > 0$ (with b independent of μ), such that the following holds whenever $0 < \varepsilon \le c$. Let $B(\varepsilon) = \{H \in B' \colon \|H^s - H^0\|_\rho < 2\varepsilon\}$, and for every integer $k \ge 0$, define $d_k = \rho/(a|b\mu|^{2k}\varepsilon)$ and $D_k = \{z \in \mathbb{C} \colon |z| < d_k + 2\}$. Then the map $H \mapsto V_k(H^s) - \mathrm{I}$ is analytic from $B(\varepsilon)$ to $\mathcal{A}_{k+1}$, for any given $k \ge 0$, and it satisfies the bound

$$\bigl\|z[V_k(H^s) - \mathrm{I}]\bigr\|_{k,k+1} \le \rho_k(1-\eta)^2/2\,, \qquad H \in B(\varepsilon)\,,\ z \in D_k\,. \tag{5.17}$$

In what follows, we assume that $|\mu| > 0$ is "sufficiently small", and that $0 < \varepsilon \le c$.

Let $n \ge 1$ be fixed, and define $G_n(z, H) = z[V_n(H^s) - \mathrm{I}]$. Then for $k = n-1, \ldots, 1, 0$, the equation

$$G_k(z, H) = V_k(H^s) \circ \ldots \circ V_{n-1}(H^s) \circ \bigl(\mathrm{I} + z[V_n(H^s) - \mathrm{I}]\bigr) - \mathrm{I} \tag{5.18}$$

defines an analytic map G_k from $D_n \times B(\varepsilon)$ to $\mathcal{A}_{n+1}$, satisfying

$$\|G_k(z, H)\|_{k,n+1} \le \|V_k(H^s) - \mathrm{I}\|_{k,k+1} + \|G_{k+1}(z, H)\|_{k+1,n+1}\,, \tag{5.19}$$

for all $z \in D_n$ and $H \in B(\varepsilon)$. This follows by inductively applying Proposition 5.4, using the identity $G_k(z, H) = [V_k(H^s) - \mathrm{I}] \circ [\mathrm{I} + G_{k+1}(z, H)] + G_{k+1}(z, H)$. The inductive estimate on G_k, obtained from (5.19) and (5.17), and valid for all $z \in D_n$, is $\|G_k(z, H)\|_{k,n+1} \le \rho_k(1-\eta)/2$. Here, and in what follows, we assume that $H \in B(\varepsilon)$. If $z = 1$, then we can use (5.9) instead of (5.17), and the resulting bound on the norm of $G_0(1, H)$ implies the second inequality in (5.15). For a general $z \in D_n$, an upper bound on $\|G_0(z, H)\|_{0,n+1}$ is given by the number ρ. Thus, by using Cauchy's formula

$$\frac{\partial}{\partial z} G_0(z, H) = \frac{1}{2\pi i} \oint\limits_{|\zeta| = d_n + 1} \frac{d\zeta}{(\zeta - z)^2}\, G_0(\zeta, H)\,, \qquad |z| \le 1\,, \tag{5.20}$$

and the mean value theorem, we obtain the bound

$$\begin{aligned} \bigl\|V'_{n+1}(H^s) - V'_n(H^s)\bigr\|_{0,n+1} &= \|G_0(1, H) - G_0(0, H)\|_{0,n+1} \\ &\le \rho/d_n = a|b\mu|^{2n}\varepsilon\,. \end{aligned} \tag{5.21}$$

Let now H be a fixed but arbitrary function in $B(c)$. If $H^s \ne H^0$ then the first inequality in (5.15) is obtained from (5.21) by choosing $\varepsilon = \|H^s - H^0\|_\rho$. The case $H^s = H^0$ is trivial, since $H^s = H^0$ implies that $V_k(H^s) = \mathrm{I}$ for all $k \ge 0$.

The analyticity of the functions $H \mapsto \xi_n'(H^s)$, in an open neighborhood $B \subset B(c)$ of H^0, follows from the analyticity of $H \mapsto H^s$ and $\Re_\mu$. And (5.16) is a consequence of the bound (4.27) in Lemma 4.7, and of the fact that $\xi(H^0) = 1$.

■

The following is an approximate version of equation (1.18).

Proposition 5.6. *If $|\mu| > 0$ is sufficiently small then there exists a real number $a > 0$, and an open neighborhood B of H^0 in $\mathcal{A}(\rho)$, such that*

$$\big\|\{\mathrm{I}, H\} \circ V_n'(H) - \xi_n'(H)\{V_n'(H), H^0\}\big\|_{0,n} \le a|b\mu|^n \big\|H - H^0\big\|_\rho\,, \tag{5.22}$$

for all $H \in \mathcal{W}^s \cap B$ and $n \ge 1$, where $b > 0$ is some fixed constant.

