CONSTRUCTION OF WHISKERS FOR THE QUASIPERIODICALLY FORCED PENDULUM

MIKKO STENLUND

Abstract. We study a Hamiltonian describing a pendulum coupled with several anisochronous oscillators, giving a simple construction of unstable KAM tori and their stable and unstable manifolds for analytic perturbations.

We extend analytically the solutions of the equations of motion, order by order in the perturbation parameter, to a uniform neighbourhood of the time axis.

1. Main Concepts and Results

1.1. **Background and history.** A quasiperiodic motion of a mechanical system is composed of incommensurable periodic motions; the trajectory in phase space winds around on a torus filling its surface densely. An integrable Hamiltonian system has a great profusion of quasiperiodic motions: if one picks an initial phase point according to a uniform distribution, the trajectory will be quasiperiodic with probability one. The remaining trajectories are periodic.

KAM theory deals with the stability of quasiperiodic motions, or persistence of invariant tori, under small perturbations. Poincaré [Poi93a] called this the general problem of dynamics.

In 1954, Kolmogorov [Kol54] outlined a result, made rigorous by Arnold in 1963 [Arn63], that quasiperiodic motions are typical also for nearly integrable analytic Hamiltonians under suitable nondegeneracy conditions. Thus, only a small fraction of the tori would be destroyed by the perturbation. Moser managed to prove the same for twist maps [Mos62] in 1962, and later for Hamiltonians [Mos66a, Mos66b], in the smooth (non-analytic) setting (see also [Mos67]).

The difficult problem to overcome is the following. Suppose that the Hamiltonian reads $\mathcal{H} = \mathcal{H}_0 + \lambda\mathcal{H}_1$, where $\mathcal{H}_0$ is integrable and λ is considered small. Then one can formally represent a solution to the equations of motion by a power series in λ, known as the Lindstedt series in this context, conditioned to agree for $\lambda = 0$ with a quasiperiodic solution obtained in the integrable case. When one computes the coefficients of the Lindstedt series, however, one encounters expressions containing arbitrarily small denominators. The latter seem to imply that the kth coefficient grows like $k!^\alpha$ with a large power α. Thus, there is little hope of being able to sum the series and obtain a true solution, unless a miracle occurs.

The proofs mentioned above relied on a rapidly convergent Newton-type iteration scheme, which is interesting in its own right, and yields solutions analytic in λ. On the other hand, one is then left to wonder why the Lindstedt series *does* converge.

2000 *Mathematics Subject Classification.* 70K43; Secondary 37J40, 37D10, 70H08, 70K44.

In 1988, an answer was provided by Eliasson [Eli96], who managed to identify enormous cancellations among the small denominator contributions and to sum the Lindstedt series "manually". Gallavotti [Gal94a, Gal94b] interpreted the cancellations in a Renormalization Group (RG) framework. For a review and some extensions, see Gentile and Mastropietro [GM96]. The importance of these achievements has to be stressed: they prove the existence of quasiperiodic solutions in an essentially constructive way.

Motivated by the RG approach of Gallavotti, in the 1999 paper [BGK99] Bricmont, Gawędzki, and Kupiainen identified the cancellations as a consequence of Ward identities (corresponding to a translation invariance of an action functional) in a suitable field theory.

Returning to much earlier works, Moser [Mos67] and Graff [Gra74] showed that also hyperbolic tori—tori having local stable and unstable manifolds—would typically persist under small perturbations. In another landmark paper [Arn64], Arnold had described a mechanism how a chain of such "whiskered" tori could provide a way of escape for special trajectories, resulting in instability in the system. (As discussed above, a trajectory would typically lie on a torus and therefore stay eternally within a bounded region in phase space.) The latter is often called Arnold mechanism and the general idea of instability goes by the name Arnold diffusion. It is conjectured in [AA68] that Arnold diffusion due to Arnold mechanism is present quite generically, among others in the three body problem.

Arnold mechanism is based on Poincaré's concept of biasymptotic solutions, discussed in the last chapter of [Poi93b], that are formed at intersections of whiskers of tori. Following such intersections a trajectory can "diffuse" in a finite time from a neighbourhood of one torus to a neighbourhood of another, and so on.

Chirikov's work [Chi79] is a very nice physical account on Arnold diffusion. Lochak's compendium [Loc99] discusses more recent developments in a readable fashion and is a good point to start learning about diffusion.

The proofs of Moser and Graff mentioned above use the rapidly convergent method of Kolmogorov, but there now exist also constructive proofs in the spirit of Eliasson and Gallavotti. We refer here to Gallavotti [Gal94b] and Gentile [Gen95a, Gen95b].

1.2. **The model.**

We consider the Hamiltonian

$$\mathcal{H}(I,\phi,A,\psi) = \tfrac{1}{2}I^2 + g^2\cos\phi + \tfrac{1}{2}A^2 - \lambda f(\phi,\psi) \tag{1.1}$$

of a pendulum coupled to d rotators, with $\phi \in \mathbb{S}^1 := \mathbb{R}/2\pi\mathbb{Z}$ and $I \in \mathbb{R}$ the coordinate and momentum of the pendulum, and $\psi \in \mathbb{T}^d := (\mathbb{S}^1)^d$ and $A \in \mathbb{R}^d$ the angles and actions of the rotators, respectively. The perturbation f is assumed to be real-valued and real-analytic in its arguments, and λ is a (small) real number, whereas the gravitational coupling constant g is taken to be positive. This Hamiltonian is sometimes called *the generalized Arnold model* or the *Thirring model.* It is *the* prototype of a nearly integrable Hamiltonian system close to a simple resonance, as is explained in the introduction of [Gen95b]. A review of applications can be found in [Chi79].

The equations of motion are

$$\dot\phi = I, \quad \dot\psi = A, \quad \dot I = g^2\sin\phi + \lambda\,\partial_\phi f, \quad \dot A = \lambda\,\partial_\psi f. \tag{1.2}$$

For the parameter value $\lambda = 0$, which is addressed as the unperturbed case, the pendulum and the rotators decouple. The former then has the separatrix flow $\phi : \mathbb{R} \to \mathbb{S}^1$ given by

$$\phi(t) = \Phi^0(e^{gt}),$$

where

$$\Phi^0(z) = 4 \arctan z.$$

By elementary trigonometry, this function possesses the symmetry property

$$\Phi^0(z) = 2\pi - \Phi^0(z^{-1}). \tag{1.3}$$

It is also odd,

$$\Phi^0(-z) = -\Phi^0(z).$$

The phase space of the unperturbed pendulum looks as in Figure 1, where the separatrix, given by Φ^0, separates closed trajectories (libration) from open ones (rotation).

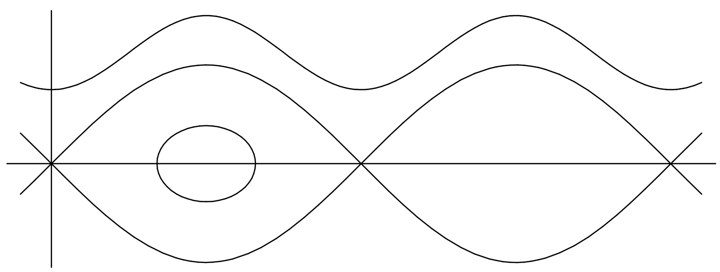


FIGURE 1. A (ϕ, I) plot showing the unperturbed pendulum separatrix that intersects the ϕ axis at integer multiples of 2π—the upright position of the pendulum.

On the other hand, $\psi : \mathbb{R} \to \mathbb{T}^d$ is quasiperiodic:

$$\psi(t) = \psi(0) + \omega t \pmod{2\pi},$$

such that the vector

$$\omega := A(0) \equiv A(t)$$

satisfies the Diophantine condition

$$|\omega \cdot q| > a\,|q|^{-\nu} \quad \text{for} \quad q \in \mathbb{Z}^d,\ q \neq 0, \tag{1.4}$$

with a and ν positive. Thus, at the instability point of the pendulum, the flow possesses the invariant tori

$$\mathcal{T}_0 := \Big\{ (\phi, \psi, I, A) = (0, \theta, 0, \omega) \;\Big|\; \theta \in T^d \Big\}$$

indexed by ω, with stable and unstable manifolds ($\mathcal{W}_0^s$ and $\mathcal{W}_0^u$, respectively) coinciding:

$$\mathcal{W}_0^{s,u} = \Big\{ (\phi, \psi, I, A) = \big(\Phi^0(z), \theta, gz\partial_z \Phi^0(z), \omega\big) \;\Big|\; z \in [-\infty, \infty],\ \theta \in T^d \Big\}. \tag{1.5}$$

Remark 1.1. The constant g is the Lyapunov exponent for the unstable fixed point of the pendulum motion; in the limit $s \to -\infty$ two nearby initial angles $\phi(s)$ and $\phi(s+\delta s)$ separate at the exponential rate e^{gs}. As $\phi(t) = \Phi^0(e^{t/g^{-1}})$, the Lyapunov exponent fixes a natural time scale of g^{-1} units, characteristic of the pendulum motion in the unperturbed Hamiltonian system (1.1).

When the perturbation is switched on ($\lambda \neq 0$), we show that some of the invariant tori survive and have stable and unstable manifolds—or "whiskers" as Arnold has called them—that may not coincide anymore.

1.3. **Main theorems.** Our approach will be to construct the perturbed manifolds in a form similar to (1.5) as graphs of analytic functions over a piece of $[-\infty, \infty] \times \mathbb{T}^d$. To see how this can be achieved, note that the unperturbed stable and unstable manifolds, $\mathcal{W}_0^s$ and $\mathcal{W}_0^u$, consist of trajectories

$$(\phi(t), \psi(t)) = (\Phi^0(e^{gt}), \omega t)$$

that at time $\pm\infty$ become quasiperiodic, as they wrap tighter and tighter around the invariant torus $\mathcal{T}_0$; indeed $(\phi(t), \psi(t)) \sim (0, \omega t)$ in the limit $t \to \pm\infty$.

Analogously, we will find the stable and unstable manifolds of the perturbed tori by looking for solutions of the form

$$(\phi(t), \psi(t)) = (\Phi(e^{\gamma t}, \omega t), \omega t + \Psi(e^{\gamma t}, \omega t)) = (0, \omega t) + (\Phi, \Psi)(e^{\gamma t}, \omega t) \qquad (1.6)$$

with quasiperiodic behavior in *one* of the two limits $t \to \pm\infty$. Note especially that we anticipate the Lyapunov exponent $\gamma > 0$ to depend on λ, with $\gamma|_{\lambda=0} = g$.

Remark 1.2. One should not assume asymptotic quasiperiodicity in both of the limits $t \to \pm\infty$, as the unstable and stable manifolds, which we denote $\mathcal{W}_\lambda^u$ and $\mathcal{W}_\lambda^s$, are generically expected to depart for nonzero values of the perturbation parameter λ. Therefore, either the past *or* future asymptotic of a trajectory will evolve so as to ultimately reach the (deformed) invariant torus $\mathcal{T}_\lambda$. The separatrix in Figure 1 is thus transformed into something like the pair of curves in Figure 2.

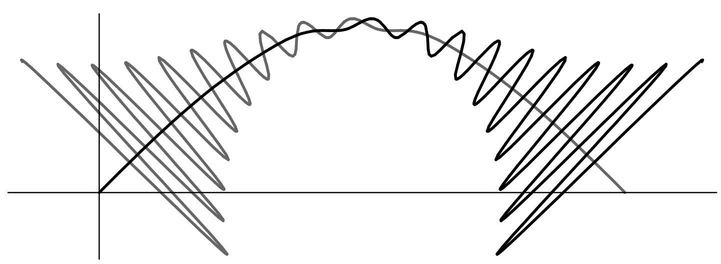


FIGURE 2. A schematic I-versus-ϕ plot, on a section of constant ψ ($d = 1$). The stable and unstable manifolds are expected to split, as opposed to coincide. The origin has been shifted for convenience.

Let us denote the total derivative d/dt by ∂_t and the complete angular gradient $(\partial_\phi, \partial_\psi)$ by ∂ for short. Substituting (1.6) into the equations of motion

$$\partial_t^2(\phi, \psi) = (\dot{I}, \dot{A}) = (g^2 \sin\phi, 0) + \lambda\, \partial f(\phi, \psi),$$

we get for $X := (\Phi, \Psi)$ the equation

$$(\omega \cdot \partial_\theta + \gamma e^{\gamma t}\partial_z)^2 X(e^{\gamma t}, \omega t) = [(g^2 \sin\Phi, 0) + \lambda\, \partial f(X + (0, \theta))](e^{\gamma t}, \omega t),$$

where θ stands for the canonical projection $[-\infty, \infty] \times \mathbb{T}^d \to \mathbb{T}^d$.

Notice that the partial differential operator

$$\mathcal{L} := \omega \cdot \partial_\theta + \gamma z \partial_z$$

satisfies the characteristic identity

$$\mathcal{L}F(ze^{\gamma t},\theta+\omega t)=\partial_t F(ze^{\gamma t},\theta+\omega t) \tag{1.7}$$

for a differentiable map $(z,\theta)\mapsto F(z,\theta)$. Equation (1.7) simply reflects the time derivative nature of $\mathcal{L}$. In fact, if T is the "time-reversal map"

$$T(z,\theta)\equiv(z^{-1},-\theta), \tag{1.8}$$

then, by the chain rule,

$$\mathcal{L}(F\circ T)=-(\mathcal{L}F)\circ T. \tag{1.9}$$

Let us abbreviate

$$\Omega(X):=(g^2\sin\Phi,0)+\lambda\,\widetilde{\Omega}(X)\quad\text{with}\quad\widetilde{\Omega}(X):=\partial f(X+(0,\theta)). \tag{1.10}$$

As a consequence, we have reduced the equations of motion to the PDE

$$\mathcal{L}^2X=\Omega(X) \tag{1.11}$$

for the map $(z,\theta)\mapsto X(z,\theta)$ in a suitable Banach space of analytic functions, albeit its restriction to the set ("characteristic")

$$\big\{(z,\theta)=(e^{\gamma t},\omega t)\;\big|\;t\in\mathbb{R}\big\} \tag{1.12}$$

is what one is physically interested in. Our preference of working directly with the invariant manifolds, as opposed to individual trajectories traversing along them, motivates us encoding the time derivative in the operator $\mathcal{L}$. Nevertheless, it will be harmless—and indeed quite informative—for the reader to keep in mind that the objects we deal with originate from (1.12) and therefore have a direct physical interpretation.

The action variables trivially follow from the knowledge of $X(z,\theta)$:

$$(I(t),A(t))=(0,\omega)+Y(e^{\gamma t},\omega t),\quad Y:=\mathcal{L}X.$$

The solutions X will provide a parametrization of the deformed tori and their stable and unstable manifolds. As hinted below (1.6), we find two kinds of solutions, $X^u(z,\theta)$ defined for $z\in[-z_0,z_0]=:\mathbb{I}^u$ and $X^s(z,\theta)$ defined for $z\in[-\infty,-z_0^{-1}]\cup[z_0^{-1},\infty]=:\mathbb{I}^s$. Here, $z_0>1$. The tori will have the three parametrizations

$$\begin{aligned}\mathcal{T}_\lambda &= \Big\{(\phi,\psi,I,A)=((0,\theta)+X^u(0,\theta),(0,\omega)+Y^u(0,\theta))\;\Big|\;\theta\in T^d\Big\}\\ &= \Big\{(\phi,\psi,I,A)=((0,\theta)+X^s(\pm\infty,\theta),(0,\omega)+Y^s(\pm\infty,\theta))\;\Big|\;\theta\in T^d\Big\},\end{aligned}$$

whereas the parametrizations of their stable and unstable manifolds then read

$$\mathcal{W}^{s,u}_\lambda=\Big\{(\phi,\psi,I,A)=((0,\theta)+X^{s,u}(z,\theta),(0,\omega)+Y^{s,u}(z,\theta))\;\Big|\;z\in\mathbb{I}^{s,u},\;\theta\in T^d\Big\}.$$

In order to enable solving (1.11), we need to deal with quantities of the form $(\omega\cdot q)^{-1}$, $q\in\mathbb{Z}^d\setminus\{0\}$, stemming from the Fourier representation of the operator $\mathcal{L}$. Here the Diophantine property of the vector $\omega\in\mathbb{R}^d$ stated in (1.4) steps in. Since $\omega\equiv A|_{\lambda=0}=\dot\psi|_{\lambda=0}$, by rescaling time (and the actions, correspondingly) in the equations of motion (1.2), the constant a can be absorbed into g^2 and λ in the equations of motion, leaving the ratio λg^{-2} unchanged: $(g,\lambda)\mapsto(g/a,\lambda/a^2)$ [1]. Thus, we may as well take a to be 1 below, transforming the condition on ω into

$$|\omega\cdot q|>|q|^{-\nu}\quad\text{for}\quad q\in\mathbb{Z}^d\setminus\{0\}. \tag{1.13}$$

[1] This scaling is responsible for the usual requirement $\lambda=\mathcal{O}(a^2)$ for KAM tori.

We will moreover consider λ small in a g-dependent fashion, taking

$$\epsilon := \lambda g^{-2} \tag{1.14}$$

small. This should be seen as an outreach towards the experimenter, albeit there is a technical wherefore: such a choice is needed for studying the limit $g \to \infty$, which corresponds to rapid forcing; see Remark 1.1. The domain we restrict ourselves to is

$$D := \left\{(\epsilon, g) \in \mathbb{C} \times \mathbb{R} \,\middle|\, |\epsilon| < \epsilon_0,\ 0 < g < g_0\right\}, \tag{1.15}$$

for some positive values of ϵ_0 and g_0.

Finally, note that if $X = (\Phi, \Psi)$ solves (1.11) on some domain $D' \subset [-\infty, \infty] \times \mathbb{T}^d$, then so does

$$X_{\alpha,\beta}(z, \theta) := X(\alpha z, \theta + \beta) + (0, \beta), \tag{1.16}$$

as long as $(\alpha z, \theta + \beta \pmod{2\pi}) \in D'$. The aforementioned invariance is a manifestation of the freedom of choosing initial conditions for (ϕ, ψ)—we may choose the origin of time and the configuration of the physical system there.

For $\epsilon = 0$, the solutions are obtained from

$$X^0(z, \theta) := (\Phi^0(z), 0) \tag{1.17}$$

using (1.16). In particular, $X^0(1, 0) = (\pi, 0)$. This will provide us with a natural way of fixing α and β below.

We are now ready to state the first of the two main theorems of this article. It is a version of a classical result, and by no means new; earlier treatments include *for instance* [Mel63, Mos67, Gra74, Eli94, Gal94b, Gen95a, Gen95b]. However, the interest here lies in the new techniques used in the proof.

Theorem 1 (Tori and their whiskers)**.** *Let f be real-analytic and even, i.e.,*

$$f(\phi, \psi) = f(-\phi, -\psi).$$

Also, suppose ω satisfies the Diophantine condition (1.13)*, and fix $g_0 > 0$. Then there exist a positive number ϵ_0 and a function $\gamma(\epsilon, g)$ on D, analytic in ϵ with $|\gamma - g| < Cg|\epsilon|$, such that equation* (1.11) *has a solution X^u which is analytic in ϵ as well as in (z, θ) in a neighbourhood of $[-1, 1] \times \mathbb{T}^d$ and which satisfies*

$$X^u(1, 0) = (\pi, 0), \quad X^u(z, \theta) = X^0(z) + \mathcal{O}(\epsilon). \tag{1.18}$$

Corresponding to the same γ, there exists a solution $X^s(z, \theta) = X^0(z) + \mathcal{O}(\epsilon)$ which is an analytic function of $(z^{-1}, -\theta)$ in a neighbourhood of $[-1, 1] \times \mathbb{T}^d$. The maps

$$W^{s,u}(z, \theta) = (X^{s,u}, Y^{s,u})(z, \theta) + ((0, \theta), (0, \omega)), \quad Y^{s,u} := \mathcal{L}X^{s,u}, \tag{1.19}$$

provide analytic parametrizations of the stable and unstable manifolds $\mathcal{W}^{s,u}_\lambda$ of the torus $\mathcal{T}_\lambda$.

Remark 1.3. The number ϵ_0 above depends on the Diophantine exponent ν and on f. The perturbation $(\phi, \psi) \mapsto f(\phi, \psi)$ is analytic on the compact set $\mathbb{S}^1 \times \mathbb{T}^d$. By Abel's Lemma (multivariate power series converge on polydisks), it extends to an analytic map on a "strip" $|\Im\mathfrak{m}\, \phi|, |\Im\mathfrak{m}\, \psi| \leq \eta$ $(\eta > 0)$ around $\mathbb{S}^1 \times \mathbb{T}^d$. By Theorem 1, there exists some $0 < \sigma < \eta$ such that each $\theta \mapsto X^{s,u}(\,\cdot\,, \theta)$ is analytic on $|\Im\mathfrak{m}\, \theta| \leq \sigma$.

An important part of Theorem 1 is that the domains of X^u and X^s overlap. Namely, if $(z,\theta)\mapsto X(z,\theta)$ solves equation (1.11), then so does $(z,\theta)\mapsto(2\pi,0)-(X\circ T)(z,\theta)$. This is due to (1.9) and the parity of f. Consequently, by a simple time-reversal consideration (set $t\mapsto -t$ in (1.12)), the stable and unstable manifolds are related through

$$X^s=(2\pi,0)-X^u\circ T. \tag{1.20}$$

In particular, as $T(1,0)=(1,0)$,

$$X^s(1,0)=X^u(1,0).$$

Moreover, the actions $Y^{s,u}=\mathcal{L}X^{s,u}$ satisfy

$$Y^s=Y^u\circ T, \tag{1.21}$$

yielding

$$Y^s(1,0)=Y^u(1,0).$$

In other words, a *homoclinic intersection* of the stable and the unstable manifolds $\mathcal{W}_\lambda^{s,u}$ occurs at $(z,\theta)=(1,0)$, as their parametrizations (1.19) coincide at this *homoclinic point*. Since the manifolds $\mathcal{W}_\lambda^{s,u}$ are invariant, there in fact exists a *homoclinic trajectory* on which the parametrizations agree:

$$W^s(e^{\gamma t},\omega t)\equiv W^u(e^{\gamma t},\omega t). \tag{1.22}$$

Remark 1.4. Equation (1.20) is what remains of the symmetry $X^0=(2\pi,0)-X^0\circ T$, which is just another way of writing (1.3), after the onset of *even* perturbation. This is an instance of *spontaneous symmetry breaking*: The equations of motion, (1.11), remain unchanged under the transformation $X\mapsto(2\pi,0)-X\circ T$, but the individual solutions do not respect this symmetry; $X^u\neq X^s=(2\pi,0)-X^u\circ T$, if $\lambda\neq 0$.

Coming to the second one of our main results, let us expand

$$X^u=\sum_{\ell=0}^{\infty}\epsilon^\ell X^{u,\ell}.$$

In Section 5, we will show that the common analyticity domain of each $X^{u,\ell}$ in the z-variable is in fact much larger than the (small) neighbourhood of $[-1,1]$—the corresponding analyticity domain of X^u according to Theorem 1; namely it includes the wedgelike region

$$\mathbb{U}_{\tau,\vartheta}:=\big\{|z|\le\tau\big\}\bigcup\big\{\arg z\in[-\vartheta,\vartheta]\cup[\pi-\vartheta,\pi+\vartheta]\big\}\subset\mathbb{C}$$

(with some positive τ and ϑ):

Theorem 2 (Analytic continuation)**.** *Each order $X^{u,\ell}$ of the solution extends analytically to a common region $\mathbb{U}_{\tau,\vartheta}\times\{|\Im\mathfrak{m}\,\theta|\le\sigma\}$. Moreover, if $\psi\mapsto f(\,\cdot\,,\psi)$ is a trigonometric polynomial of degree N, i.e., N is the minimal nonnegative integer such that $\hat f(\,\cdot\,,q)=0$ whenever $|q|>N$, then $\theta\mapsto X^{u,\ell}(\,\cdot\,,\theta)$ is a trigonometric polynomial of degree ℓN, at most.*

Remark 1.5. With η and σ as in Remark 1.3, the numbers τ and ϑ are specified by the following observation: $\Phi^0(z)=4\arctan z$ implies that $|\Im\mathfrak{m}\,\Phi^0(z)|\le\eta$ in $\mathbb{U}_{\tau,\vartheta}$ with τ and ϑ sufficiently small. By Remark 1.3, $(z,\theta)\mapsto f(\Phi^0(z),\theta)$ is analytic on $\mathbb{U}_{\tau,\vartheta}\times\{|\Im\mathfrak{m}\,\theta|\le\sigma\}$, which we will use as the basis of the proof.

In spite of Theorem 2, (a straightforward upper bound on) $X^{u,\ell}$ grows without a limit as $|\mathfrak{Re}\, z| \to \infty$, such that there is no reason whatsoever to expect absolute convergence of the series $\sum_{\ell=0}^{\infty} \epsilon^\ell X^{u,\ell}$ in an unbounded z-domain with a fixed ϵ. In fact, it is known that the behavior of the unstable manifold gets extremely complicated for large values of z even with innocent looking Hamiltonian systems. Still, it seems to us that the possibility of a uniform analytic extension of the coefficients $X^{u,\ell}$ has not been appreciated in the literature.

Due to (1.20), an analog of Theorem 2 and the subsequent discussion are seen to hold for the solution X^s, with z replaced by z^{-1}.

Theorem 2 is interesting, because it allows one (at each order in ϵ) to track trajectories $t \mapsto W^{s,u}(e^{\gamma t}, \theta + \omega t)$ on the invariant manifolds $\mathcal{W}_\lambda^{s,u}$ for arbitrarily long times in a uniform complex neighbourhood $|\mathfrak{Im}\, t| \leq g^{-1}\vartheta$ of the real line, for arbitrary $\theta \in \mathbb{T}^d$. The motivation for doing this stems from studying the splitting of the manifolds $\mathcal{W}_\lambda^{s,u}$ in the vicinity of the homoclinic trajectory (1.22), and is the topic of another article. The general ideology that, being able to extend "splitting related functions" to a large complex domain yields good estimates, is due to Lazutkin [Laz03], as is emphasized in [LMS03].

1.4. **Strategy.** Let us briefly explain how Theorem 1 will be proved in three steps. Due to (1.20), we may concentrate on studying the unstable manifold. Thus, we write

$$X(z,\theta) := X^u(z,\theta) = X_0(\theta) + zX_1(\theta) + \delta_2 X(z,\theta).$$

From (1.11) we first get an equation for $X_0 := X^u(0,\,\cdot\,)$ alone. Second, *given* X_0, an equation for $X_1 := \partial_z X^u(0,\,\cdot\,)$ and γ alone is obtained. Third, *given* X_0, X_1, *and* γ, an equation for the remainder $\delta_2 X$ is obtained.

It turns out that solving for X_0 and X_1 (together with γ), *i.e.*, the invariant torus and the linearization of the unstable manifold around it, is difficult. Namely, these problems involve the small denominators of KAM theory. In contrast, solving for $\delta_2 X$ amounts to a simple Contraction Mapping argument.

We deduce the existence of X_0 from [BGK99]. The existence proof of X_1 is reminiscent of the RG argument in the latter paper, except that the Lyapunov exponent γ has to be fine-tuned to a proper value such that the renormalization flow converges.

At this point we would like to draw the readers attention to the interesting reference [Gen95b], where the author takes a different approach. Gentile fixes the perturbed Lyapunov exponent γ in advance and replaces g by $\tilde{g}(\epsilon,\gamma)$ in the Hamiltonian, which is analogous to introducing counterterms in quantum field theory, and finds the corresponding manifolds. One could then solve the implicit equation $\tilde{g}(\epsilon,\gamma) = g$ and to obtain γ as a function of g and ϵ.

Acknowledgements. I am indebted to Antti Kupiainen for his help during the course of this work. Guido Gentile, Kari Astala, and Jean Bricmont provided sharp remarks and critical comments on the manuscript that made it more comprehensible and mathematically accurate. I wish to express my gratitude to all of them. I thank Giovanni Gallavotti and Emiliano De Simone for discussions at Rutgers University and University of Helsinki, respectively.

2. Perturbed Tori

The perturbed tori will be found by looking for solutions having the general form

$$\phi(t) = \Phi_0(\omega t), \quad \psi(t) = \omega t + \Psi_0(\omega t),$$

with $\Phi_0 : \mathbb{T}^d \to \mathbb{R}$ and $\Psi_0 : \mathbb{T}^d \to \mathbb{R}^d$ satisfying the "$t \to -\infty$ asymptotics"

$$\mathcal{D}^2\Phi_0(\theta) = g^2 \sin\Phi_0(\theta) + \lambda\,\partial_\phi f(\Phi_0(\theta), \theta + \Psi_0(\theta)) \tag{2.1}$$

$$\mathcal{D}^2\Psi_0(\theta) = \lambda\,\partial_\psi f(\Phi_0(\theta), \theta + \Psi_0(\theta)) \tag{2.2}$$

obtained from equation (1.11) by putting $z = 0$ and $\mathcal{D} = \omega \cdot \partial_\theta$. Note that if $X_0 = (\Phi_0, \Psi_0)$ is a solution to equations (2.1) and (2.2), then so is

$$\sigma_\beta X_0(\theta) := (\Phi_0(\theta + \beta), \Psi_0(\theta + \beta) + \beta) \tag{2.3}$$

for $\beta \in \mathbb{T}^d$. We point out that together (2.1) and (2.2) are equivalent to

$$\mathcal{D}^2 X_0 = \Omega(X_0). \tag{2.4}$$

2.1. **Spaces of analytic functions.** Let us define the spaces we shall be working in. As linear subspaces of ℓ^1, the Banach spaces

$$\mathcal{B}_\sigma^\Phi := \Big\{\Phi : \mathbb{T}^d \to \mathbb{C} \;\Big|\; \|\Phi\|_\sigma := \sum_{q\in\mathbb{Z}^d} |\hat\Phi(q)| e^{\sigma|q|} < \infty\Big\},$$

$$\mathcal{B}_\sigma^\Psi := \Big\{\Psi : \mathbb{T}^d \to \mathbb{C}^d \;\Big|\; \|\Psi\|_\sigma := \sum_{q\in\mathbb{Z}^d} |\hat\Psi(q)| e^{\sigma|q|} < \infty\Big\},$$

for any $\sigma \geq 0$, have the advantage that Fourier analysis on their elements is convenient. Furthermore, we are trying to find a solution $X = (\Phi, \Psi)$ analytic on the torus, and, for a suitably small σ, such a function belongs to $\mathcal{B}_\sigma^\Phi \times \mathcal{B}_\sigma^\Psi$ because of the exponential decay of its Fourier coefficients; $|\hat X(q)| < Ce^{-\sigma|q|}$ with some positive constant C. Indeed, if $\sigma > 0$, the spaces above comprise precisely those functions on the torus that admit an analytic extension to the "strip" $|\Im\mathfrak{m}\,\theta| < \sigma$. We will occasionally write $\mathcal{B}_\sigma$ when referring to either one of $\mathcal{B}_\sigma^\Phi$ and $\mathcal{B}_\sigma^\Psi$.

Of course, as our analysis proceeds, the perturbation f will appear all over the place. This, in turn, dictates the analyticity properties of a plethora of maps, in practice introducing the constraint $\sigma \leq \eta$ for the spaces $\mathcal{B}_\sigma$; see Remark 1.3.

Notice the natural embeddings

$$\mathcal{B}_{\sigma+\alpha} \subset \mathcal{B}_\sigma,$$

for $\alpha \geq 0$, due to the inequality

$$\|\cdot\|_\sigma \leq \|\cdot\|_{\sigma+\alpha}. \tag{2.5}$$

Consider the linear operator $\tau_\beta : \mathcal{B}_{\sigma+\alpha} \to \mathcal{B}_\sigma$ defined through setting $\widehat{\tau_\beta X}(q) = e^{iq\cdot\beta}\hat X(q)$, with $\beta \in \mathbb{C}^d$. Whenever $|\Im\mathfrak{m}\,\beta| \leq \alpha$, $\|\tau_\beta\|_{\mathcal{L}(\mathcal{B}_{\sigma+\alpha};\mathcal{B}_\sigma)} \leq 1$. The realization of τ_β in terms of the variable θ is just the translation $\Psi(\theta) \mapsto \Psi(\theta + \beta)$. τ_β will serve as a useful device in encoding the real-analyticity of f as an algebraic property into the Fourier series of certain other functions. This is due to the the fact that exponential smallness of $|\widehat X(q)|$ in q implies real-analyticity of a function X on the torus, and vice versa.

We shall encounter n-linear maps from $\mathbb{C}^{d+1}$ into $\mathbb{C}$. Endowed with the norm

$$\|A\|_{\mathcal{L}(^n(\mathbb{C}^{d+1});\mathbb{C})} := \inf\Big\{M \geq 0 \;\Big|\; |A(z_1,\dots,z_n)| \leq M|z_1|\dots|z_n| \quad \forall\, z_i \in \mathbb{C}^{d+1}\Big\}$$

they form the Banach space $\mathcal{L}(^n(\mathbb{C}^{d+1});\mathbb{C})$; see [Cha85].

2.2. **Past and future asymptotics of the solution in the perturbed case.** This subsection discusses the $t \to \pm\infty$ asymptotics of the solution X. In these limits the motion settles onto the "distorted version" $\mathcal{T}_\lambda$ of the invariant torus $\mathcal{T}_0$ with the pendulum seizing to swing, but wiggling quasiperiodically about its unstable equilibrium.

Theorem 2.1. *Under the assumptions of Theorem 1, there exist positive numbers r and ϵ_0 such that, for $(\epsilon, g) \in D$, equations* (2.1) *and* (2.2) *have a unique solution $X_0 = (\Phi_0, \Psi_0)$ in the class of those real-analytic functions of $\theta \in \mathbb{T}^d$ that satisfy $\|\Psi_0\|_{\ell^1} < r$ and $\langle\Psi_0\rangle = 0$ (zero average). The function X_0, defined on $\{|\Im\mathfrak{m}\,\theta| \leq \sigma\} \times D$ for some $\sigma > 0$, is analytic and uniformly bounded by $(C|\epsilon|, Cg^2|\epsilon|)$. Moreover, it is $\mathbb{R}\times\mathbb{R}^d$-valued on $\mathbb{T}^d$ for ϵ real. Thus, any real-analytic solution $X_0' = (\Phi_0', \Psi_0')$ with $\langle\Psi_0'\rangle = \beta \in \mathbb{R}^d$ and $\|\Psi_0' - \beta\|_{\ell^1} < r$ must be the one given by*

$$X_0'(\theta) \equiv X_0(\theta + \beta) + (0, \beta),$$

i.e., $X_0' = \sigma_\beta X_0$, using the notation of (2.3).

Remark 2.2. Remark 1.3 below Theorem 1 holds true. Recall that we have defined $\epsilon := \lambda g^{-2}$ in (1.14) and the domain D in (1.15). This is a version of the KAM Theorem. Notice that $X_0 \in \mathcal{B}_\sigma^\Phi \times \mathcal{B}_\sigma^\Psi$.