Proof. First, we note that for each $n \ge 1$, if $r_n = \rho_n$ then

$$\big\|V_n'(H) \circ S_\mu^n\big\|_{r_n} \le 2\,, \qquad \big\|\big[H_n - H^0\big] \circ T_1^{-n}\big\|_{r_n} \le \big\|H_n - H^0\big\|_\rho\,, \tag{5.23}$$

for all $H \in \mathcal{W}^s$ sufficiently close to H^0. These two bounds follow from Lemma 5.5 and Proposition 3.8. The same is true if we set $r_n = (1+\varepsilon)\rho_n$, with $\varepsilon > 0$ sufficiently small, since the positive constants $\eta < 1/m(T)$ and $\kappa' < \kappa$ in the definition (5.7) can be chosen arbitrarily. Thus, by Proposition 2.2.b, we have

$$\big\|V_n'(H) \circ S_\mu^n\big\|_{\rho_n}' \le \frac{2}{\varepsilon\rho_n}\,, \qquad \big\|\big[H_n - H^0\big] \circ T_1^{-n}\big\|_{\rho_n}' \le \frac{1}{\varepsilon\rho_n}\big\|H_n - H^0\big\|_\rho\,, \tag{5.24}$$

where $\|.\|_{\rho_n}'$ is the norm described in Definition 2.3. To be more precise, this holds for all n, and for all H in some (sufficiently small) open neighborhood of H^0 in $\mathcal{A}(\rho)$. The same comment applies to the statements made below, but it will not always be mentioned.

By Proposition 5.3, $V_n'(H)$ is a canonical transformation from A_n into $A_0 = \mathcal{D}(\rho_0)$, and the bound (5.15) shows that this transformation is one-to-one. As a consequence, we have $\{f, g\} \circ V_n'(H) = \{f \circ V_n'(H), g \circ V_n'(H)\}$, whenever f and g are analytic on $\mathcal{D}(\rho_0)$. This, together with (5.4), implies that

$$\{\mathrm{I}, H\} \circ V_n'(H) = \big\{V_n'(H), H \circ V_n'(H)\big\} = \xi'(H)\left(\frac{\mu}{\vartheta}\right)^n \big\{V_n'(H), H_n \circ T_\mu^{-n}\big\}\,. \tag{5.25}$$

By using first the identities $(\mu/\vartheta)H^0 \circ T_\mu^{-1} = H^0$ and $\{\mathcal{S}_z^{-1}f, \mathcal{S}_z^{-1}g\} = \mathcal{S}_z^{-1}\{f, g\}$, and then the bound (5.24), the left hand side of (5.22) can now be estimated as follows:

$$\begin{aligned}
&\big\|\{\mathrm{I}, H\} \circ V_n'(H) - \xi_n'(H)\{V_n'(H), H^0\}\big\|_{0,n} \\
&\qquad = \big\|\xi_n'(H)(\mu/\vartheta)^n\big\{V_n'(H), \big[H_n - H^0\big] \circ T_\mu^{-n}\big\}\big\|_{0,n} \\
&\qquad = \xi_n'(H)|\vartheta|^{-n}\big\|\mu^n\big\{V_n'(H), \big[H_n - H^0\big] \circ T_\mu^{-n}\big\} \circ S_\mu^n\big\|_{\rho_n} \\
&\qquad = \xi_n'(H)|\vartheta|^{-n}\big\|\big\{V_n'(H) \circ S_\mu^n, \big[H_n - H^0\big] \circ T_1^{-n}\big\}\big\|_{\rho_n} \\
&\qquad \le \xi_n'(H)|\vartheta|^{-n}4d(\varepsilon\rho_n)^{-2}\big\|H_n - H^0\big\|_\rho\,.
\end{aligned} \tag{5.26}$$

This implies (5.22), if we substitute for $\|H_n - H^0\|_\rho$ the bound given in Lemma 4.7.

$\blacksquare$

In order to translate between the two different notations used in (5.22) and (1.18), let us write $F = (F_1, F_2)$, if $F \in \mathcal{A}_n$ and $F(q,p) = (F_1(q,p), F_2(q,p))$, for all (q,p) in the domain of A_n. In particular, $\mathrm{I} = (\mathrm{I}_1, \mathrm{I}_2)$, where $\mathrm{I}_1(q,p) = q$ and $\mathrm{I}_2(q,p) = p$ for all q and p. Then

$$\begin{aligned} \{\mathrm{I}, F\} &= \bigl\{(\mathrm{I}_1, \mathrm{I}_2), F\bigr\} = \bigl(\{\mathrm{I}_1, F\}, \{\mathrm{I}_2, F\}\bigr) = \bigl(\nabla_2 F, -\nabla_1 F\bigr) = \mathbb{J}\nabla F\,, \\ \{F, H^0\} &= \{F, \omega\cdot \mathrm{I}_2\} = \omega\cdot\{F, \mathrm{I}_2\} = \omega\cdot\nabla_1 F\,, \end{aligned} \tag{5.27}$$

where $\mathbb{J}(q,p) = (p, -q)$.