Proof. The proof is a reduction to the one given in [BGK99]. Here we systematically omit the subindex 0 of Φ_0, Ψ_0, and X_0. Let us concentrate on the pendulum part, equation (2.1), first. We expect Φ to be close to its unperturbed value, zero, and it pays to cancel the leading term of $g^2 \sin\Phi(\theta)$ on the right-hand side by subtracting $g^2\Phi(\theta)$ from both sides. We then have

$$(\mathcal{D}^2 - g^2)\Phi = U(\Phi, \Psi) =: U(X) \tag{2.6}$$

with

$$U(X)(\theta) := g^2(\sin\Phi(\theta) - \Phi(\theta)) + \lambda\, \partial_\phi f(\Phi(\theta), \theta + \Psi(\theta)). \tag{2.7}$$

Pay attention to the fact that $U(X)(\theta)$ depends locally on X—only through $X(\theta)$, that is. Abusing notation, we shall use $U(X)(\theta)$, $U(X,\theta)$, $U(X(\theta),\theta)$, *etc.*, in the same meaning, whichever is the most convenient form. Now, $U(\chi,\theta)$ is analytic in the vector argument $\chi = (\chi_\phi, \chi_\psi)$ in the region $|\chi_\phi|, |\chi_\psi| \leq \eta$, where $\eta > 0$ depends on the analyticity domain of f; see Remark 1.3 on page 6.

Let us now write down the Fourier–Taylor expansion

$$\begin{aligned} U(X(\theta),\theta) &= \sum_{n=0}^{\infty} \frac{1}{n!} D^n U(0,\theta)\,(X(\theta),\dots,X(\theta)) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\substack{\mathbf{q}=(q_1,\dots,q_n)\\ q_i\in\mathbb{Z}^d}} e^{i\theta\cdot\sum_i q_i}\, D^n U(0,\theta)\,(\hat{X}(q_1),\dots,\hat{X}(q_n)), \end{aligned} \tag{2.8}$$

where $D^n U(0,\theta) \in \mathcal{L}({}^n(\mathbb{C}^{d+1});\mathbb{C})$ is the nth Fréchet derivative of the map $U(\,\cdot\,,\theta) : \mathbb{C}^{d+1} \to \mathbb{C} : \chi \mapsto U(\chi,\theta)$.

The map $\theta \mapsto U_n(\theta) := \frac{1}{n!} D^n U(0,\theta)$ is analytic in the same domain as $\theta \mapsto U(0,\theta) = \lambda\,\partial_\phi f(0,\theta)$, *i.e.*, $|\mathfrak{Im}\,\theta| \leq \eta$. Its Fourier representation $U_n(\theta) = \sum_{q\in\mathbb{Z}^d} e^{iq\cdot\theta} u_n(q)$ has coefficients

$$u_n(q) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} e^{-iq\cdot\theta}\,\frac{1}{n!} D^n U(0,\theta)\,d\theta \tag{2.9}$$

in $\mathcal{L}({}^n(\mathbb{C}^{d+1});\mathbb{C})$. Using this notation, we translate (2.8) into the Fourier language;

$$\widehat{U(X)}(q) = \sum_{n=0}^{\infty} \sum_{\mathbf{q}\in(\mathbb{Z}^d)^n} u_n(q - \sum_{i=1}^{n} q_i)\,(\hat{X}(q_1),\ldots,\hat{X}(q_n)). \tag{2.10}$$

The right-hand side of equation (2.10) is a power series in $\hat{X}$, converging whenever $\hat{X}$ is sufficiently close to zero. Namely, we have

Lemma 2.3. *The multilinear maps $u_n(q)$ obey the bound*

$$\|u_n(q)\|_{\mathcal{L}({}^n(\mathbb{C}^{d+1});\mathbb{C})} \leq Cg^2(r_0^3 + |\epsilon|)(r_0/e)^{-n} e^{-\rho|q|}, \tag{2.11}$$

where ρ and r_0 is any pair of positive numbers satisfying $\rho + r_0 = \eta$, $\eta > 0$ being the width of the analyticity domain of f as explained in Remark 1.3.

The proof of Lemma 2.3 is straightforward, but, for the sake of continuity, is given in Subsection 2.3 below.

Considering the closed origin-centered balls of radius $r < r_0/2$ in $\mathcal{B}^\Phi_\sigma$ and $\mathcal{B}^\Psi_\sigma$—$\bar{B}^\Phi_{\sigma,r}$ and $\bar{B}^\Psi_{\sigma,r}$, respectively—we next study $U_\beta : \bar{B}^\Phi_{\sigma,r} \times \bar{B}^\Psi_{\sigma,r} \to \mathcal{B}^\Phi_\sigma : (\Phi,\Psi) \mapsto \tau_\beta U(\tau_{-\beta}\Phi, \tau_{-\beta}\Psi)$. By equation (2.7),

$$U_\beta(\Phi(\theta),\Psi(\theta),\theta) = U(\Phi(\theta), \theta + \beta + \Psi(\theta)), \tag{2.12}$$

when $\beta \in \mathbb{R}^d$. The right-hand side is analytic in β, and extends to $|\mathfrak{Im}\,\beta| + \sigma + r < \eta$ through the same expression, leaving U_β analytic with respect to X.

More quantitatively, one checks using the bound (2.11) that the power series

$$\widehat{U_\beta(X)}(q) = \sum_{n=0}^{\infty} \sum_{\mathbf{q}\in(\mathbb{Z}^d)^n} e^{i\beta\cdot(q-\sum_i q_i)}\,u_n(q - \sum_{i=0}^{n} q_i)\,(\hat{X}(q_1),\ldots,\hat{X}(q_n)), \tag{2.13}$$

converges uniformly with respect to X and β, even if the latter has a small imaginary part. In fact, $\|U_\beta(X)\|_\sigma$ obeys the upper bound

$$\begin{aligned}
&\sum_{n=0}^{\infty} \sum_{\mathbf{q}\in(\mathbb{Z}^d)^n} \sum_{q\in\mathbb{Z}^d} e^{\sigma|q|} \left| e^{i\beta\cdot(q-\sum_i q_i)}\,u_n(q - \sum_{i=0}^{n} q_i)\,(\hat{X}(q_1),\ldots,\hat{X}(q_n)) \right| \\
&\leq \sum_{n=0}^{\infty} \left(\sum_{q\in\mathbb{Z}^d} e^{(|\mathfrak{Im}\,\beta|+\sigma)|q|} \|u_n(q)\|_{\mathcal{L}({}^n(\mathbb{C}^{d+1});\mathbb{C})} \right) \sum_{\mathbf{q}\in(\mathbb{Z}^d)^n} \prod_{i=1}^{n} |\hat{X}(q_i)| e^{\sigma|q_i|} \\
&\leq Cg^2(r_0^3 + |\epsilon|) \sum_{n=0}^{\infty} \sum_{q\in\mathbb{Z}^d} e^{(|\mathfrak{Im}\,\beta|+\sigma-\rho)|q|}\,(r_0/e)^{-n}\,\|X\|_\sigma^n \leq Cg^2(r_0^3 + |\epsilon|),
\end{aligned}$$

if we choose $|\Im\mathfrak{m}\,\beta| + \sigma < \rho = \eta - r_0$ and $r < r_0/2e$, since $\|X\|_\sigma \le 2r$. Thus, fixing $r = r_0/6$, say, we obtain

$$\sup_{X \in \bar{B}^\Phi_{\sigma,r} \times \bar{B}^\Psi_{\sigma,r}} \|U_\beta(X)\|_\sigma \le C g^2 (r^3 + |\epsilon|) \tag{2.14}$$

whenever

$$|\Im\mathfrak{m}\,\beta| + \sigma + 6r < \eta. \tag{2.15}$$

Lemma 2.4. *Suppose* (2.15) *holds, and* $\Psi \in \bar{B}^\Psi_{\sigma,r}$. *Then, for* r *and* ϵ_0 *small enough,*

$$(\mathcal{D}^2 - g^2)\Phi = U_\beta(\Phi, \Psi)$$

has a solution $\Phi_\beta(\Psi) \in \bar{B}^\Phi_{\sigma,r}$*, real-valued provided* β*,* ϵ*, and* Ψ *are, and there are no other solutions in the* ℓ^1*-ball* $\bar{B}^\Phi_{0,r} \supset \bar{B}^\Phi_{\sigma,r}$*. In fact,* $\Phi_\beta(\Psi) = \tau_\beta \Phi_0(\tau_{-\beta}\Psi)$*. The map* $\Psi \mapsto \Phi_\beta(\Psi)$ *is analytic on* $\bar{B}^\Psi_{\sigma,r}$*.* $\Phi_\beta(\Psi)$ *also depends analytically on* β *as well as on* $(\epsilon, g) \in D$ *(see* (1.15)*), and obeys the bound*

$$\|\Phi_\beta(\Psi)\|_\sigma \le C|\epsilon| \tag{2.16}$$

uniformly in Ψ*,* β*, and* g*.*

Remark 2.5. The smallness condition is $C(r^3 + \epsilon_0) \le r$, where C is the same constant as in (2.14) and contains the norm of the perturbation f.

The standard but lengthy proof of Lemma 2.4 may be found in Subsection 2.3.

Let us come back to equation (2.2), whose right-hand side may now be written solely in terms of $\Psi \in \bar{B}^\Psi_{\sigma,r}$, amounting to

$$\mathcal{D}^2\Psi = V(\Psi) \tag{2.17}$$

with $V(\Psi)(\theta) \equiv \lambda\,\partial_\psi f(\Phi(\Psi)(\theta), \theta + \Psi(\theta))$. Consider then $V_\beta(\Psi) := \tau_\beta V(\tau_{-\beta}\Psi)$. By Lemma 2.4, it reads

$$V_\beta(\Psi)(\theta) \equiv V(\tau_{-\beta}\Psi)(\theta + \beta) \equiv \lambda\,\partial_\psi f(\Phi_\beta(\Psi)(\theta), \theta + \beta + \Psi(\theta))$$

and is analytic in the domain

$$\bar{B}^\Psi_{\sigma,r} \times D \times \{|\Im\mathfrak{m}\,\theta| \le \sigma\} \times \{\beta \mid |\Im\mathfrak{m}\,\beta| + \sigma + 6r < \eta\} \tag{2.18}$$

with the uniform bound

$$\|V_\beta(\Psi)\|_\sigma \le \sup_{|\Im\mathfrak{m}\,\phi|, |\Im\mathfrak{m}\,\psi| \le \eta} |\lambda\,\partial_\psi f(\phi, \psi)| \le C g^2 |\epsilon|,$$

provided $C|\epsilon| \le \eta$ (see (2.16)).

Equation (2.17) is the variational equation corresponding to the action functional

$$S : \bar{B}^\Psi_{\sigma,r} \to \mathbb{R} : \Psi \mapsto S(\Psi) = \int_{\mathbb{T}^d} s(\Psi, \theta)\, d\theta$$

given by the integrand

$$s(\Psi, \theta) = \tfrac{1}{2}(\Phi\mathcal{D}^2\Phi + \Psi \cdot \mathcal{D}^2\Psi) + g^2 \cos\Phi - \lambda f(\Phi, \theta + \Psi),$$

where $\Phi = \Phi(\Psi)$. S is invariant under the $\mathbb{T}^d$-action $\Psi(\theta) \mapsto \Psi_\beta(\theta) := \Psi(\theta + \beta) + \beta$, $\beta \in \mathbb{R}^d$. Hence, $\partial_\beta S(\Psi_\beta)|_{\beta=0} = 0$ yields the *Ward identity*

$$\int_{\mathbb{T}^d} \frac{\delta S(\Psi)}{\delta \Psi^i(\theta)}\, d\theta = \int_{\mathbb{T}^d} \Psi(\theta) \cdot \partial_{\theta^i} \frac{\delta S(\Psi)}{\delta \Psi(\theta)}\, d\theta \quad (i = 1, \dots, d) \tag{2.19}$$

of the symmetry in the functional derivative notation. In fact,

$$\frac{\delta S(\Psi)}{\delta \Psi(\theta)} = (\mathcal{D}^2 \Psi - V(\Psi))(\theta).$$

Integrating by parts three times one sees that

$$\int_{\mathbb{T}^d} \Psi(\theta) \cdot \partial_{\theta^i} \mathcal{D}^2 \Psi(\theta)\, d\theta = -\int_{\mathbb{T}^d} \Psi(\theta) \cdot \partial_{\theta^i} \mathcal{D}^2 \Psi(\theta)\, d\theta = 0.$$

The general identity (2.19) therefore reduces to the identity

$$\int_{\mathbb{T}^d} V^i(\Psi, \theta)\, d\theta = \int_{\mathbb{T}^d} \Psi(\theta) \cdot \partial_{\theta^i} V(\Psi, \theta)\, d\theta \tag{2.20}$$

for the map V.

In conclusion, we have the KAM-type small denominator problem (2.17) with $V_\beta(\Psi, \theta)$ analytic in the domain (2.18) and bounded there by $C|\lambda|$, together with the Ward identity (2.20) stemming from a translation symmetry of the action that generates the equation. Furthermore, $V_\beta(\Psi, \theta)$ is real-valued whenever β, ϵ, and Ψ are. For $0 < \sigma < \eta - 6r$—so that we may choose $\Im\mathfrak{m}\, \beta \neq 0$—this is precisely the setup in [BGK99], where the authors devise a method for dealing with such problems using a Renormalization approach.

The subtle analysis in [BGK99] yields a unique solution $\Psi \in \bar{B}^\Psi_{\sigma,r}$ to (2.17) with zero average and analytic in $(\epsilon, g) \in D$. The inevitable loss of analyticity takes place in the domain of β. The map $\theta \mapsto \Psi(\theta)$ is $\mathbb{R}^d$-valued on the torus for real ϵ and satisfies $\|\Psi\|_\sigma \le C|\lambda| = Cg^2|\epsilon|$.

Denote by Ψ_n, $n \in \mathbb{Z}_+$, the unique solution to (2.17) in the ball $\bar{B}^\Psi_{\sigma/n,r}$. Since $\bar{B}^\Psi_{\sigma,r} \subset \bar{B}^\Psi_{\sigma/n,r}$, Ψ has to coincide with Ψ_n. Hence, Ψ is the unique solution in

$$\bigcup_{n=1}^{\infty} \bar{B}^\Psi_{\sigma/n,r} \supset \left\{ \Psi : \mathbb{T}^d \to \mathbb{R}^d \,\middle|\, \Psi \text{ real-analytic and } \|\Psi\|_{\ell^1} < r \right\}.$$

Indeed, assuming the map $\theta \mapsto \Psi(\theta)$ is real-analytic, $\|\Psi\|_{\sigma/n} < \infty$ for some n, and we have that $\|\Psi\|_{\sigma/n} \searrow \|\Psi\|_0 \equiv \|\Psi\|_{\ell^1}$ as $n \to \infty$. Thus, if $\|\Psi\|_{\ell^1} < r$, we gather that $\|\Psi\|_{\sigma/n} < r$ for sufficiently large values of n.

This concludes the proof of Theorem 2.1. □

2.3. **Proofs of Lemmata 2.3 and 2.4.**

Proof of Lemma 2.3. Write $\|\cdot\| = \|\cdot\|_{\mathcal{L}(^n(\mathbb{C}^{d+1});\mathbb{C})}$ for short. From (2.9) and the Cauchy Integral Theorem,

$$\begin{aligned}\|u_n(q)\| &= \left\|\frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} e^{-iq\cdot(\theta+i\beta)}\,\frac{1}{n!}D^nU(0,\theta+i\beta)\,d\theta\right\| \\ &\le e^{q\cdot\beta}\frac{1}{n!}\sup_{\theta\in\mathbb{T}^d}\|D^nU(0,\theta+i\beta)\|,\end{aligned}$$

for $\beta\in\mathbb{R}^d$ and $|\beta|\le\eta$. Take $0<\rho<\eta$ and choose $\beta=-\rho q/|q|$. We compute the standard norm of n-homogeneous polynomials,

$$\|D^nU(0,\theta+i\beta)\|_{\mathcal{P}^n(\mathbb{C}^{d+1};\mathbb{C})} := \sup_{|z|\le 1}|D^nU(0,\theta+i\beta)\,(z,\dots,z)|,$$

which, using the Cauchy Integral Formula, turns into

$$\sup_{|z|\le 1}\left|\frac{n!}{2\pi i}\oint_{\partial\mathbb{D}(0,r_0)}\frac{U(\zeta z,\theta+i\beta)\,d\zeta}{\zeta^{n+1}}\right| \le n!\,r_0^{-n}\sup_{|z|\le 1}\sup_{\zeta\in\partial\mathbb{D}(0,r_0)}|U(\zeta z,\theta+i\beta)|.$$

Here $\mathbb{D}(0,r_0)$ is the origin-centered circle of radius r_0 in the complex plane, with the constraint $r_0+\rho\le\eta$. For $|z|\le r_0$ and $|\Im\mathfrak{m}\,\theta|\le\rho$ we estimate

$$|U(z,\theta)|\le Cg^2(r_0^3+|\epsilon|);$$

see equation (2.7). Here we have singled out $\lambda g^{-2}=\epsilon$, and C is independent of g. We stress that $U(z,\theta)$ simply stands for the expression obtained from the expression of $U(X,\theta)$ in (2.7) by replacing $X(\theta)$ by $z\in\mathbb{C}^{d+1}$.

Symmetric multilinear maps are fully determined by their diagonal—the corresponding homogeneous polynomial, that is—which is explicitly confirmed by the Polarization Formula [Cha85, Din99]. Hence, in order to obtain the estimate in (2.11), we multiply the corresponding polynomial estimate by the factor $n^n/n!\sim e^n/\sqrt{2\pi n}$. □

Proof of Lemma 2.4. The proof is a simple application of the Banach Fixed Point Theorem. We fix $\Psi\in\bar{B}^{\Psi}_{\sigma,r}$ and study the operator $F(\Phi):=(\mathcal{D}^2-g^2)^{-1}U_\beta(\Phi,\Psi)$.