Now we are ready to take the limit of the sequence $\{V_n'(H)\}$, considered as C^ℓ functions on the torus $\mathbb{T}^d$. The limit $V'(H)$ is essentially the function Γ_H described in Theorem 1.7, restricted to real arguments. Later on, we will show that this function can be analytically continued to a complex neighborhood of the torus.

Definition 5.7. *For every integer $\ell \ge 0$, denote by $C^\ell(\mathbb{T}^d\times\{0\}, \mathbb{C}^{2d})$, or $C^\ell(\mathbb{T}^d)$ for short, the vector space of all ℓ–times continuously differentiable functions from $\mathbb{R}^{2d}$ to $\mathbb{C}^{2d}$ that are 2π–periodic in the first d variables (when considered as functions of $2d$ real arguments) and constant in the last d variables. When equipped with the norm*

$$\|g\|_{(\ell)} = \max_{|\alpha|\le\ell}\sup_{q\in\mathbb{R}^d}\bigl|\bigl(\partial^\alpha g\bigr)(q,0)\bigr|\,, \qquad g\in C^\ell\,, \tag{5.28}$$

$C^\ell(\mathbb{T}^d)$ is a Banach space. Here, we have used the notation $\partial^\alpha = \partial_1^{\alpha_1}\partial_2^{\alpha_2}\cdots\partial_{2d}^{\alpha_{2d}}$.

Notice that, by definition, if $g\in C^\ell(\mathbb{T}^d)$ then $g(q,p) = g(q,0)$ for all q. The extra variable $p\in\mathbb{R}^d$ has been introduced (kept) only to simplify notation. In particular, if $G\in\mathcal{A}_n$, then the restriction of $G\circ S_0$ to $\mathbb{R}^{2d}$, which we will also denote by $G\circ S_0$, belongs to $C^\ell(\mathbb{T}^d)$, for every ℓ. Recall that $S_0(q,p) = (q,0)$. The map $G\mapsto G\circ S_0$ is continuous as a linear operator from $\mathcal{A}_n$ to C^ℓ, and it satisfies the bound

$$\|G\circ S_0\|_{(\ell)} \le \frac{\ell!}{\rho_n^\ell}\sup_{x\in A_n}|G(x)| \le K_\ell\,\eta^{-\ell n}\|G\|_{0,n}\,, \qquad G\in\mathcal{A}_n\,, \tag{5.29}$$

for some constant K_ℓ that is independent of n. The second inequality in (5.29) is obtained by applying Proposition 2.2.a to the function $G\circ S_\mu^n$ in $\mathcal{A}(\rho_n)$.

Proposition 5.8. *Given $\ell\ge 1$, there exists a constant $a_\ell > 0$ and an open neighborhood B_ℓ of H^0 in $\mathcal{A}(\rho)$, such that the following holds. For every $H\in B_\ell$, the two limits in the equation*

$$V'(H) = S_0 + \lim_{n\to\infty}\bigl[V_n'(H) - \mathrm{I}\bigr]\circ S_0\,, \qquad \xi'(H) = \lim_{n\to\infty}\xi_n'(H)\,, \tag{5.30}$$

converge in $C^\ell(\mathbb{T}^d)$ and $\mathbb{C}$, respectively, and the convergence is uniform on B_ℓ. The functions $H\mapsto V'(H) - S_0$ and $H\mapsto\xi'(H)$ are analytic on B_ℓ, and they satisfy the bounds

$$\bigl\|V'(H) - S_0\bigr\|_{(\ell)} \le a_\ell\|H^s - H^0\|_\rho\,, \qquad \bigl|\xi'(H) - 1\bigr| \le a_\ell\|H^s - H^0\|_\rho\,, \tag{5.31}$$

for all $H \in B_\ell$. Furthermore, if $H \in \mathcal{W}^s \cap B_\ell$ then

$$\big(\mathbb{J}\nabla H\big) \circ V'(H) = \xi'(H)\omega \cdot \nabla_1 V'(H)\,, \tag{5.32}$$

and if in addition, $|\mu| > 0$ is sufficiently small, then

$$V'(H) = V'_k(H) \circ T^k_\mu V'\big(\Re^k_\mu(H)\big) \circ T^{-k}_\mu\,, \qquad k = 1, 2, \ldots\,. \tag{5.33}$$

Proof. First we prove the assertions under the additional assumption that $|\mu| > 0$ be sufficiently small. In this case, the uniform convergence (and hence also the analyticity) on some open neighborhood of H^0 of the two limits in (5.30), as well as the bounds (5.31), follow directly from Lemma 5.5 and inequality (5.29). Thus, $X_n = \xi'_n(H)\omega \cdot \nabla_1 V'_n(H) \circ S_0$ converges in $C^{\ell-1}(\mathbb{T}^d)$ to $X = \xi'(H)\omega \cdot \nabla_1 V'(H)$, and by the chain rule, $Y_n = (\mathbb{J}\nabla H) \circ V'_n(H) \circ S_0$ converges in $C^\ell(\mathbb{T}^d)$ to $Y = (\mathbb{J}\nabla H) \circ V'(H)$. Consider now $H \in \mathcal{W}^s$ sufficiently close to H^0 for Proposition 5.6 to apply. Then $Y_n - X_n \to 0$ by (5.22) and (5.29). Consequently, $Y = X$, as claimed in (5.32).