First, $(\mathcal{D}^2-g^2)^{-1}$ is a linear operator bounded in norm by g^{-2}. From (2.14),

$$\|F(\Phi)\|_\sigma\le g^{-2}\|U_\beta(\Phi,\Psi)\|_\sigma\le C(r^3+|\epsilon|)\le r$$

for sufficiently small r and ϵ, which means that $F(\bar{B}^{\Phi}_{\sigma,r})\subset\bar{B}^{\Phi}_{\sigma,r}$. Proving contractiveness resembles estimating the norm of U_β in the proof of Theorem 2.1, and is omitted. The existence and uniqueness of the solution $\Phi(\Psi,\beta)\in\bar{B}^{\Phi}_{\sigma,r}$ now follow.

For β, ϵ, and Ψ real, F maps the closed subset of real-valued functions $\Phi\in\bar{B}^{\Phi}_{\sigma,r}$ into itself and is a contraction there, so $\Phi(\Psi,\beta)$ is real-valued by uniqueness.

The operator F depends analytically on the parameter Ψ in $\bar{B}^{\Psi}_{\sigma,r}$. Consider the sequence $(F^k(0))_{k\in\mathbb{N}}$ of successive substitutions. Each element $F^k(0)$ is analytic in $\Psi\in\bar{B}^{\Psi}_{\sigma,r}$. Furthermore, the Banach Fixed Point Theorem guarantees that such a sequence converges to the fixed point $\Phi(\Psi,\beta)$ in geometric progression;

$$\|F^k(0)-\Phi(\Psi,\beta)\|_\sigma\le\frac{\mu^n}{1-\mu}\,\|F(0)\|_\sigma<\frac{r\mu^n}{1-\mu}.$$

Consequently, $\Phi(\Psi, \beta)$ is the uniform limit of a sequence of analytic functions, and, as such, analytic itself. The same argument goes for $(\epsilon, g) \in D$ (see (1.15)), as well as for β in the domain specified by (2.15).

Because (2.5) implies $\Psi \in \bar{B}^{\Psi}_{\sigma,r} \subset \bar{B}^{\Psi}_{0,r}$, in fact $\Phi(\Psi, \beta)$ is the unique solution in $\bar{B}^{\Phi}_{0,r}$.

Let us denote $\Phi(\Psi) = \Phi(\Psi, 0)$. If $\Psi \in \bar{B}^{\Psi}_{\sigma,r}$ and $|\Im\mathfrak{m}\,\beta| \le \sigma/2$, then $\tau_{-\beta}\Psi \in \bar{B}^{\Psi}_{\sigma/2,r}$, such that $\Phi = \Phi(\tau_{-\beta}\Psi)$ is the unique element in $\bar{B}^{\Phi}_{\sigma/2,r}$ solving $\Phi = (\mathcal{D}^2 - g^2)^{-1} U(\Phi, \tau_{-\beta}\Psi)$. The diagonality of τ_β and $\mathcal{D}$ yields

$$\Phi_\beta(\Psi) = (\mathcal{D}^2 - g^2)^{-1} U_\beta(\Phi_\beta(\Psi), \Psi),$$

where $\Phi_\beta(\Psi) = \tau_\beta \Phi(\tau_{-\beta}\Psi) \in \bar{B}^{\Phi}_{0,r}$. But $\Phi(\Psi, \beta)$ was the unique solution in $\bar{B}^{\Phi}_{0,r}$, such that $\Phi_\beta(\Psi) = \Phi(\Psi, \beta) \in \bar{B}^{\Phi}_{\sigma,r}$. For larger $|\Im\mathfrak{m}\,\beta|$ one obtains an analytic continuation.

A priori, we know that $\|\Phi_\beta(\Psi)\|_\sigma = \|F(\Phi_\beta(\Psi))\|_\sigma \le r$. On the other hand, we know that $\Phi_\beta(\Psi)|_{\epsilon=0} = 0$ by uniqueness, whence the estimate (2.16) follows. □

3. Lyapunov Exponent—Linearizing the Unstable Manifold

In this section we study the motion in the immediate vicinity of the torus $\mathcal{T}_\lambda$ corresponding to the solution $X_0(\theta)$ of Theorem 2.1. To that end, suppose $X(z, \theta)$ is an analytic solution to equation (1.11) with $X(0, \theta) = X_0(\theta)$. Then $X_1(\theta) := \partial_z X(0, \theta)$ should satisfy the equation

$$(\mathcal{D} + \gamma)^2 X_1 = D\Omega(X_0) X_1, \tag{3.1}$$

as $\Omega(X)(z, \theta)$ depends on z only through X evaluated at (z, θ).

Note that (3.1) is a problem of "eigenvalue type"; recalling $\gamma|_{\epsilon=0} = g$, we will strive to choose $\gamma = \gamma(\epsilon, g)$ in a g-dependent neighbourhood, say

$$|\gamma - g| < g/2, \tag{3.2}$$

of its unperturbed value g, such that (3.1) has a nontrivial solution. That we succeed is the content of Theorem 3.3. Consequently, our γ will depend analytically on ϵ, nicely controlled by $|\gamma - g| < Cg|\epsilon|$.

The subtlety of proving Theorem 3.3 lies in solving a small denominator problem. We go about dealing with it using a Renormalization Group method, treating such small denominators scale by scale. Here we show that the framework of [BGK99] is applicable. The proof, though, is self-contained.

First, view the map $X \mapsto \Omega(X)$ as the map that takes the pair (Φ, Ψ) to $(\Omega_\Phi(\Phi, \Psi), \Omega_\Psi(\Phi, \Psi))$ with the components $\Omega_\Phi(\Phi, \Psi) = g^2 \sin\Phi + \lambda\,\partial_\phi f(\Phi, \theta + \Psi)$ and $\Omega_\Psi(\Phi, \Psi) = \lambda\,\partial_\psi f(\Phi, \theta + \Psi)$. Then the component form of (3.1) reads

$$(\mathcal{D} + \gamma)^2 \begin{pmatrix} \Phi_1 \\ \Psi_1 \end{pmatrix} = \begin{pmatrix} g^2 \cos\Phi_0 + \lambda f_{\phi,\phi} & \lambda f_{\phi,\psi} \\ \lambda f_{\psi,\phi} & \lambda f_{\psi,\psi} \end{pmatrix} \begin{pmatrix} \Phi_1 \\ \Psi_1 \end{pmatrix}. \tag{3.3}$$

In each entry, $f_{a,b}$ stands for the matrix $(\partial_b \partial_a f)(\Phi_0, \theta + \Psi_0)$.

From (3.3) we get for Ψ_1 the equation

$$\Psi_1 = \left[(\mathcal{D} + \gamma)^2 - \lambda f_{\psi,\psi}\right]^{-1} (\lambda f_{\psi,\phi} \Phi_1) =: J\Phi_1, \tag{3.4}$$

Here J is a well-defined bounded linear operator from $\mathcal{B}_\sigma^\Phi$ to $\mathcal{B}_\sigma^\Psi$, provided that ϵ_0 is small. Checking this is straightforward implementation of Neumann series and the fact that the operator $(\mathcal{D}+\gamma)^{-2}$ has the diagonal Fourier kernel

$$(\mathcal{D}+\gamma)^{-2}(p,q) = \delta_{p,q}(i\omega\cdot q+\gamma)^{-2}, \quad p,q\in\mathbb{Z}^d. \tag{3.5}$$

Using (3.2), one obtains the bound

$$\|J\|_{\mathcal{L}(\mathcal{B}_\sigma^\Phi;\mathcal{B}_\sigma^\Psi)} \le C|\epsilon|. \tag{3.6}$$

Remark 3.1. The definition of J is an instance where demanding smallness of $\epsilon := \lambda g^{-2}$ is natural, indeed necessary.

Consequently, using (3.4), we get for Φ_1 the equation

$$[(\mathcal{D}+\gamma)^2 - g^2]\Phi_1 = g^2(\cos\Phi_0 - 1)\Phi_1 + \lambda f_{\phi,\phi}\Phi_1 + \lambda f_{\phi,\psi}J\Phi_1 =: H\Phi_1. \tag{3.7}$$

Recall that $\Phi_0|_{\epsilon=0} = 0$ by Lemma 2.4. Therefore $H|_{\epsilon=0} = 0$, and $\Phi_1|_{\epsilon=0} = 4$ (due to $\Phi^0(z) = 4\arctan z$) is a physically motivated nontrivial solution to (3.7). In other words, the differential operator $(\mathcal{D}+g)^2 - g^2$ is *singular*. On the other hand, when $\epsilon \ne 0$ is small, we know that Φ_0 remains close to zero, making the whole right-hand side in (3.7) small. We then hope to find a Lyapunov exponent γ, close to g, such that $(\mathcal{D}+\gamma)^2 - g^2 - H$ stays singular and the equation still admits a nontrivial solution close to the constant function 4.

It follows from (3.6) that the operator H appearing in (3.7), which lies in $\mathcal{L}(\mathcal{B}_\sigma^\Phi) \equiv \mathcal{L}(\mathcal{B}_\sigma^\Phi;\mathcal{B}_\sigma^\Phi)$, has the useful properties below. The proof comprises Subsection 3.1.

Lemma 3.2. *Denote the kernel of $H \in \mathcal{L}(\mathcal{B}_\sigma^\Phi)$ by $H(p,q)$, $(p,q)\in\mathbb{Z}^d\times\mathbb{Z}^d$. For $|\Im\mathfrak{m}\,\kappa| \le g/3$, there exists an operator $H(\kappa)\in\mathcal{L}(\mathcal{B}_\sigma^\Phi)$ related to H by*

$$(t_s H)(p,q) := H(p+s,q+s) = H(\omega\cdot s;p,q), \quad s\in\mathbb{Z}^d.$$

Let $0<\sigma'<\sigma$. The kernel $H(\kappa;p,q)$ is analytic on

$$\left\{(\kappa,\epsilon,g,\gamma) \;\middle|\; |\Im\mathfrak{m}\,\kappa| \le g/3,\, (\epsilon,g)\in D,\, |\gamma-g|<g/2\right\}$$

and it satisfies the bound

$$|H(\kappa;p,q)| \le Cg^2|\epsilon|\, e^{-\sigma'|p-q|}$$

with $C = C(\sigma')$. As for the κ-derivatives,

$$|H^{(k)}(\kappa;p,q)| \le Ck!\,(g/3 - |\Im\mathfrak{m}\,\kappa|)^{-k} g^2|\epsilon|^2\, e^{-\sigma'|p-q|}, \quad k\ge 1.$$

Moreover,

$$\left|\frac{\partial}{\partial\gamma}H(0;0,0)\right| \le C|\epsilon|^2 g\,\frac{1}{1-2|\gamma-g|/g}.$$

3.1. **Proof of Lemma 3.2.**

Proof. To simplify notations, we decompose

$$H = H_1 + H_2 \quad \text{with} \quad H_2 = \lambda f_{\phi,\psi} J.$$

Let Φ and Ψ be arbitrary functions in the spaces $\mathcal{B}_\sigma^\Phi$ and $\mathcal{B}_\sigma^\Psi$, respectively.

H_1 acts as ordinary multiplication: $H_1\Phi(\theta) = H_1(\theta)\Phi(\theta)$ with $H_1(\theta) \in \mathbb{C}$. We write $\widehat{H}_1$ for the Fourier transform of the map $\theta \mapsto H_1(\theta)$. Denoting a kernel element of the operator H_1 by $H_1(p,q)$, we have $H_1(p,q) \equiv \widehat{H}_1(p-q)$. We gather that

$$t_s H_1 = H_1 \tag{3.8}$$

holds, and that the kernel of H_1 satisfies

$$|H_1(p,q)| \le C|\lambda|\, e^{-\sigma|p-q|}, \quad p,q \in \mathbb{Z}^d.$$

Here $\sigma > 0$ is the width of the analyticity strip around the real $\mathbb{T}^d$ of the map $\theta \mapsto H_1(\theta)$, *i.e.*, of X_0. Since, by Theorem 2.1, X_0 is analytic with respect to $(\epsilon, g) \in D$, so is $H_1(p,q)$.

Observe that the expression defining J in (3.4) may be cast as

$$J\Phi = \left[\mathbb{1} - (\mathcal{D}+\gamma)^{-2}(\lambda f_{\psi,\psi})\right]^{-1} (\mathcal{D}+\gamma)^{-2}(\lambda f_{\psi,\phi}\Phi) = B\Lambda O\Phi,$$

where B, Λ, and O stand for $[\mathbb{1} - (\mathcal{D}+\gamma)^{-2}(\lambda f_{\psi,\psi})]^{-1}$, $(\mathcal{D}+\gamma)^{-2}$, and $\lambda f_{\psi,\phi}$, respectively. Assuming each index a and b in $f_{a,b}$ stands either for ϕ or ψ, the reader should bear in mind that $f_{a,b}$ refers to the multiplication operator corresponding to the Jacobian matrix $(\partial_b\partial_a f)(\Phi_0, \theta + \Psi_0)$. Its Fourier kernel reads $f_{a,b}(p,q) = \hat{f}_{a,b}(p-q)$, whence the translation invariance

$$t_s f_{a,b} = f_{a,b}. \tag{3.9}$$

Denoting $\Lambda(q) \equiv \Lambda(q,q) \equiv (i\omega\cdot q + \gamma)^{-2}$, we are interested in the kernel

$$J(p,q) = \sum_{r\in\mathbb{Z}^d} B(p,r)\Lambda(r)O(r,q), \quad p,q \in \mathbb{Z}^d, \tag{3.10}$$

of J. We shall also need the “shifted version” of $\Lambda(q)$,

$$\Lambda(\kappa; q) := (i\omega\cdot q + i\kappa + \gamma)^{-2}, \quad \kappa \in \mathbb{C}. \tag{3.11}$$

It is related to $\Lambda(q)$ by the property

$$t_s\Lambda(q) = \Lambda(\omega\cdot s; q). \tag{3.12}$$

Further, $\Lambda(\kappa;q)$ is analytic on $\{\kappa \mid |\Im\mathfrak{m}\,\kappa| \le g/3\} \times \{\gamma \mid |\gamma - g| < g/2\}$ and satisfies

$$|\Lambda(\kappa;q)| \le 36 g^{-2}. \tag{3.13}$$

Equation (3.13) also means that the operator $\Lambda(\kappa)$ corresponding to the kernel in (3.11) belongs to $\mathcal{L}(\mathcal{B}_\sigma)$ with $\|\Lambda(\kappa)\|_{\mathcal{L}(\mathcal{B}_\sigma)} \le 36g^{-2}$. Interpreting $f_{a,b}$ as a multiplication operator, $\|f_{a,b}\|_{\mathcal{L}(\mathcal{B}_\sigma)} \le \|f_{a,b}\|_\sigma$ shows that $B, O \in \mathcal{L}(\mathcal{B}_\sigma)$.

As in the case of H_1, O acts as multiplication by a real-analytic function whose modulus is bounded by $C|\lambda|$. Thus, we estimate

$$|O(p,q)| \le C|\lambda|\, e^{-\sigma|p-q|} \quad \text{and} \quad |\Lambda(p)O(p,q)| \le C|\epsilon|\, e^{-\sigma|p-q|}. \tag{3.14}$$

Bounding the kernel of B calls for the Neumann series

$$B = \sum_{k=0}^{\infty} B_k, \quad \text{with} \quad B_k := (\lambda\, \Lambda f_{\psi,\psi})^k\,. \tag{3.15}$$

Clearly $\|B_k\|_{\mathcal{L}(\mathcal{B}_\sigma)} \le (C|\epsilon|)^k$ and $|B_k(p,q)| \le (C|\epsilon|)^k$ such that, by Fubini's Theorem,

$$\widehat{B\Psi}(p) = \sum_{k=0}^{\infty} \widehat{B_k\Psi}(p) = \sum_{q\in\mathbb{Z}^d}\sum_{k=0}^{\infty} B_k(p,q)\hat{\Psi}(q). \tag{3.16}$$

The expression of B_k contains $k-1$ products of the operator $\lambda\,\Lambda f_{\psi,\psi}$ with itself, which appear as convolutions in terms of Fourier transforms. Explicitly,

$$B_k(p,q) = \lambda^k \sum_{q_i\in\mathbb{Z}^d} \Lambda(p)\hat{f}_{\psi,\psi}(p-q_1)\cdots\Lambda(q_{k-1})\hat{f}_{\psi,\psi}(q_{k-1}-q). \tag{3.17}$$

Using the bound $|\Lambda(p)\hat{f}_{\psi,\psi}(q)| \le Cg^{-2}\, e^{-\sigma|q|}$ we see that, for $0<\sigma'<\sigma$,

$$|B_k(p,q)| \le \left(Cg^{-2}|\lambda|\right)^k e^{-\sigma'|p-q|} \sum_{q_i\in\mathbb{Z}^d} e^{-(\sigma-\sigma')(|p-q_1|+\cdots+|q_{k-1}-q|)} \le (C|\epsilon|)^k\, e^{-\sigma'|p-q|}.$$

Thus, choosing ϵ appropriately small we make the geometric series arising in (3.16) convergent and obtain

$$|B(p,q)|,\, |J(p,q)| \le C\, e^{-\sigma'|p-q|}$$

with the aid of (3.14) in (3.10). Finally,

$$|H_2(p,q)| \le Cg^2|\epsilon|^2 e^{-\sigma'|p-q|}. \tag{3.18}$$

Exploiting (3.9), we compute

$$t_s H_2 = \lambda\, t_s(f_{\phi,\psi}J) = \lambda\, f_{\phi,\psi}\, t_s J = \lambda\, f_{\phi,\psi}\,(t_s B)(t_s\Lambda)O. \tag{3.19}$$

With the aid of (3.15) and (3.12), $t_s B_k = \lambda^k\,[\Lambda(\omega\cdot s)f_{\psi,\psi}]^k$. Thus, $(t_sB_k)(p,q)$ depends on s only through $\omega\cdot s$. Moreover, the dependence on $\omega\cdot s$ is analytic in a neighbourhood of the real line: Consider the shifted quantity

$$B_k(\kappa;p,q) := \lambda^k \sum_{q_i\in\mathbb{Z}^d} \Lambda(\kappa;p)\hat{f}_{\psi,\psi}(p-q_1)\cdots\Lambda(\kappa;q_{k-1})\hat{f}_{\psi,\psi}(q_{k-1}-q),$$

which for $\kappa=\omega\cdot s$ becomes $(t_sB_k)(p,q)$. The summand above is analytic on

$$D_g := \{\epsilon \mid |\epsilon|<\epsilon_0\} \times \{\kappa \mid |\Im\mathfrak{m}\,\kappa| \le g/3\} \times \{\gamma \mid |\gamma - g| < g/2\},$$

and the sum converges uniformly, as is readily observed after recalling the bound (3.13) on $\Lambda(\kappa;q)$ and looking at the estimation of $|B_k(p,q)|$. Thus, $B_k(\kappa;p,q)$ is analytic. But the Neumann series $\sum_{k=0}^{\infty} B_k(\kappa;p,q)$ also converges uniformly, making the limit $B(\kappa;p,q)$ analytic on D_g. Evidently, $(t_sB)(p,q) = B(\omega\cdot s;p,q)$. The kernel $B_k(\kappa;p,q)$ defines an operator $B(\kappa)$. Motivated by equation (3.19), we extend the definition of H_2 and set $H_2(\kappa) := \lambda f_{\phi,\psi}B(\kappa)\Lambda(\kappa)O$. Using (3.13), a straightforward computation shows that also $H_2(\kappa;p,q)$ obeys (3.18) and is analytic on D_g. Furthermore,

$$(t_sH_2)(p,q) = H_2(\omega\cdot s;p,q).$$

Recalling the translation invariance (3.8) of H_1, we simply take $H(\kappa) := H_1 + H_2(\kappa)$.