In what follows, we assume that $H \in \mathcal{W}^s \cap B$ where B is some suitable open neighborhood of H^0 (sufficiently small, as needed by the propositions used). We may also assume that $\mathcal{W}^s \cap B$ is mapped into itself by $\Re_\mu$. Then $V'_{n-k}(H_k) \to V'(H_k)$ as $n \to \infty$, for all $k \geq 1$. Furthermore, by Proposition 5.8 and Lemma 4.7, there are constants $a, b > 0$, with b independent of μ, such that

$$\|V'_{n-k}(H_k) \circ S_0 - S_0\|_{(0)}\,,\ \|V'(H_k) - S_0\|_{(0)} \leq a|b\mu|^k \|H - H^0\|_\rho\,, \tag{5.34}$$

whenever $n > k \geq 1$. Let now $k \geq 1$ be fixed. ¿From the definition of $V'_1, V'_2, \ldots$, it follows that

$$V'_n(H) = V'_k(H) \circ T^k_\mu V'_{n-k}(H_k) \circ T^{-k}_\mu\,, \tag{5.35}$$

for all $n > k$, and in particular,

$$V'_n(H) \circ S_0 = V'_k(H) \circ S^k_\mu F_n\,, \qquad F_n = T^k_1 V'_{n-k}(H_k) \circ S_0 \circ T^{-k}_1\,. \tag{5.36}$$

The goal now is to take $n \to \infty$. Since $[V'_{n-k}(H_k) - \mathrm{I}] \circ S_0$ converges to $V'(H_k) - S_0$ in $C^0(\mathbb{T}^d)$, we have $F_n \to F = T^k_1 \circ V'(H_k) \circ T^{-k}_1$ pointwise on $\mathbb{R}^{2d}$. And from the bound (5.34) we see that if $|\mu| > 0$ is chosen sufficiently small (as assumed), then the ranges of $F_1, F_2, \ldots$ and F are contained in the domain $\mathcal{D}(\rho_k)$ where $V'_k(H) \circ S^k_\mu$ is analytic. This shows that $V'_k(H) \circ S^k_\mu F_n$ converges to $V'_k(H) \circ S^k_\mu F$, pointwise on $\mathbb{R}^{2d}$. Thus, since $V'_n(H) \circ S_0 \to V'(H)$, we conclude from (5.36) that $V'(H) = V'_k(H) \circ S^k_\mu F$, which is equivalent to (5.33).

In order to complete the proof, we need to eliminate the smallness condition on μ. Let μ_0 be a nonzero complex number such that $\mu = \mu_0$ satisfies the condition (1.15). As was shown above, there exists $z \neq 0$ of modulus ≤ 1, such that for the choice $\mu = z\mu_0$, the limits in (5.30) converge in $C^\ell(\mathbb{T}^d)$ and $\mathbb{C}$, respectively, for every $H \in B'$. But by Proposition 5.1 (which also uses Lemma 4.4), the two choices $\mu = \mu_0$ and $\mu = z\mu_0$ lead to the same functions $V'_n(H) \circ S_0$, and thus to the same limit $V'(H)$, for all H in some open set $B_\ell \subset B$ containing H^0. And by Lemma 4.4, the condition $H \in \mathcal{W}^s \cap B_\ell$ needed for (5.32) is independent of the choice (μ_0 or $z\mu_0$) of μ, if B_ℓ is sufficiently small. (The identity (5.33) could also be extended from $\mu = z\mu_0$ to $\mu = \mu_0$, but this will not be needed later.)

■

Consider $H \in \mathcal{W}^s \cap B_1$. Then for every positive integer ℓ, the function $H_k = \Re_\mu^k(H)$ belongs to B_ℓ, if k is sufficiently large. Thus, the identity (5.33) suggests that $V'(H)$ is in fact of class C^ℓ. And since $\ell > 0$ was arbitrary, ...

Lemma 5.9. *The map $V': B_1 \to C^1(\mathbb{T}^d)$, described in Proposition 5.8 has the following property: If $H \in B_1$ then $V'(H)$ is of class C^∞, and for every $\ell \geq 0$, V' is analytic as a map from B_1 to $C^\ell(\mathbb{T}^d)$.*

Proof. Let $\ell \geq 2$ be fixed but arbitrary. As in the proof of Proposition 5.8, we can (will) choose $|\mu| > 0$ as small as needed, without loosing generality. Let $m \geq 1$ be such that $\Re_\mu^m(H)$ belongs to $B = B_1 \cap B_\ell$ whenever $H \in \mathcal{W}^s \cap B_1$. The existence of such an integer m follows e.g. from the bound (4.27) in Lemma 4.7. By Proposition 5.8 and Lemma 4.7, there are real numbers $a, a' > 0$, and a universal constant $b > 0$, such that

$$\big\|V'(H_k^s) - S_0\big\|_{(\ell)} \leq a\big\|H_k^s - H_0\big\|_\rho \leq a'|b\mu|^k \|H - H^0\|_\rho\,, \tag{5.37}$$

for all $H \in B_1$ and $k \geq m$, where $H_k^s = \Re_\mu^k(H^s)$. This bound shows that if $|\mu| > 0$ has been chosen sufficiently small, then there exists $k \geq m$ (sufficiently large) such that the range of $F(H) = T_1^k V'(H_k^s) \circ T_1^{-k}$ is contained in $\mathcal{D}(\rho_k)$, for all $H \in B_1$. With this choice of k, define $G(H) = [V_k'(H) - \mathrm{I}] \circ S_\mu^k$. Then

$$\begin{aligned} V'(H) = V'(H^s) &= \big[S_\mu^k + G(H)\big] \circ F(H) \\ &= \mathrm{I} + S_\mu^k\big[F(H) - S_0\big] + G(H) \circ F(H)\,, \end{aligned} \tag{5.38}$$

for all $H \in B_1$, as was shown in Proposition 5.8. Since $G(H): \mathcal{D}(\rho_k) \to \mathbb{C}^{2d}$ is analytic and $F(H): \mathbb{R}^{2d} \to \mathcal{D}(\rho_k)$ is of class C^ℓ, equation (5.38) shows that $V'(H)$ is of class C^ℓ.

By Proposition 5.6, $G: B_1 \to \mathcal{A}^{2d}(\rho_k)$ is analytic. Furthermore, $H \mapsto [F(H) - \mathrm{I}]$ is analytic as a map from B_1 to $C^\ell(\mathbb{T}^d)$, since it is the composition of $H \in B_1 \mapsto H_k^s \in B$ with $H \in B \mapsto [V'(H^s) - S_0] \in C^\ell(\mathbb{T}^d)$, and both of these maps are analytic (the second one by Proposition 5.8). By using the last expression for $V'(H)$ in (5.38), it is now straightforward to show that $H \mapsto [V'(H) - S_0]$ is analytic as a map from B_1 to $C^\ell(\mathbb{T}^d)$, as claimed.

■

If $H \in \mathcal{W}^s$ has some "excess analyticity" then the relationship between $V'(H)$ and $V'(\mathcal{J}_q^{-1}H)$, valid a priori only for real q's, can be used to analytically continue $V'(H)$ away from $\mathbb{R}^{2d}$.

Given two sets X, Y, and a point $x \in X$, define the evaluation functional $\mathcal{E}_x$ on Y^X by setting $\mathcal{E}_x f = f(x)$, for all $f \in Y^X$.

Lemma 5.10. *Let $r - \rho > \varepsilon > 0$. Then there exists an open neighborhood B of H^0 in $\mathcal{A}(r)$ such that the following holds. For every $H \in B$, the equation*

$$\mathcal{V}_H(q,p) = \mathcal{E}_0 J_q \circ V'\bigl(\mathcal{J}_q^{-1} H\bigr), \qquad q \in \mathcal{D}_1(\varepsilon),\ p \in \mathbb{C}^d, \tag{5.39}$$

defines an analytic function $\mathcal{V}_H$ on $\mathcal{D}_1(\varepsilon) \times \mathbb{C}^d$ whose restriction to $\mathbb{R}^{2d}$ coincides with $V'(H)$. Furthermore, $\mathcal{V}_H - S_0$ belongs to $\mathcal{A}^{2d}(\varepsilon)$, for every $H \in B$, and $H \mapsto \mathcal{V}_H - S_0$ is analytic as a map from $\mathcal{A}(r)$ to $\mathcal{A}^{2d}(\varepsilon)$.

Proof. Choose δ in $(\varepsilon, r-\rho)$. Clearly, $(q,H) \mapsto \mathcal{J}_q^{-1} H$ is differentiable (analytic) from $\mathcal{D}_1(\delta) \times \mathcal{A}(r)$ to $\mathcal{A}(\rho)$. This follows from the fact that $\mathcal{J}_q^{-1}\colon \mathcal{A}(\rho+\delta) \to \mathcal{A}(\rho)$ is bounded (of norm 1) for all $q \in \mathcal{D}_1(\delta)$, and that differentiation is bounded from $\mathcal{A}(r)$ to $\mathcal{A}(\rho+\delta)$. Define $G\colon B_1 \to \mathcal{C}^1(\mathbb{T}^d)$ by setting $G(H) = V'(H) - S_0$, for all $H \in B_1$; see Proposition 5.8. The boundedness of $\mathcal{J}_q^{-1}\colon \mathcal{A}(r) \to \mathcal{A}(\rho)$ also implies that we can find $B \ni H^0$ open in $\mathcal{A}(r)$ such that both H and $\mathcal{J}_q^{-1} H$ belong to the domain B_1 of the map V'. Thus, since G is analytic on B_1, and since $\mathcal{E}_0$ is bounded on $C^1(\mathbb{T}^d)$, we conclude that $(q,H) \mapsto E_0 G(\mathcal{J}_q^{-1} H)$ is analytic on $\mathcal{D}_1(\delta) \times B$. The analyticity of $\mathcal{V}_H$, for $H \in B$, follows from the identity $[\mathcal{V}_H - S_0](q,p) = E_0 G(\mathcal{J}_q^{-1} H)$.