The bound on the derivative $H^{(k)}(\kappa;p,q)$ is achieved by a Cauchy estimate. To that end, one observes $H'(\kappa) = H_2'(\kappa)$ and uses the bound (3.18) on D_g. Similarly, because X_0 is independent of γ, $\partial H/\partial\gamma = \partial H_2/\partial\gamma$, and we get the bound on $\partial H(0;0,0)/\partial\gamma$.

The constants above are independent of g, as long as $0 < g < g_0$. That is to say, the estimates hold on $\bigcup_{0<g<g_0} D_g = \big\{(\kappa,\epsilon,g,\gamma) \;\big|\; |\Im\mathfrak{m}\,\kappa| \le g/3,\ (\epsilon,g)\in D,\ |\gamma - g| < g/2\big\}$. □

3.2. **Linearized invariant manifolds: rudiments of renormalization.** We now proceed to stating the main theorem of this section, discussing the linearization X_1. Our proof is based on a Renormalization Group (RG) technique we present below.

Theorem 3.3. *Under the assumptions of Theorem 1, there exist a number ϵ_0 and a map $\gamma = \gamma(\epsilon, g)$ on D, analytic in ϵ, with $|\gamma - g| \le Cg|\epsilon|$, such that equation* (3.1) *has a nontrivial solution X_1 which is*

(1) *analytic in $|\epsilon| < \epsilon_0$ and*
(2) *analytic in θ in a complex neighbourhood $\mathcal{U}$ of $\mathbb{T}^d$,*

and satisfies the physical constraint

$$\Phi_1\big|_{\epsilon=0} \equiv 4 = \langle\Phi_1\rangle.$$

Furthermore, it is real-valued if ϵ and θ are real, and

$$\sup_{\theta\in\mathcal{U}} |\Psi_1(\theta)| \le C|\epsilon| \quad \textit{and} \quad \sup_{\theta\in\mathcal{U}} |\Phi_1(\theta) - 4| \le Cg|\epsilon|.$$

The map γ is independent of $\langle\Psi_0\rangle$. If X_1 and X_1' correspond to X_0 and X_0' of Theorem 2.1, respectively, with $\langle\Psi_0\rangle = 0$ and $\langle\Psi_0'\rangle = \beta \in \mathbb{R}^d$, then

$$X_1'(\theta) \equiv X_1(\theta + \beta).$$

Remark 3.4. We chose the normalization 4, because $X^0(z,\theta) = (4\arctan z, 0)$ is the unperturbed solution (separatrix) and $\arctan z = z + \mathcal{O}(z^3)$.

Remark 3.5. The pair (γ, X_1) of Theorem 3.3 is unique in the sense that it is the only one making *our construction* work, which is manifested by Lemma 3.11 below. We do not prove the uniqueness of X_1. However, for a given solution X_1 the value of γ is unique: If γ' is another one, (3.1) yields $(\mathcal{D}+\gamma)^2 X_1 = (\mathcal{D}+\gamma')^2 X_1$ because $D\Omega(X_0)$ is independent of γ. This shows that $\gamma' = \gamma$, because $\hat{\Phi}_1(0) = 4 \ne 0$.

Let us commence sketching the backbone of Theorem 3.3 by recalling equation (3.7). We expand the square on the left-hand side and obtain

$$(\mathcal{D}^2 + 2\gamma\mathcal{D})\Phi_1 = (H + g^2 - \gamma^2)\Phi_1. \tag{3.20}$$

For a small $\epsilon \ne 0$, Φ_1 should remain close to the unperturbed value $4 = \partial_z\Phi^0(0)$. Due to the linearity of (3.20) such a solution may be normalized as $\langle\Phi_1\rangle = 4$. Thus, we set

$$\Phi_1(\theta) = 4 + \xi(\theta), \tag{3.21}$$

where we *demand* the function $\xi : \mathbb{T}^d \to \mathbb{R}$ to vanish on the average, *i.e.*,

$$\hat{\xi}(0) = 0. \tag{3.22}$$

Plugging (3.21) into (3.20) results in

$$(\mathcal{D}^2 + 2\gamma\mathcal{D})\xi = \pi_0(\xi + 4), \quad \text{where} \quad \pi_0 := H + g^2 - \gamma^2.$$

After switching into Fourier representation, this reads

$$\hat{\xi}(q) = G(q)\Big[\sum_{p\in\mathbb{Z}^d} \pi_0(q,p)\hat{\xi}(p) + \hat{\rho}_0(q)\Big] \quad \text{if } q \in \mathbb{Z}^d \setminus \{0\}, \tag{3.23}$$

$$0 = \sum_{p\in\mathbb{Z}^d} \pi_0(0,p)\hat{\xi}(p) + \hat{\rho}_0(0), \tag{3.24}$$

where ρ_0 is a function defined through its Fourier transform by setting

$$\hat{\rho}_0(q) := 4\pi_0(q,0). \tag{3.25}$$

The symbol $G(q)$ stands for the diagonal element $G(q,q)$ of the operator G whose Fourier kernel is given by

$$G(p,q) := \delta_{p,q}\begin{cases} (2i\gamma\, \omega\cdot q - (\omega\cdot q)^2)^{-1} & \text{if } q \in \mathbb{Z}^d \setminus \{0\}, \\ 0 & \text{if } q = 0. \end{cases} \tag{3.26}$$

The matter of the fact is that, in terms of our new notations, any solution ξ of

$$\xi = G(\pi_0\xi + \rho_0) \tag{3.27}$$

also solves (3.23); only the zero mode constraint (3.22) has been included here. After finding such a ξ, we go on to show that it is a solution to (3.24), as well.

As is apparent from the definition of G, this problem involves arbitrarily small denominators $\omega \cdot q$. Our strategy is to recursively decompose G into parts, each of which corresponds to denominators up to a given order of magnitude. We then end up solving "partial problems" of (3.27) scale by scale, and show that these solutions converge to a true solution of (3.27) as the recursion proceeds and smaller and smaller denominators become dealt with.

Leaving the all-important scaling parameter $\aleph \in \,]0,1[$ to be decided later[2], we shall need the entire functions

$$\chi_n : \mathbb{C} \to \mathbb{C} : \chi_n(\kappa) = \begin{cases} e^{-(\aleph^{-n}\kappa)^6} & \text{if } n \in \mathbb{Z}_+, \\ 1 & \text{if } n = 0. \end{cases}$$

Their importance lies in the fact that the sequence $(\chi_n - \chi_{n+1})_{n\in\mathbb{N}}$ of functions is an analytic partition of unity on $\mathbb{R} \setminus \{0\}$; on this set $0 \le 1 - \chi_N \nearrow 1$ pointwise, as $N \to \infty$. Some of the first members of the sequence appear plotted in Figure 3. The number 6 in the exponent is a choice of convenience; it is the one used in [BGK99].

Let us now introduce the diagonal operators G_n and Γ_n, $n \in \mathbb{N}$, defined by

$$G_n(q) := \chi_n(\omega\cdot q)G(q) \quad \text{and} \quad \Gamma_n := G_n - G_{n+1},$$

respectively. Observe that $G_0 = G$ and $G_n(0) = 0$. The point here is that in $\Gamma_n(q)$ the functions $\chi_n(\omega\cdot q) - \chi_{n+1}(\omega\cdot q)$ act as cutoffs for the values of $\omega\cdot q$. Each Γ_n deals with

[2] Aleph, $\aleph$, is the first letter in the Hebrew alphabet.

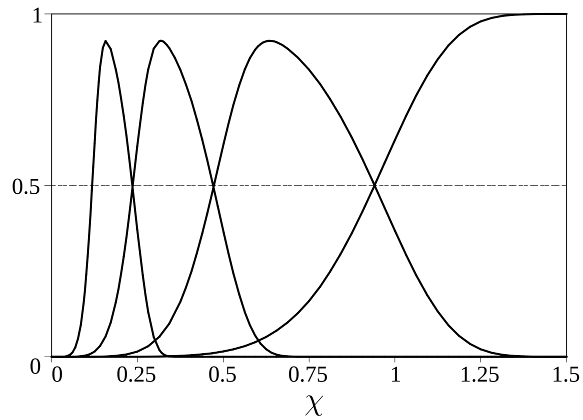


FIGURE 3. Graphs of $\chi_n - \chi_{n+1}$ with $n = 0, 1, 2, 3$, and $\aleph = \frac{1}{2}$. The maxima are located roughly at $\aleph^n$.

the denominators $\omega \cdot q$ that are roughly of order $\aleph^n$ and, intuitively,

$$\Gamma_{<n} := \sum_{k=0}^{n-1} \Gamma_k = G - G_n \tag{3.28}$$

gets closer and closer to G as n tends to infinity. Instead of the full equation (3.27), consider the easier, approximate problem

$$x_n = \Gamma_{<n}(\pi_0 x_n + \rho_0), \tag{3.29}$$

obtained by replacing G with $\Gamma_{<n}$. It is easier since $\Gamma_{<n}$ discards the most dangerous ones of the small denominators. However, its solution should become a better and better approximation of the solution of (3.27) with increasing n.

Having $G_0 = G_1 + \Gamma_0$, we decompose $\xi = \xi_1 + \eta_0$ and assume that $\eta_0 = \eta_0(\xi_1)$ solves the "large denominator problem"

$$\eta_0 = \Gamma_0(\pi_0(\xi_1 + \eta_0) + \rho_0). \tag{3.30}$$

Then, solving the original problem (3.27) for ξ amounts to solving

$$\xi_1 = G_1(\pi_0(\xi_1 + \eta_0) + \rho_0) \tag{3.31}$$

for ξ_1.

Assuming $\mathbb{1} - \Gamma_0\pi_0$ is invertible[3], we can extract η_0 out of (3.30) and get

$$\eta_0 = (\mathbb{1} - \Gamma_0\pi_0)^{-1}\Gamma_0(\pi_0\xi_1 + \rho_0). \tag{3.32}$$

Therefore, (3.31) transforms into

$$\xi_1 = G_1(\mathbb{1} - \pi_0\Gamma_0)^{-1}(\pi_0\xi_1 + \rho_0)$$

with the aid of the identities

$$\pi_0(\mathbb{1} - \Gamma_0\pi_0)^{-1} = (\mathbb{1} - \pi_0\Gamma_0)^{-1}\pi_0 \quad \text{and} \quad (\mathbb{1} - \pi_0\Gamma_0)^{-1}\pi_0\Gamma_0 = (\mathbb{1} - \pi_0\Gamma_0)^{-1} - \mathbb{1}.$$

Thus, defining the new objects

$$\pi_1 := (\mathbb{1} - \pi_0\Gamma_0)^{-1}\pi_0 \quad \text{and} \quad \rho_1 := (\mathbb{1} - \pi_0\Gamma_0)^{-1}\rho_0,$$

[3]Think of Γ_0 as comprising only large denominators and π_0 being proportional to ϵ.

we obtain

$$\eta_0 = \Gamma_0(\pi_1\xi_1 + \rho_1)$$

and

$$\xi_1 = G_1(\pi_1\xi_1 + \rho_1). \tag{3.33}$$

Indeed, equation (3.33) has *precisely the same form* as the original problem (3.27). Now, relaxing the assumption that η_0 be *a priori* known, suppose we are able to solve (3.33), and take (3.32) as the definition of η_0, instead. Then the solution of the full problem is recovered using the simple relation

$$\xi = \xi_1 + \eta_0 = (\mathbb{1} - \Gamma_0\pi_0)^{-1}(\xi_1 + \Gamma_0\rho_0).$$

Owing to the aforementioned formal covariance between equations (3.27) and (3.33), we may iterate the construction above. Thus, in general, solving

$$\xi_{n+1} = G_{n+1}(\pi_{n+1}\xi_{n+1} + \rho_{n+1}) \tag{3.34}$$

for ξ_{n+1} with the definitions

$$\pi_{n+1} := (\mathbb{1} - \pi_n\Gamma_n)^{-1}\pi_n, \tag{3.35}$$

$$\rho_{n+1} := (\mathbb{1} - \pi_n\Gamma_n)^{-1}\rho_n, \tag{3.36}$$

$$\eta_n := \Gamma_n(\pi_{n+1}\xi_{n+1} + \rho_{n+1}), \tag{3.37}$$

produces $\xi_n = \xi_{n+1} + \eta_n$, or

$$\xi_n = (\mathbb{1} - \Gamma_n\pi_n)^{-1}\left(\xi_{n+1} + \Gamma_n\rho_n\right) \tag{3.38}$$

for the solution of $\xi_n = G_n(\pi_n\xi_n + \rho_n)$.

Equations (3.38) and (3.36) reveal through

$$\pi_n\xi_n + \rho_n = \pi_n\left[(\mathbb{1} - \Gamma_n\pi_n)^{-1}\xi_{n+1} + \Gamma_n\rho_{n+1}\right] + (\mathbb{1} - \pi_n\Gamma_n)\rho_{n+1}$$

the recursion invariance

$$\pi_0\xi_0 + \rho_0 = \pi_1\xi_1 + \rho_1 = \cdots = \pi_n\xi_n + \rho_n = \cdots \tag{3.39}$$

in our construction.

Let us tidy up the notation by giving the definitions

$$v_n(y) \equiv \pi_n y + \rho_n \quad \text{and} \quad f_n := \mathbb{1} + \Gamma_{<n}v_n \quad \text{with} \quad \Gamma_{<0} = 0. \tag{3.40}$$

In particular, (3.39) takes the form $v_n(\xi_n) = v_0(\xi_0)$. We also set

$$\Xi_n(y) \equiv (\mathbb{1} - \Gamma_n\pi_n)^{-1}\left(y + \Gamma_n\rho_n\right), \tag{3.41}$$

such that (3.38) reads $\xi_n = \Xi_n(\xi_{n+1})$, and (3.39) reduces to

$$v_{n+1} = v_n \circ \Xi_n. \tag{3.42}$$

The latter is a convenient way of writing $v_{n+1} = (\mathbb{1} - \pi_n\Gamma_n)^{-1}v_n$. Notice also that Ξ_n is formally invertible.

One easily verifies

$$\Xi_n = \mathbb{1} + \Gamma_n v_{n+1}. \tag{3.43}$$

As a consequence,

$$f_{n+1} = f_n \circ \Xi_n. \tag{3.44}$$

Since $f_0 = \mathbb{1}$, we have the cumulative formula

$$f_n = \Xi_0 \circ \Xi_1 \circ \cdots \circ \Xi_{n-1}. \tag{3.45}$$

Hence, a similar expansion of (3.42) implies

$$v_n = v_0 \circ f_n.$$

Inserting here the definition of f_n, we get

$$v_n = v_0 \circ (\mathbb{1} + \Gamma_{<n} v_n). \tag{3.46}$$

Proposition 3.6. *Let $\xi_n = \Xi_n(\xi_{n+1})$. If ξ_{n+1} satisfies $\xi_{n+1} = G_{n+1}(\pi_{n+1}\xi_{n+1} + \rho_{n+1})$, then ξ_n satisfies $\xi_n = G_n(\pi_n\xi_n + \rho_n)$, and vice versa.*

Proof. Suppose $\xi_{n+1} = G_{n+1}v_{n+1}(\xi_{n+1})$. By $G_n = G_{n+1} + \Gamma_n$ and (3.42),

$$G_n v_n \circ \Xi_n = G_{n+1}v_{n+1} - \mathbb{1} + \mathbb{1} + \Gamma_n v_{n+1}.$$

But, with the aid of (3.43), this transforms into

$$(G_n v_n - \mathbb{1}) \circ \Xi_n = G_{n+1}v_{n+1} - \mathbb{1}.$$

As Ξ_n is invertible with $\xi_n = \Xi_n(\xi_{n+1})$, the identity above proves the formal equivalence of the small denominator problems (3.34), or $G_n v_n(\xi_n) = \xi_n$, with differing n. □

Recalling (3.45), we immediately arrive at

Corollary 3.7. *If $\xi_n = G_n(\pi_n\xi_n + \rho_n)$, then*

$$\xi_0 := f_n(\xi_n) = \xi_n + \Gamma_{<n}v_n(\xi_n)$$

solves the complete problem: $\xi_0 = G_0(\pi_0\xi_0 + \rho_0)$.