Let $\varepsilon < \delta' < \delta$. Since $\mathcal{J}_q^{-1} H$ is 2π-periodic in each component of q, the same is true for $\mathcal{E}_0 G(\mathcal{J}_q^{-1} H)$. Thus, $\mathcal{V}_H - S_0$ belongs to $\mathcal{A}^{2d}(\delta')$, for all $H \in B$. Consider now the sup-norm on $\mathcal{A}^{2d}(\delta')$. Then the analyticity of $H \mapsto \mathcal{V}_H - S_0$, as a map from B to $\mathcal{A}^{2d}(\delta')$, follows from the fact that all derivatives with respect to the second argument of $(q,H) \mapsto \mathcal{E}_0 G(\mathcal{J}_q^{-1} H)$, at any $H \in B$, are bounded uniformly in the first argument, $q \in \mathcal{D}_1(\delta')$.

Consider now a fixed but arbitrary $q \in \mathbb{R}^d$, and assume that $H \in \mathcal{W}^s$. Recall that $\mathcal{J}_q^{-1}$ is an isometry on $\mathcal{A}(\rho)$. Thus, given any open neighborhood B'' of H^0 in $\mathcal{A}(\rho)$, if $\|H - H^0\|_\rho$ is sufficiently small, then the functions $H' = \mathcal{J}_q^{-1} H$, $H'_k = \mathfrak{R}_\mu^k(H')$, and $H_k = \mathfrak{R}_\mu^k(H)$ belong to B'', for all $k \geq 0$. Here, we have also used Lemma 4.6, and the bound (4.27) in Lemma 4.7. By using Theorem 1.6 and Lemma 3.10, together with the identity $T_\mu^k \circ J_{T^{-k}q} = J_q \circ T_\mu^k$, we first obtain $H'_k = \mathcal{J}_{T^{-k}q}^{-1} H_k$, and then $V_k(H') = J_q^{-1} \circ V_k(H) \circ J_q$, for all $k \geq 0$. This holds for all H in some open neighborhood B' of H^0 in $\mathcal{A}(\rho)$, that is independent of $q \in \mathbb{R}^d$. The conclusion is that $V'\bigl(\mathcal{J}_q^{-1} H\bigr) = J_q^{-1} \circ V'(H) \circ J_q$, and thus

$$\mathcal{E}_{(q,p)} V'(H) = \mathcal{E}_0\bigl(V'(H) \circ J_q\bigr) = \mathcal{E}_0 J_q \circ V'\bigl(\mathcal{J}_q^{-1} H\bigr) = \mathcal{V}_H(q,p), \tag{5.40}$$

for all $q \in \mathbb{R}^d$ and $H \in B'$. This shows that $\mathcal{V}_H$ and $V'(H)$ agree on $\mathbb{R}^d$, as claimed.

■

Proof of Theorem 1.7. Define $\Gamma_H(q) = \mathcal{V}_H(q,0)$, for all $q \in \mathcal{D}_1(r')$, and for all H in the domain of the map $H \mapsto \mathcal{V}_H$ described in Lemma 5.10 (with $\varepsilon = r'$). Consider first a function $H \in \mathcal{W}^s$ in this domain, with the property that $H(q,p)$ only depends on p, or equivalently, that $\mathcal{J}_q^{-1} H = H$ for all q. In this case, $\xi'(H) = \xi(H)$ and $V'(H) = S_0$, as can be verified form the definitions. Thus, $\xi'(H) = (\Omega \cdot \nabla \mathbb{E} H)(0)$ and $\Gamma_{H^0}(u) = E_0 J_q \circ S_0 = (q,0)$, for all $q \in \mathcal{D}_1(r')$, as claimed.