Remark 3.8. The solution ξ_0 above comprises two terms having clear interpretations. The first term, ξ_n, solves the small denominator problem, namely $\xi_n = G_n(\pi_n\xi_n + \rho_n)$, at the nth step. The second term, $\Gamma_{<n}v_n(\xi_n)$, on the other hand, consists of the sum

$$\eta_{<n}(\xi_n) := \sum_{k=0}^{n-1} \eta_k(\xi_{k+1}) \quad \text{with} \quad \xi_{k+1} = (\Xi_{k+1} \circ \cdots \circ \Xi_{n-1})\,(\xi_n),$$

where $\eta_k = \eta_k(\xi_{k+1})$ solves the large denominator problem $\eta_k = \Gamma_k v_k(\xi_{k+1} + \eta_k)$ in analogy with (3.30). Indeed, $\Gamma_k v_k(\xi_{k+1} + \eta_k) = \Gamma_k v_k(\xi_k) = \Gamma_k v_{k+1}(\xi_{k+1}) = \eta_k$.

Finally, we make a crucial observation. If we operate on (3.46) by $\Gamma_{<n}$ and set

$$x_n := f_n(0) = \Gamma_{<n}v_n(0), \tag{3.47}$$

we solve the approximate problem (3.29):

$$x_n = \Gamma_{<n}(\pi_0 x_n + \rho_0).$$

We shall demonstrate that the approximate solutions x_n form a Cauchy sequence in a simple Banach space, and that their limit

$$\xi := \lim_{n\to\infty} x_n \tag{3.48}$$

solves the original equation (3.27).

Just to motivate the above discussion, think of an abstract map $\mathcal{R}_n$ that takes (π_n, ρ_n, G_n) to $(\pi_{n+1}, \rho_{n+1}, G_{n+1})$. The recursion scheme

$$\xi = G(\pi_0\xi + \rho_0) \overset{\mathcal{R}_0}{\mapsto} \xi_1 = G_1(\pi_1\xi_1 + \rho_1) \overset{\mathcal{R}_1}{\mapsto} \cdots \overset{\mathcal{R}_{n-1}}{\mapsto} \xi_n = G_n(\pi_n\xi_n + \rho_n) \overset{\mathcal{R}_n}{\mapsto} \cdots$$

is called *renormalization* of the problem, and $\mathcal{R}_n$ is the corresponding renormalization transformation. Then, in view of Proposition 3.6, it remains for one to demonstrate that this process "converges", in order to be able to solve the original equation $\xi = G(\pi_0\xi + \rho_0)$. That is to say, one wishes that the *renormalization flow* of the triplet (π_0, ρ_0, G_0), $(\pi_n, \rho_n, G_n) = (\prod_{k=0}^{n-1} \mathcal{R}_k)(\pi_0, \rho_0, G_0)$, in a sense tends to a fixed point (π^*, ρ^*, G^*) of some limiting operator "$\mathcal{R}_\infty = \lim_{k\to\infty} \mathcal{R}_k$" as $n \to \infty$, and that the equation

$$\xi^* = G^*(\pi^*\xi^* + \rho^*) \tag{3.49}$$

is well-defined and solvable.

In our case $G^*\rho^* = 0$, such that the equation is linear and possesses the trivial solution $\xi^* = 0$. Corollary 3.7 then throws light on why (3.48) should solve (3.27); $f_n(\xi_n)$ solves it, and ξ_n approaches zero with increasing n. Therefore, it is fair to expect that also $\lim_{n\to\infty} f_n(0)$ is a solution.

3.3. **Banach spaces.** Technically speaking, we need to control the renormalization flow (3.35)–(3.37) by estimating the kernel elements of Γ_n and π_n, for the operators $\mathbb{1} - \pi_n\Gamma_n$ and $\mathbb{1} - \Gamma_n\pi_n$ had better be invertible between suitable spaces. Such Banach spaces will be defined in this subsection.

We begin by analyzing the properties of the operators Γ_n. *A priori*, one expects the most significant contribution to arise from such q's that $\omega \cdot q = \mathcal{O}(\aleph^n)$, due to the cutoff $\chi_n - \chi_{n+1}$ in the definition of these operators. Therefore, (3.26) implies

$$|\Gamma_n(q)| = \mathcal{O}(g^{-1}\aleph^{-n}). \tag{3.50}$$

More accurately, it is fairly easy to obtain

$$|\chi_n(\kappa) - \chi_{n+1}(\kappa)| \le C|\aleph^{-n}\kappa|^\ell \begin{cases} e^{-\frac{1}{2}|\aleph^{-n}\kappa|^6} & \text{if } n \ge 1, \\ 1 & \text{if } n = 0, \end{cases} \tag{3.51}$$

for $\ell = 0, 1, \ldots, 6$, in a strip $|\mathfrak{Im}\, \kappa| < \aleph^n b$. C only depends on the parameter b. Pay attention to the fact that $G(q)$, which was defined in (3.26), only depends on q through $\omega \cdot q$. Therefore, it is handy to introduce the analytic function

$$\iota : \mathbb{C} \setminus \{0, 2i\gamma\} \to \mathbb{C} : \iota(\kappa) = (2i\gamma\kappa - \kappa^2)^{-1}.$$

In particular, $\iota(\omega \cdot q) = G(q)$ for $q \neq 0$. This motivates the further definition

$$\Gamma_n(\kappa; p, q) := \delta_{p,q} \begin{cases} [\chi_n(\omega \cdot q + \kappa) - \chi_{n+1}(\omega \cdot q + \kappa)]\iota(\omega \cdot q + \kappa) & \text{if } \omega \cdot q + \kappa \neq 0, \\ 0 & \text{if } \omega \cdot q + \kappa = 0. \end{cases}$$

The importance of the resulting operator $\Gamma_n(\kappa)$ is based on the possibility of viewing $\omega \cdot q$ as a complex "variable":

$$\Gamma_n(q, q) = \Gamma_n(0; q, q) = \Gamma_n(\omega \cdot q; 0, 0).$$

We shall often write $\Gamma_n(\kappa; q)$ instead of the complete $\Gamma_n(\kappa; q, q)$.

Imposing the condition $b \leq g/2$ on b we get within $|\Im\mathfrak{m}\,\kappa| < \aleph^n b$ that

$$\begin{aligned} |\Gamma_n(\kappa; q)| &= \aleph^{-n} \frac{|\chi_n(\omega \cdot q + \kappa) - \chi_{n+1}(\omega \cdot q + \kappa)|}{|\aleph^{-n}(\omega \cdot q + \kappa)|} \, |(\omega \cdot q + \kappa)\, \iota(\omega \cdot q + \kappa)| \\ &\leq C_\Gamma \, g^{-1} \aleph^{-n} \min(1, |\aleph^{-n}(\omega \cdot q + \kappa)|^5) \begin{cases} e^{-\frac{1}{2}|\aleph^{-n}(\omega \cdot q + \kappa)|^6} & \text{if } n \geq 1, \\ 1 & \text{if } n = 0, \end{cases} \end{aligned} \tag{3.52}$$

making use of (3.51). In particular, we have confirmed the heuristic estimate (3.50).

Now to the spaces promised. Ultimately the solution of (3.27), namely ξ (and therefore Φ_1) will live in the space $\mathcal{B}^{\Phi}_{\alpha^*} \subset \ell^1(\mathbb{Z}^d; \mathbb{C})$ for a sufficiently small width α^* of the analyticity strip—see Subsection 2.1. The following weights will come in handy:

$$w_n(q) := \begin{cases} e^{\aleph^{-n}|\omega \cdot q|} & \text{if } n \geq 1, \\ 1 & \text{if } n = 0. \end{cases} \tag{3.53}$$

We extend these to negative indices by setting $w_{-n}(q) \equiv w_n(q)^{-1}$.

Definition 3.9 (Spaces h_n)**.** For $n \in \mathbb{Z}$, let

$$\|\xi\|_n := \sum_{q \in \mathbb{Z}^d} |\hat{\xi}(q)| w_n(q).$$

These norms induce the Banach spaces h_n. Observe that h_0 is the space $\ell^1(\mathbb{Z}^d; \mathbb{C})$.

Notice that our weights satisfy

$$w_{n+1}(q)^{\aleph} = w_n(q) \quad \text{and} \quad w_n(q) \geq 1 \qquad (n \geq 1). \tag{3.54}$$

The spaces at hand thus realize the embedding hierarchy

$$h_{n+1} \subset h_n \qquad (n \in \mathbb{Z})$$

due to the trivial inequalities

$$\| \cdot \|_n \leq \| \cdot \|_{n+1} \qquad (n \in \mathbb{Z}). \tag{3.55}$$

Operator norms $\| \cdot \|_{\mathcal{L}(h_n; h_m)}$ between such spaces h_n and h_m will be denoted by $\| \cdot \|_{n;m}$ for short. We actually have,

$$\|L\|_{n;m} = \sup_{q \in \mathbb{Z}^d} \sum_{p \in \mathbb{Z}^d} |L(p, q)| w_m(p) w_{-n}(q) \qquad (m, n \in \mathbb{Z}). \tag{3.56}$$

Either from this or from (3.55) by the Schwarz inequality one infers

$$\| \cdot \|_{n+1;m} \leq \| \cdot \|_{n;m} \leq \| \cdot \|_{n;m+1} \qquad (m, n \in \mathbb{Z})$$

such that the operator spaces satisfy

$$\mathcal{L}(h_n; h_{m+1}) \subset \mathcal{L}(h_n; h_m) \subset \mathcal{L}(h_{n+1}; h_m) \qquad (m, n \in \mathbb{Z}).$$

Moreover, the Schwarz inequality implies the useful bounds

$$\|L_1 L_2\|_{n;m} \leq \|L_1\|_{l;m} \|L_2\|_{n;l} \qquad (l, m, n \in \mathbb{Z}). \tag{3.57}$$

From now on, n will always assume *nonnegative* values. Define the domain

$$D_n := \left\{ \kappa \in \mathbb{C} \,\middle|\, |\kappa| < \aleph^n b \right\}, \tag{3.58}$$

recalling that $b \le g/2$. Then (3.52) easily validates the bounds

$$\|\Gamma_n(\kappa)\|_{-n;n} \le C_\Gamma\, g^{-1}\aleph^{-n} \tag{3.59}$$

for $\kappa \in D_n$, where the (new) constant C_Γ is independent of κ and g as long as $g < g_0$. This shows, in particular, that

$$\Gamma_n(\kappa) \in \mathcal{L}(h_{-n}; h_n) \subset \mathcal{L}(h_0; h_0).$$

Remark 3.10. The weights $w_n(q)$ arise as follows. The diagonal kernel of Γ_n is strongly concentrated around small denominators $\omega \cdot q$ of order $\aleph^n$; for large $\omega \cdot q$ the value of $\Gamma_n(q)$ is very close to zero, *but not quite equal to zero.* Therefore, in an expression such as $\widehat{\Gamma_n \xi}(q) = \Gamma_n(q)\hat{\xi}(q)$ we cannot let $|\hat{\xi}(q)|$ be *arbitrarily* large for large values of $\omega \cdot q$. This "tail" can be of the order of $w_n(q) = e^{\aleph^{-n}|\omega\cdot q|}$, say, which amounts to $\xi \in h_{-n}$.

It has to be emphasized that having the same power of $\aleph^{-n}$ and $|\omega \cdot q|$ in $w_n(q)$ is crucial, which can be read off from (3.52). This way $\omega \cdot q$ "scales" as $\aleph^n$ in all estimates in the nth step of the iteration.

The motivation for introducing the spaces h_n, on the other hand, comes from the fact that in the recursion (3.35) the domain of π_n will shrink. So, in the norms $\|\cdot\|_n$ we incorporate a weight that increases as n grows. It is a matter of convenience to use the inverse of the weight $w_n(q)^{-1}$ appearing in $\|\cdot\|_{-n}$.

3.4. **Renormalization made rigorous: estimates and the Lyapunov exponent.**

The rest of this section is devoted to demonstrating that the renormalization flow of π_n in (3.35) is controlled in the norms $\|\cdot\|_{n;-n}$ such that the products $\|\pi_n\|_{n;-n}\|\Gamma_n\|_{-n;n}$ are small, so as to make the recursion formulae (3.35)–(3.37) well-defined through Neumann series. Recalling (3.59), the task roughly amounts to making sure that $\|\pi_n\|_{n;-n}$ decays at least as rapidly as $\aleph^n$ with increasing n.

According to Lemma 3.2, $\pi_0 = H + g^2 - \gamma^2 \in \mathcal{L}(\mathcal{B}^\Phi_\sigma)$ can be written as

$$\pi_0(p,q) = p_0(\omega \cdot q)\delta_{p,q} + \tilde{\pi}_0(p,q),$$

where $\tilde{\pi}_0$ vanishes on the diagonal, and in the first term

$$p_0(\kappa) := \delta_0 + \bar{p}_0(\kappa), \qquad \bar{p}_0(0) = 0,$$

depends analytically on κ, as long as $|\mathfrak{Im}\,\kappa| \le g/3$; explicitly $\delta_0 = H(0;0,0) + g^2 - \gamma^2$ and $\bar{p}_0(\kappa) = H(\kappa;0,0) - H(0;0,0)$.

Similarly, we split π_n into its diagonal and off-diagonal parts:

$$\pi_n(p,q) = p_n(\omega \cdot q)\delta_{p,q} + \tilde{\pi}_n(p,q), \qquad \tilde{\pi}_n(q,q) = 0,$$

with

$$p_n(\kappa) = \delta_n + \bar{p}_n(\kappa), \qquad \bar{p}_n(0) = 0.$$

The possibility of doing this follows from the computation

$$t_s\pi_0 = t_sH + g^2 - \gamma^2 = H(\omega\cdot s) + g^2 - \gamma^2 =: \pi_0(\omega \cdot s)$$

and its recursive consequence

$$t_s\pi_{n+1} = (\mathbb{1} - \pi_n(\omega\cdot s)\,\Gamma_n(\omega\cdot s))^{-1}\,\pi_n(\omega\cdot s) =: \pi_{n+1}(\omega\cdot s).$$

Motivated by the computation above, let us inductively define the maps

$$\pi_{n+1,\beta}(\kappa) := (\mathbb{1} - \pi_{n\beta}(\kappa)\,\Gamma_n(\kappa))^{-1}\,\pi_{n\beta}(\kappa), \qquad \kappa \in D_n, \quad |\mathfrak{Im}\,\beta| < \alpha_n, \tag{3.60}$$

starting at

$$\pi_{0\beta}(\kappa) := P_0(\kappa) + \tilde{\pi}_{0\beta}(\kappa), \qquad \kappa \in D_0, \quad |\Im\mathfrak{m}\,\beta| < \alpha_0,$$

by setting $b \le g/3$ in (3.58). Here

$$\begin{aligned} P_0(\kappa; p, q) &:= p_0(\kappa + \omega\cdot q)\delta_{p,q}, \\ \tilde{\pi}_{0\beta}(\kappa; p, q) &:= e^{i\beta\cdot(p-q)} H(\kappa; p, q)(1 - \delta_{p,q}), \end{aligned}$$

and, with σ' coming from Lemma 3.2,

$$\alpha_{n+1} := \left(1 - \frac{4}{(n+3)^2}\right)\alpha_n, \qquad \alpha_0 < \sigma'. \tag{3.61}$$

In particular, Eric Weisstein's World of Mathematics [Wei] tells us that

$$\alpha_n \searrow \alpha_0 \cdot \prod_{k=3}^{\infty}\left(1 - \frac{4}{k^2}\right) = \frac{\alpha_0}{6} > 0 \qquad \text{as} \qquad n \to \infty. \tag{3.62}$$

As far as notation is concerned, we may omit β if it equals zero: $\pi_n(\kappa) \equiv \pi_{n0}(\kappa)$, and so forth. By a straightforward induction argument,

$$\pi_{n\beta}(\kappa; p, q) := e^{i\beta\cdot(p-q)}\pi_n(\kappa; p, q),$$

such that β does not enter the diagonal of $\pi_{n\beta}$. Of course,

$$|\pi_{n\beta}(\kappa; p, q)| = e^{-\Im\mathfrak{m}\,\beta\cdot(p-q)}|\pi_n(\kappa; p, q)|. \tag{3.63}$$

For clarity, set

$$P_n(\kappa) := \delta_n \mathbb{1} + \bar{P}_n(\kappa) \quad \text{with} \quad \bar{P}_n(\kappa; p, q) := \bar{p}_n(\kappa + \omega\cdot q)\delta_{p,q},$$

so that we may express the operator $\pi_{n\beta}(\kappa)$ itself, without reference to its kernel, as

$$\pi_{n\beta}(\kappa) = P_n(\kappa) + \tilde{\pi}_{n\beta}(\kappa) = \delta_n + \bar{P}_n(\kappa) + \tilde{\pi}_{n\beta}(\kappa), \qquad \delta_n \equiv \delta_n\mathbb{1},$$

for short. This decomposition satisfies

$$\|\pi_{n\beta}(\kappa)\|_{n;-n} \le |\delta_n| + \|\bar{P}_n(\kappa)\|_{n;-n} + \|\tilde{\pi}_{n\beta}(\kappa)\|_{n;-n}. \tag{3.64}$$

It will turn out that the sum in (3.64) is finite if $\kappa \in D_n$ and $|\Im\mathfrak{m}\,\beta| < \alpha_n$—indeed very small, as we are trying to prove—meaning that $\pi_{n\beta}(\kappa) \in \mathcal{L}(h_n; h_{-n})$.

The crux of analyzing the renormalization flow is the following lemma, for which we provide an inductive proof later on in this section. The reader is advised to take the result as granted for now.