The asserted analyticity properties of the functions Γ_H, and of the map $H \mapsto \Gamma_H - \Gamma_{H^0}$ follow directly from Lemma 5.10. (The Banach space of analytic 2π-periodic functions, mentioned in Theorem 1.7, consists of all functions g on $\mathcal{D}_1(r')$ such that $G\colon (q,p) \mapsto g(q)$ belongs to $\mathcal{A}(r')$; in particular, if $g = \Gamma_H$ then $G = \mathcal{V}_H$.) Since Γ_{H^0} maps $\mathcal{D}_1(r')$ into $\mathcal{D}(r')$, the analyticity of $H \mapsto \Gamma_H - \Gamma_{H^0}$ also implies that Γ_H maps $\mathcal{D}_1(r')$ into $\mathcal{D}(r)$, whenever $\|H - H^0\|_{r'}$ is sufficiently small (which we assume).

As was shown in Lemma 5.10, the restriction of $\mathcal{V}_H$ to $\mathbb{R}^d$ is equal to $V'(H)$, Thus, by Proposition 5.8, if $H \in \mathcal{W}^s$ is sufficiently close to H^0, then the function Γ_H satisfies the equation $(\mathbb{J}\nabla H) \circ \Gamma_H = \xi'(H)\omega \cdot \nabla\Gamma_H$ on $\mathbb{R}^d$. But since both sides of this equation are analytic functions on $\mathcal{D}_1(r')$, the same equation holds on $\mathcal{D}_1(r')$. This proves equation (1.18). The analyticity of the function ξ' was shown in Proposition 5.8.

∎

Proof of Theorem 1.9. Assume that the hypotheses of the theorem are satisfied, including the conditions on ∇f. Let E_H be a $d-1$ dimensional subspace of $\mathbb{C}^d$ that is mapped onto E by $(D\nabla f)(0)$, and for every $\varepsilon > 0$, define $G_\varepsilon = \{v \in E_H\colon \|v\| < \varepsilon\}$.

Choose a positive real number $2\varepsilon < r'' - r$, and consider the map $(h,v) \mapsto (H+h) \circ R_v$ from $\mathcal{A}(r''-\varepsilon) \times G_\varepsilon$ to $\mathcal{A}(r)$. This map is analytic, as can be seen e.g. from Proposition 2.2. Notice also that $H^0 \circ R_v = 0$ for any v. Thus, if $\|H - H^0\|_{r''-\varepsilon}$ is sufficiently small (as we assume here), then $(H+h) \circ R_v$ belongs to the domain of $\Gamma : H' \mapsto \Gamma_{H'}$, for all $v \in G_\varepsilon$, and for all h in some open neighborhood B'' of 0 in $\mathcal{A}(r'' - \varepsilon)$. In this case, $(h,v) \mapsto (H+h) \circ R_v$ can be composed with the function $\mathcal{F} = \mathbb{P}^u - W \circ \mathbb{P}^s$, which takes values in $\mathcal{A}^u$ and is analytic on the domain of Γ; see Section 4. The resulting function $(h,v) \mapsto F_h(v) = \mathcal{F}((H+h) \circ R_v)$ is analytic on $B'' \times G_\varepsilon$. And it vanishes precisely when $(H+h) \circ R_v$ belongs to $\mathcal{W}^s$. This follows from the fact that $\mathcal{W}^s$ is the graph of W. Thus, part of our goal reduces to solving the equation $F_h(v) = 0$, for a given function h.

By using equation (4.22), together with the fact that $\mathbb{P}^s(H \circ R_v)$ lies on $\mathcal{W}^s$ (see Corollary 4.5), we find that

$$DF_0(v)w = -\sum_{j=2}^{d} \Omega^{(j)} \cdot \bigl(D\nabla f\bigr)(-v)w\, h_j\,, \qquad w \in E_H\,, \tag{5.41}$$

for all $v \in G_\varepsilon$. Given that $(D\nabla f)(0)$ maps E_H onto E, this equation shows that the map $DF_0(0)\colon E_H \to \mathcal{A}^u$ is invertible. Furthermore, $F_0(0) = 0$, since $H \in \mathcal{W}^s$. Thus, by the implicit function theorem, there exists an open neighborhood $B' \subset B''$ of the origin in $\mathcal{A}(r'' - \varepsilon)$, and an analytic function $v'\colon B' \to G_\varepsilon$, such that $v'(0) = 0$ and $F_h(v'(h)) = 0$, for all $h \in B'$. If we define $v(H+h) = v'(h)$, then the maps $H' \mapsto v(H') \in G_\varepsilon$ and $H' \mapsto H'_{v(H')} \in \mathcal{A}(r)$ are analytic on $B'_H = \{H' \in \mathcal{A}(r'' - \varepsilon)\colon H' - H \in B'\}$, and $H'_{v(H')}$ lies on $\mathcal{W}^s$, for all $H' \in B'_H$.