Lemma 3.11 (Modified Lyapunov exponent controls the flow)**.** *Set $b = g/3$ and $\aleph = \min(\frac{1}{8}, b^2)$. There exist constants $c_\gamma > 0$, $C > 0$, $c > 0$, $\mu > 1$, and a unique Lyapunov exponent γ satisfying*

$$|\gamma - g| < c_\gamma|\epsilon|g \tag{3.65}$$

such that, for any $n \in \mathbb{N}$, the bounds

$$\|\tilde{\pi}_{n\beta}(\kappa)\|_{n;-n} \leq C|\epsilon|g \begin{cases} g & \text{if } n = 0, \\ \aleph^n e^{-c\mu^n} & \text{if } n \geq 1, \end{cases} \tag{3.66}$$

$$\|\bar{P}_n(\kappa)\|_{n;-n} \leq C|\epsilon|g \begin{cases} g & \text{if } n = 0, \\ \aleph^n & \text{if } n \geq 1, \end{cases} \tag{3.67}$$

$$|\delta_n| \leq C|\epsilon|g \begin{cases} g & \text{if } n = 0, \\ \aleph^{2n} & \text{if } n \geq 1, \end{cases} \tag{3.68}$$

hold true for $(\epsilon, g) \in D$, $\kappa \in D_n$ and $|\Im\mathfrak{m}\,\beta| < \alpha_n$. Moreover, c is bounded away from zero and $\mu \to \infty$ in the limit $g \to 0$.

Remark 3.12. The sole purpose of introducing the complex variable κ is to go about proving the bound (3.67) on the *diagonal* part of π_n. We use analyticity in κ and restrict the latter to a domain of ever decreasing size.

The *possibility* of including the complex parameter β in the analysis, on the other hand, facilitates proving exponential decay of $\pi_n(\kappa; p, q)$ in the quantity $|p - q|$. This is sufficiently rapid for obtaining the bound (3.66) on the *off-diagonal* part of π_n. Also the analyticity strip of β around $\mathbb{R}$ is taken narrower and narrower upon iteration, but no narrower than a certain limit $(\alpha_0/6)$.

Corollary 3.13. *The bounds of Lemma 3.11 imply*

$$\|\pi_{n\beta}(\kappa)\|_{n;-n} \leq C|\epsilon|g \begin{cases} g & \text{if } n = 0 \\ \aleph^n & \text{if } n \geq 1, \end{cases}$$

The caveat to get around in the proof of Lemma 3.11 is that δ_n is reluctant to go to zero along the recursion. To change the state of affairs, we fine-tune the Lyapunov exponent γ such that also $\delta_n \to 0$ as $n \to \infty$. As stated in the lemma, there turns out to exist precisely one such value of γ. This is what ultimately enables us to prove the convergence of our renormalization scheme, consequently validating Theorem 3.3 discussing the linearized solution X_1. For the sake of continuity, we first give the simple proof of Theorem 3.3 and only then prove Lemma 3.11.

3.5. **Proof of Theorem 3.3.** With x_n as in (3.47), the task is to show that the limiting function ξ—see (3.48)—is an analytic solution to (3.27).

Given the formal definition $y_\beta := \tau_\beta y$, (3.47) implies

$$x_{n\beta} = f_{n\beta}(0).$$

Recalling (3.44) and (3.45), one clearly has

$$f_{n+1,\beta} = f_{n\beta} \circ \Xi_{n\beta} \quad \text{and} \quad f_{n\beta} = \Xi_{0\beta} \circ \Xi_{1\beta} \circ \cdots \circ \Xi_{n-1,\beta}.$$

Hence, the recursion relation

$$x_{n+1,\beta} = x_{n\beta} + \big(f_{n\beta}(\Xi_{n\beta}(0)) - f_{n\beta}(0)\big)$$

follows. Here (3.41) extends to

$$\Xi_{n\beta}(y) \equiv (\mathbb{1} - \Gamma_n \pi_{n\beta})^{-1}(y + \Gamma_n \rho_{n\beta}). \tag{3.69}$$

Notice that the flows of $\rho_{n\beta}$ and $4\pi_{n\beta}(\,\cdot\,,0)$ [4] are identical. Furthermore, the initial conditions agree according to (3.25), such that

$$\hat{\rho}_{n\beta}(q) \equiv 4\pi_{n\beta}(q,0).$$

Because $D\Xi_{n\beta}(y) \equiv (\mathbb{1} - \Gamma_n \pi_{n\beta})^{-1}$, the chain rule reveals

$$Df_{n\beta}(y) \equiv (\mathbb{1} - \Gamma_0\pi_{0\beta})^{-1}(\mathbb{1} - \Gamma_1\pi_{1\beta})^{-1}\cdots(\mathbb{1} - \Gamma_{n-1}\pi_{n-1,\beta})^{-1}.$$

Recursive implementation of Corollary 3.13 in the form

$$\|(\mathbb{1} - \Gamma_n\pi_{n\beta})^{-1}\|_{n;n-1} \leq \|(\mathbb{1} - \Gamma_n\pi_{n\beta})^{-1}\|_{n;n} \leq 2$$

implies that $Df_{n\beta}(y) \in \mathcal{L}(h_{n-1}; h_0)$ with $\sup_{y\in h_n}\|Df_{n\beta}(y)\|_{n-1;0} \leq 2^n$. By the Mean-Value Theorem we go on to estimate

$$\|x_{n+1,\beta} - x_{n\beta}\|_0 \leq 2^n \|\Xi_{n\beta}(0)\|_n \tag{3.70}$$

with the aid of the inequality $\|\cdot\|_{n-1} \leq \|\cdot\|_n$.

Lemma 3.14. *For parameters as in Lemma 3.11 and ϵ_0 small, we may perceive $\Xi_{n\beta}$ as an analytic map from h_n to $h_n \subset h_{n-1}$ with*

$$\|\Xi_{n\beta}(0)\|_n \leq C|\epsilon| \begin{cases} g & \text{if } n = 0, \\ e^{-c\mu^n} & \text{if } n \geq 1. \end{cases}$$

Proof. Since Γ_n annihilates the zero mode ($\Gamma_n(0) = 0$), $\Gamma_n\rho_{n\beta} = 4(\Gamma_n\tilde{\pi}_{n\beta})(\,\cdot\,,0)$, which is super-exponentially small in the norm $\|\cdot\|_n$ by $\|\tilde{\pi}_{n\beta}(\,\cdot\,,0)\|_{-n} \leq \|\tilde{\pi}_{n\beta}\|_{n;-n}$ and Lemma 3.11. According to (3.69), Lemma 3.14 clearly holds if we take ϵ small enough so as to validate $\|\Gamma_n\|_{-n;n}\|\pi_{n\beta}\|_{n;-n} \leq \frac{1}{2}$, say, for each n. For the bounds on $\pi_{n\beta}$ and Γ_n we refer the reader to Corollary 3.13 and (3.59), respectively. □

By Lemma 3.14, $f_{n\beta}$ maps h_{n-1} to h_0, confirming that $x_{n\beta} \in h_0$ for each n. Coming back to (3.70) and taking $|\mathfrak{Im}\,\beta| < \alpha^* := \alpha_0/6$ (see (3.62)), the sequence $(x_{n\beta})_{n\in\mathbb{N}}$ is Cauchy in the Banach space h_0. Moreover, $x_{0\beta} = 0$ gives us

$$\|\xi_\beta\|_0 \leq \sum_{n=0}^{\infty} \|x_{n+1,\beta} - x_{n\beta}\|_0 \leq C|\epsilon|g,$$

where the factor g is due to the last statement in Lemma 3.11. It is implied that

$$|\hat{\xi}(q)| \leq C|\epsilon|g\, e^{-\alpha^*|q|} \qquad (q \in \mathbb{Z}^d).$$

We infer that ξ is real-analytic on $\mathbb{T}^d$.

Recalling that $\lim_{n\to\infty}\Gamma_{<n}(q) = G(q)$ for each $q \in \mathbb{Z}^d$, let us take the *pointwise* limit $n \to \infty$ of (3.27) in the Fourier representation: $\hat{\xi}(q)$ equals

$$\lim_{n\to\infty}\Gamma_{<n}(q)\left(\widehat{\pi_0 x_n}(q) + \hat{\rho}_0(q)\right) = G(q)\lim_{n\to\infty}\left(\widehat{\pi_0 x_n}(q) + \hat{\rho}_0(q)\right) = G(q)\big(\widehat{\pi_0\xi} + \hat{\rho}_0\big)(q),$$

because π_0 is a continuous operator on h_0. Indeed, ξ solves (3.27)!

Out of curiosity, we conclude by the recursion invariance (3.39) that

$$\xi_n = G_n(\pi_n\xi_n + \rho_n) = G_n(\pi_0\xi + \rho_0) \tag{3.71}$$

[4] $\pi_{n\beta}(\,\cdot\,,0)$ is shorthand for the function $\theta \mapsto \sum_q e^{iq\cdot\theta}\pi_{n\beta}(q,0)$.

converges to $\xi^* = 0$, pointwise in terms of the Fourier representation. Hence, equation (3.34) really trivializes in the large-n-limit. Another way of seeing this is the pointwise bound $|\widehat{G_n\rho_n}(q)| \leq C|G(q)|\|\tilde{\pi}_n\|_{n;-n}$, which tends to zero and paraphrases $G^*\rho^* = 0$ below (3.49).

We still need to demonstrate that the solution ξ of (3.23) also solves (3.24), *i.e.*, that

$$\big(\widehat{\pi_0\xi} + \hat{\rho}_0\big)(0) = \pi_0(0,\,\cdot\,)\hat{\xi} + \hat{\rho}_0(0) = \sum_{p\in\mathbb{Z}^d} \pi_0(0,p)\hat{\xi}(p) + \hat{\rho}_0(0) = 0.$$

From (3.71), $G_n(q) := \chi_n(\omega\cdot q)G(q)$, $G(0) = 0$, and (3.27),

$$\begin{aligned}
\big|\widehat{\pi_n\xi_n}(0)\big| &\leq \sum_{q\in\mathbb{Z}^d} |\tilde{\pi}_n(0,q)|\chi_n(\omega\cdot q)\,\big|G(q)\big(\widehat{\pi_0\xi} + \hat{\rho}_0\big)(q)\big| \\
&\leq \|\xi\|_0 \sup_{q\in\mathbb{Z}^d} |\tilde{\pi}_n(0,q)|\chi_n(\omega\cdot q) \\
&\leq \|\xi\|_0\, \|\tilde{\pi}_n\|_{n;-n} \sup_{q\in\mathbb{Z}^d\setminus\{0\}} w_n(q)\chi_n(\omega\cdot q) \\
&\leq C\|\xi\|_0\, \|\tilde{\pi}_n\|_{n;-n} \quad\longrightarrow\quad 0 \quad \text{as} \quad n\to\infty,
\end{aligned}$$

Thus,

$$\big(\widehat{\pi_0\xi} + \hat{\rho}_0\big)(0) = \lim_{n\to\infty} \big(\widehat{\pi_n\xi_n} + \hat{\rho}_n\big)(0) = \lim_{n\to\infty} \hat{\rho}_n(0).$$

But

$$\lim_{n\to\infty} \hat{\rho}_n(0) = 4\lim_{n\to\infty} \pi_n(0,0) = 4\lim_{n\to\infty} \delta_n = 0,$$

and we are done with the construction of (γ, X_1) under the assumption $\langle\Psi_0\rangle = 0$.

The case $\langle\Psi_0\rangle \neq 0$. If X_1 solves (3.1), it is a matter of applying the translation τ_β on both sides of the equation to get $(\mathcal{D}+\gamma)^2 X_1' = D\Omega(X_0')X_1'$, where $X_0' = \tau_\beta X_0 + (0,\beta)$ and $X_1' = \tau_\beta X_1$. In other words, the translation property in the formulation of the theorem holds, and the value of γ does not change under such translations. □

3.6. **Proof of Lemma 3.11.** We begin by deriving several identities that are easy to refer to below. To this end, let us look at the flow (3.60) more closely, observing that we may formally split

$$\big(\mathbb{1} - \pi_{n\beta}(\kappa)\Gamma_n(\kappa)\big)^{-1} = \left(\mathbb{1} - P_n(\kappa)\Gamma_n(\kappa)\right)^{-1} + r_{n\beta}(\kappa).$$

The remainder $r_{n\beta}(\kappa)$ reads explicitly

$$r_{n\beta}(\kappa) := \left(\mathbb{1} - \pi_{n\beta}(\kappa)\Gamma_n(\kappa)\right)^{-1}\, \tilde{\pi}_{n\beta}(\kappa)\Gamma_n(\kappa)\, \left(\mathbb{1} - P_n(\kappa)\Gamma_n(\kappa)\right)^{-1}.$$

In fact, this quantity is asymptotically *very small* in $\mathcal{L}(h_{-n}; h_{-n})$ due to the explicit factor $\tilde{\pi}_{n\beta}$; given the bounds (3.66)–(3.68) for some particular value of n,

$$\|r_{n\beta}(\kappa)\|_{-n;-n} \leq C|\epsilon|e^{-c\mu^n}. \tag{3.72}$$

Continuing abstractly, (3.60) becomes

$$\pi_{n+1,\beta}(\kappa) = \left(\mathbb{1} - P_n(\kappa)\Gamma_n(\kappa)\right)^{-1} P_n(\kappa) + s_n(\kappa) + \tilde{s}_{n\beta}(\kappa),$$

where $s_n(\kappa)$ is the diagonal and $\tilde{s}_{n\beta}(\kappa)$ the off-diagonal part of the small remainder term $r_{n\beta}(\kappa)\pi_{n\beta}(\kappa)$, respectively. Therefore, the diagonal $P_n(\kappa)$—containing the problematic δ_n—and the off-diagonal $\tilde{\pi}_{n\beta}(\kappa)$ iterate according to the rules

$$\begin{cases} P_{n+1}(\kappa) = \left(\mathbb{1} - P_n(\kappa)\Gamma_n(\kappa)\right)^{-1} P_n(\kappa) + s_n(\kappa), \\ \tilde{\pi}_{n+1,\beta}(\kappa) = \tilde{s}_{n\beta}(\kappa). \end{cases} \tag{3.73}$$

Notice that $s_n(\kappa)$ is indeed free of β, because each $P_n(\kappa)$ is.

By construction, $\delta_n = \pi_{n\beta}(0;0,0) = P_n(0;0)$ for each n, such that the diagonality of $(\mathbb{1} - P_n(0)\Gamma_n(0))^{-1}$ with $\Gamma_n(0;0) = 0$ implies that changes in δ_n upon iteration only arise from the small term s_n in (3.73):

$$\delta_{n+1} = \delta_n + d_n, \qquad d_n := s_n(0;0). \tag{3.74}$$

But $r_{n\beta}(0;0,0) = 0$, again because $\Gamma_n(0;0) = 0$, such that

$$d_n = s_n(0;0) = \left(r_n\pi_n\right)(0;0,0) = (r_n\tilde{\pi}_n)(0;0,0). \tag{3.75}$$

We remind the reader of our convention of dropping one of the kernel indices of diagonal operators. For instance, $s_n(\kappa;q) \equiv s_n(\kappa;q,q)$.

It is convenient to spell out a consequence of (3.73):

$$\bar{P}_{n+1}(\kappa) = \bar{P}_n(\kappa) + P_n(\kappa)\Gamma_n(\kappa)(\mathbb{1} - P_n(\kappa)\Gamma_n(\kappa))^{-1}P_n(\kappa) + (s_n(\kappa) - d_n). \tag{3.76}$$

Proof of Lemma 3.11. Here we finally prove that the bounds (3.66)–(3.68), such that (3.73)—and indeed everything above—becomes not only formally justified. To this end, we proceed by induction. As iterating (3.66) and (3.67) is rather easy, the proof boils down to *choosing the value of our free parameter*, the Lyapunov exponent γ, so as to guarantee that δ_n satisfies (3.68) at each step.

(i) Case $n = 0$**.** Consider κ restricted to D_0 with $b \le g/3$. Lemma 3.2 and $\bar{P}_0(\kappa;q) = H(\kappa;q,q) - H(0;0,0)$ readily imply

$$\|\bar{P}_0(\kappa)\|_{0;0} \le C_0|\lambda|.$$

Furthermore, increasing C_0 and employing (3.63) with $|\Im\mathfrak{m}\,\beta| < \alpha_0 < \sigma'$,

$$\|\tilde{\pi}_{0\beta}(\kappa)\|_{0;0} \le C_0|\lambda|.$$

The leading Taylor coefficient $\bar{p}_0'(0) = H'(0;0,0)$ and the corresponding remainder of the function $\bar{p}_0 = H(\,\cdot\,;0,0) - H(0;0,0)$ satisfy

$$|\bar{p}_0'(0)| \le \tfrac{1}{4}C_0|\epsilon|^2 g \quad \text{and} \quad |\bar{p}_0(\kappa) - \bar{p}_0'(0)\kappa| \le \tfrac{1}{2}C_0|\epsilon|^2|\kappa|^2,$$

taking C_0 large enough.