If $(D\nabla f)(0)$ is non-singular, define $w \in \mathbb{C}^d$ by the equation $(D\nabla f)(0)w = \omega$; otherwise, let w be some arbitrary vector in $\mathbb{C}^d \setminus E_H$. Then there exists an open neighborhood $A \times B_H$ of $(0,H)$ in $\mathbb{C} \oplus \mathcal{A}(r'')$, such that $(s,H') \mapsto \Upsilon_{H'}(s) = H'_{sw+v(H'_{sw})}$ is well defined and analytic as a map from $A \times B_H$ to $\mathcal{A}(r)$. By decreasing the size of A, if necessary, all derivatives of this map can be bounded uniformly in s, which makes $H' \mapsto \Upsilon_{H'}$ an analytic family of curves in $C^1(A, \mathcal{A}(r))$. Assume now that $(D\nabla f)(0)$ is non-singular. Then an explicit calculation, using the expression for $\xi'(H)$ given in Theorem 1.7, shows that the derivative of $\xi' \circ \Upsilon_H$ at zero is -1. Thus, if A and B_H are chosen sufficiently small, then the derivative of $\xi' \circ \Upsilon_{H'}$ is non-vanishing on A, for all $H' \in B_H$.

$\blacksquare$

References

[1] A.N. Kolmogorov, *On Conservation of Conditionally Periodic Motions Under Small Perturbations of the Hamiltonian.* Dokl. Akad. Nauk SSSR, **98**, 527–530 (1954).

[2] V.I. Arnold, *Proof of A.N. Kolmogorov's Theorem on the Preservation of Quasi–Periodic Motions under Small Perturbations of the Hamiltonian.* Usp. Mat. Nauk, **18**, No. 5, 13–40 (1963). Russ. Math. Surv., **18**, No. 5, 9–36 (1963).

[3] J. Moser, *On Invariant Curves of Area–Preserving Mappings of an Annulus.* Nachr. Akad. Wiss. Gött., II. Math. Phys. Kl 1962, 1-20 (1962).

[4] R. de la Llave, *Introduction to KAM Theory.* Preprint U. Texas, mp_arc 93–8 (1993).

[5] L.H. Eliasson, *Absolutely Convergent Series Expansions for Quasi Periodic Motions.* Report 2–88, Dept. of Mathematics, University of Stockholm (1988). Math. Phys. EJ, **2**, No. 4, 33pp (1996).

[6] D.F. Escande, 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**, 257–284 (1981).

[7] R.S. McKay, *Renormalisation in Area Preserving Maps.* Thesis, Princeton (1982). World Scientific, London (1993).

[8] A. Mehr, D.F. Escande, *Destruction of KAM Tori in Hamiltonian Systems: Link with the Destabilization of nearby Cycles and Calculation of Residues.* Physica, **13D**, 302–338 (1984).

[9] S. Tompaidis, *Numerical Study of Invariant Sets of a Volume–Preserving Map.* Preprint U. de Rennes, mp_arc 95–441 (1995).

[10] 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, 93–118 (1986).

[11] Ya.G. Sinai, K. Khanin, *Renormalization Group Methods in the Theory of Dynamical Systems.* Int. J. Mod. Phys., **B2**, 147–165 (1988).

[12] D. Kosygin, *Multidimensional KAM Theory from the Renormalization Group Viewpoint.* In "Dynamical Systems and Statistical Mechanics", Ya.G. Sinai (ed), AMS, Adv. Sov. Math., **3**, 99–129 (1991).

[13] J. Feldman, E. Trubowitz, *Renormalization in Classical Mechanics and Many Body Quantum Field Theory.* Journal d'Analyse Mathématique, **58**, 213–247 (1992).

[14] G. Gallavotti, *Perturbation Theory.* In "Mathematical Physics towards the XXI Century", R. Sen, A. Gersten (eds), Ben Gurion University Press, Ber Sheva, 275–294 (1994).

[15] G. Gentile, V. Mastropietro, *Tree Expansion and Multiscale Decomposition for KAM Tori.* Nonlinearity, **8**, 1–20 (1995).

[16] G. Gallavotti, G. Gentile, V. Mastropietro, *Field theory and KAM tori.* Math. Phys. EJ, **1**, No. 5, 13pp (1995).

[17] M.J. Feigenbaum, L.P. Kadanoff, S.J. Shenker, *Quasi–Periodicity in Dissipative Systems: A Renormalization Group Analysis.* Physica, **5D**, 370–386 (1982).

[18] S. Ostlund, D.A. Rand, J. Sethna, E. Siggia, *Universal Properties of the Transition from Quasiperiodicity to Chaos in Dissipative Systems.* Physica, **8D**, 303–342 (1983).

[19] D.A. Rand, *Existence, Non–Existence and Universal Breakup of Dissipative Golden Invariant Tori.* Nonlinearity, **5**, 639–706 (1992).

[20] E. Hille, R.S. Phillips, *Functional Analysis and Semi–Groups.* AMS Colloquium Publications, **31** (1957).

[21] M.J. Bertin, A. Decomps–Guilloux, M. Grandet–Hugot, M. Pathiaux–Delefosse, J.P. Schreiber, *Pisot and Salem Numbers.* Birkhäuser Verlag, Basel · Boston · Berlin (1992).

[22] J.J. Abad, *work in progress.*