Assume that γ lies in the open g-centered disk of radius $c|\epsilon|g$:

$$\gamma \in I_\gamma := \mathbb{D}(g, c_\gamma|\epsilon|g). \tag{3.77}$$

Recall that $\delta_0 = \epsilon g^2 u(\epsilon,g,\gamma) + g^2 - \gamma^2$, where $\epsilon g^2 u(\epsilon,g,\gamma) = H(0;0,0)$. If $\delta_0(\gamma_1) = \delta_0(\gamma_2)$ and we denote $\gamma_i = g(1+x_i)$, the Mean-Value Theorem yields

$$|\gamma_1 - \gamma_2| \le \tfrac{1}{2}\big(|x_1 + x_2| + |\epsilon|g\|\partial_\gamma u\|_\infty\big)|\gamma_1 - \gamma_2|.$$

By Lemma 3.2, $\|\partial_\gamma u\|_\infty \le C|\epsilon|g^{-1}/(1 - 2c_\gamma|\epsilon|)$, and $|x_1 + x_2| < 2c_\gamma|\epsilon|$. For a sufficiently small $|\epsilon|$, we gather $\gamma_1 = \gamma_2$, such that $\gamma \mapsto \delta_0$ is one-to-one on I_γ. Moreover, the image of the disk I_γ contains the disk $\mathbb{D}\big(0, (2c_\gamma - c_\gamma^2|\epsilon| - \|u\|_\infty)|\epsilon|g^2\big)$. Thus, for a sufficiently large value of c_γ and small value of ϵ, there exists a *closed* set $J_0 \subset I_\gamma$ which $\gamma \mapsto \delta_0$ maps analytically and *bijectively* onto the closed disk

$$I_0 := \bar{\mathbb{D}}(0, C_0|\epsilon|g^2).$$

We are about to prove below that a correct choice of γ leads to

$$\delta_n \in I_n := \bar{\mathbb{D}}(0, C_0|\epsilon|g\,\aleph^{2n})$$

for each and every $n \in \mathbb{Z}_+$.

(ii) Induction step: hypotheses. Fix $n \in \mathbb{N}$. Suppose

$$\|\tilde{\pi}_{n\beta}(\kappa)\|_{n;-n} \leq C_n|\epsilon|g\,\aleph^n \begin{cases} g & \text{if } n = 0, \\ e^{-c\mu^n} & \text{if } n \geq 1, \end{cases}$$

for some constants $c > 0$ and $\mu > 1$—to be fixed later—and

$$\|\bar{P}_n(\kappa)\|_{n;-n} \leq C_n|\epsilon|g\,\aleph^n \begin{cases} g & \text{if } n = 0, \\ 1 & \text{if } n \geq 1, \end{cases}$$

hold true for $|\epsilon| < \epsilon_n$, $|\Im\mathfrak{m}\,\beta| < \alpha_n$, and $\kappa \in D_n$. Suppose there exists a closed set $J_n \subset I_\gamma$ and a bijective analytic map $\Delta_n : J_n \to I_n : \gamma \mapsto \delta_n$.

Further, let the kernel elements of these operators be analytic in D_n and continuous in the closure $\bar{D}_n$. Also the estimates

$$|\bar{p}_n'(0)| \leq \left(1 - \frac{1}{n+2}\right)\frac{1}{2}C_n g \begin{cases} |\epsilon|^2 & \text{if } n = 0, \\ |\epsilon|^{3/2} & \text{if } n \geq 1, \end{cases} \tag{3.78}$$

and

$$|\bar{p}_n(\kappa) - \bar{p}_n'(0)\kappa| \leq \frac{1}{2}C_n\aleph^{-n}|\kappa|^2 \begin{cases} |\epsilon|^2 & \text{if } n = 0, \\ |\epsilon|^{3/2} & \text{if } n \geq 1, \end{cases}, \tag{3.79}$$

which facilitate dealing with the Taylor expansion of $\bar{p}_n$, are supposed to be satisfied.

In particular, it follows from (3.64), $b \leq g/3$, and the inductive hypotheses that

$$|p_n(\kappa)| \leq B_n C_n|\epsilon|g\,\aleph^n \quad \text{with} \quad B_n := \begin{cases} b|\epsilon| + g & \text{if } n = 0, \\ b|\epsilon|^{1/2} + \aleph^n & \text{if } n \geq 1, \end{cases} \tag{3.80}$$

and

$$\|\pi_{n\beta}(\kappa)\|_{n;-n} \leq A_n C_n|\epsilon|g\,\aleph^n, \tag{3.81}$$

where

$$A_n := \begin{cases} g & \text{if } n = 0, \\ 1 + \aleph^n + e^{-c\mu^n} & \text{if } n \geq 1. \end{cases} \tag{3.82}$$

The strategy is to iterate the above hypotheses and prove that, in the bitter end, C_n and ϵ_n can be chosen in an n-independent fashion, *uniformly in g*.

(ii a) The off-diagonal $\tilde{\pi}_{n+1,\beta}(\kappa)$. If $\tilde{\beta} \in \mathbb{C}^d$, then

$$|\tilde{\pi}_{n+1,\tilde{\beta}}(\kappa; p, q)|\, e^{-\Im\mathfrak{m}\,(\beta-\tilde{\beta})\cdot(p-q)} w_n(p)^{-1} w_n(q)^{-1} \leq \|\tilde{\pi}_{n+1,\beta}(\kappa)\|_{n;-n}.$$

But with a modification of (3.72),

$$\|r_{n\beta}(\kappa)\|_{-n;-n} \leq 4C_\Gamma C_n|\epsilon|\tilde{B}_n \quad \text{where} \quad \tilde{B}_n := \begin{cases} g & \text{if } n = 0, \\ e^{-c\mu^n} & \text{if } n \geq 1, \end{cases} \tag{3.83}$$

such that

$$\|\tilde{\pi}_{n+1,\beta}(\kappa)\|_{n;-n} \leq \|r_{n\beta}(\kappa)\pi_{n\beta}(\kappa)\|_{n;-n} \leq 2C_n|\epsilon|g\,\aleph^n\tilde{B}_n$$

both provided ϵ meets the condition

$$|\epsilon| \leq \epsilon_{n+1} := \max\left(\epsilon_n, \tfrac{1}{2}(A_nC_nC_\Gamma)^{-1}\right). \tag{3.84}$$

Hence, if $|\mathfrak{Im}\,\beta| < \alpha_n$,

$$|\tilde{\pi}_{n+1,\tilde{\beta}}(\kappa; p, q)| \leq 2C_n|\epsilon|g\,\aleph^n\tilde{B}_n \cdot (1 - \delta_{p,q})\,e^{\mathfrak{Im}\,(\beta-\tilde{\beta})\cdot(p-q)}w_n(p)w_n(q).$$

Now assume $|\mathfrak{Im}\,\tilde{\beta}| < \alpha_{n+1}$ and, fixing p and q, take

$$\beta = \tilde{\beta} + i(\alpha_n - \alpha_{n+1})\frac{p-q}{|p-q|}.$$

Obviously $|\mathfrak{Im}\,\beta| < \alpha_n$. What we get this way is

$$|\tilde{\pi}_{n+1,\tilde{\beta}}(\kappa; p, q)| \leq 2C_n|\epsilon|g\,\aleph^n\tilde{B}_n \cdot (1 - \delta_{p,q})\,e^{-(\alpha_n-\alpha_{n+1})|p-q|}w_n(p)w_n(q)$$

for each pair $(p, q) \in \mathbb{Z}^d \times \mathbb{Z}^d$. Thus, from the expression (3.56) for the norm,

$$\begin{aligned} &\|\tilde{\pi}_{n+1,\tilde{\beta}}(\kappa)\|_{n+1;-(n+1)} \\ &\quad\leq 2C_n|\epsilon|g\,\aleph^n\tilde{B}_n \sup_{q\in\mathbb{Z}^d} \sum_{p\in\mathbb{Z}^d\setminus\{q\}} e^{-4(n+3)^{-2}\alpha_n|p-q|}\frac{w_n(p)w_n(q)}{w_{n+1}(p)w_{n+1}(q)} \\ &\quad\leq 2C_n|\epsilon|g\,\aleph^n\tilde{B}_n \sum_{p\in\mathbb{Z}^d\setminus\{0\}} e^{-4(n+3)^{-2}\alpha_n|p|}w_{n+1}(p)^{-(1-\aleph)}. \end{aligned} \tag{3.85}$$

After (3.54), the second inequality follows from shifting p to $p+q$. We control the above bound by treating the cases $|\omega \cdot p| \leq \aleph^{(n+1)/2}$ and $|\omega \cdot p| > \aleph^{(n+1)/2}$ separately. In fact, if $|\omega \cdot p| \leq \aleph^{(n+1)/2}$, then $|p| > \aleph^{-(n+1)/2\nu}$ follows from (1.13), and

$$e^{-4(n+3)^{-2}\alpha_n|p|} < e^{-2n^{-2}\alpha_n|p|} \cdot e^{-2(n+1)^{-2}\alpha_n\aleph^{-(n+1)/2\nu}}, \quad w_{n+1}(p)^{-(1-\aleph)} < 1,$$

whereas

$$|\omega \cdot p| > \aleph^{(n+1)/2} \quad \Longrightarrow \quad w_{n+1}(p)^{-(1-\aleph)} < e^{-(1-\aleph)\aleph^{-(n+1)/2}}.$$

Since $\alpha_n > \alpha_0/6$ by (3.62) and, for $a > 1$ and $m > 0$, $m^{-2}a^m \geq \frac{e^2}{4}(\ln a)^2$, we have

$$e^{-2(n+1)^{-2}\alpha_n\aleph^{-(n+1)/2\nu}} \leq e^{-\frac{1}{12}e^2\alpha_0\ln(\aleph^{-1/4\nu})\,\aleph^{-(n+1)/4\nu}}.$$

The remaining d-dimensional geometric series satisfies

$$\sum_{p\in\mathbb{Z}^d\setminus\{0\}} e^{-2(n+1)^{-2}\alpha_0|p|} \leq C(d)\left(\frac{(n+1)^2}{\alpha_0}\right)^d.$$

Hence, we infer that if $|\mathfrak{Im}\,\beta| < \alpha_{n+1}$ and $\kappa \in D_n$, then

$$\|\tilde{\pi}_{n+1,\beta}(\kappa)\|_{n+1;-(n+1)} \leq C_{n+1}|\epsilon|g\,\aleph^{n+1}e^{-c\mu^{n+1}},$$

where we finally pin down the values of the previously free parameters

$$c := \tfrac{1}{2}\min\left(\tfrac{1}{12}e^2\alpha_0\ln(\aleph^{-1/4\nu}),1-\aleph\right) > 0 \quad\text{and}\quad \mu := \aleph^{-1/\max(4\nu,2)} > 1,$$

and take

$$C_{n+1} \geq 2C(d)\,\aleph^{-1}e^{-c\mu^{n+1}}\tilde{B}_n\left(\frac{(n+1)^2}{\alpha_0}\right)^d C_n. \tag{3.86}$$

(ii b.1) The non-constant part $\bar{P}_{n+1}(\kappa)$ of the diagonal. If $\kappa\in D_{n+1}$ and $|\omega\cdot q|<\aleph^n(1-\aleph)b$, then $\kappa+\omega\cdot q\in D_n$. So, by (3.54),

$$\begin{aligned}\|\bar{P}_{n+1}(\kappa)\|_{n+1;-(n+1)} &= \sup_{q\in\mathbb{Z}^d}|\bar{P}_{n+1}(\kappa;q)|w_{n+1}(q)^{-2}\\ &\leq \max\left\{\sup_{|\omega\cdot q|<\aleph^n(1-\aleph)b}\frac{|\bar{p}_{n+1}(\kappa+\omega\cdot q)|}{w_{n+1}(q)^2},\ \sup_{|\omega\cdot q|\geq\aleph^n(1-\aleph)b}|\bar{P}_{n+1}(\kappa;q)|w_n(q)^{-2/\aleph}\right\}\\ &\leq \max\left\{\sup_{|\omega\cdot q|<\aleph^n(1-\aleph)b}\frac{|\bar{p}_{n+1}(\kappa+\omega\cdot q)|}{w_{n+1}(q)^2},\quad e^{-2b\aleph^{-1}(1-\aleph)^2}\|\bar{P}_{n+1}(\kappa)\|_{n;-n}\right\}.\end{aligned}$$

But we know that the relations $\|\bar{P}_{n+1}(\kappa)\|_{n;-n}\leq\|P_{n+1}(\kappa)\|_{n;-n}+|\delta_{n+1}|$ and $|\delta_{n+1}| = |P_{n+1}(0;0)|\leq\|P_{n+1}(0)\|_{n;-n}$ hold. Moreover, (3.81) and (3.60) yield

$$\|P_{n+1}(\kappa)\|_{n;-n}\leq\|\pi_{n+1,\beta}(\kappa)\|_{n;-n}\leq 2\|\pi_{n\beta}(\kappa)\|_{n;-n}\leq 2A_nC_n|\epsilon|g\,\aleph^n,$$

assuming (3.84) and $\kappa\in D_n\supset D_{n+1}$ hold. Observe that, for positive x and p, $x^{-1}e^{-x^{-p}/(ep)}\leq 1$. Consequently, if we demand that

$$\aleph\leq\min\left(\tfrac{1}{8},b^2\right), \tag{3.87}$$

say, and

$$C_{n+1}\geq A_nC_n, \tag{3.88}$$

it remains to be proven that

$$\sup_{\substack{\kappa\in D_{n+1}\\ |\omega\cdot q|<\aleph^n(1-\aleph)b}}\frac{|\bar{p}_{n+1}(\kappa+\omega\cdot q)|}{w_{n+1}(q)^2}\leq C_{n+1}|\epsilon|g\,\aleph^{n+1}. \tag{3.89}$$

Notice that the rather arbitrary (3.87) imposes an interrelation between $\aleph$ and g, which is needed in the limit $b\leq g/3\to 0$; since we cannot take b large, we have to take $\aleph=o(b)$ in order to guarantee $e^{-2b\aleph^{-1}(1-\aleph)^2}\leq\aleph/4$ above.

In order to verify (3.89), we use the recursion formula

$$\bar{p}_{n+1}-\bar{p}_n \;=\; p_n\,\gamma_n\,a_n\,p_n+s_n(\,\cdot\,;0)-s_n(0;0) \tag{3.90}$$

subject to

$$a_n := (1-p_n\gamma_n)^{-1}\quad\text{and}\quad\gamma_n(\kappa) := \Gamma_n(\kappa;0),$$

which is an advocate of (3.76). The bound (3.52) yields

$$|\gamma_n(\kappa)|\leq C_\Gamma g^{-1}\aleph^{-n}|\aleph^{-n}\kappa|^5\qquad(\kappa\in D_n). \tag{3.91}$$

By virtue of $|s_n(\kappa;0)|\leq\|r_n(\kappa)\pi_n(\kappa)\|_{n;-n}$, (3.83) gives

$$|s_n(\kappa;0)|\leq 4C_n^2C_\Gamma|\epsilon|^2\begin{cases}g^3 & \text{if } n=0,\\ A_n\,g\,\aleph^ne^{-c\mu^n} & \text{if } n\geq 1,\end{cases} \tag{3.92}$$

in D_n.

(ii b.2) The Taylor expansion of $\bar{p}_{n+1}(\kappa) \equiv \bar{P}_{n+1}(\kappa;0)$. Let us abbreviate

$$\sigma_n(\kappa) \equiv \bar{p}_n(\kappa) - \bar{p}_n'(0)\kappa,$$

for each natural number n. The objective is to show that the estimates

$$|\bar{p}_{n+1}'(0)| \leq \left(1 - \frac{1}{n+3}\right) \frac{C_{n+1}|\epsilon|^{3/2} g}{2}, \tag{3.93}$$

i.e., the iterate of (3.78), and

$$\sup_{\kappa \in D_n} |(\sigma_{n+1} - \sigma_n)(\kappa)| \leq C_{n+1}|\epsilon|^{7/4} g\, \aleph^{n+1} \tag{3.94}$$

hold. Indeed, with the aid of such bounds together with (3.79), (3.89) follows from

$$\sup_{x \geq 0} (x + |\kappa|)^k e^{-\alpha x} = \left(\frac{k}{\alpha}\right)^k e^{\alpha|\kappa| - k} \qquad (\alpha > 0)$$

for $k = 1, 2$ and ϵ suitably small. Moreover, the Cauchy estimate

$$|\sigma_{n+1}(\kappa)| \leq |\sigma_n(\kappa)| + b^{-2}|\kappa|^2 \frac{\aleph^{-2n}}{1 - \aleph} \sup_{\zeta \in D_n} |(\sigma_{n+1} - \sigma_n)(\zeta)| \qquad (\kappa \in D_{n+1}),$$

implies that also (3.79) gets successfully iterated.

The bound in (3.91) implies

$$\gamma_n(0) = \gamma_n'(0) = 0,$$

such that $\bar{p}_{n+1}'(0) = \bar{p}_n'(0) + s_n'(0;0)$ according to (3.90), and hence

$$\bar{p}_{n+1}'(0) - \bar{p}_n'(0) = \frac{1}{2\pi i} \oint_{\partial D_n} \frac{s_n(\zeta;0)}{\zeta^2}\, d\zeta.$$

Thus, resorting to (3.92),

$$|\bar{p}_{n+1}'(0) - \bar{p}_n'(0)| \leq \aleph^{-n} b^{-1} \sup_{\kappa \in D_n} |s_n(\kappa,0)| \leq \frac{C_{n+1}|\epsilon|^{3/2} g}{2(n+2)(n+3)},$$

if the constant C_{n+1} satisfies

$$C_{n+1} \geq 8(n+2)(n+3) C_\Gamma b^{-1} A_n \tilde{B}_n |\epsilon|^{1/2} C_n^2. \tag{3.95}$$

The bound (3.93) now follows, assuming also $C_n \leq C_{n+1}$.

We still need to demonstrate (3.94). This will be provided by (3.90), since then

$$\sigma_{n+1}(\kappa) - \sigma_n(\kappa) = (p_n\, \gamma_n\, a_n\, p_n)(\kappa) + s_n(\kappa;0) - \sum_{l=0,1} s_n^{(l)}(0;0) \frac{\kappa^l}{l!},$$

such that (3.80), (3.91) and (3.92) yield (3.94) if

$$C_{n+1} \geq 4 C_\Gamma \aleph^{-1} \left(B_n^2 b^5 + 6 A_n \tilde{B}_n\right) |\epsilon|^{1/4} C_n^2. \tag{3.96}$$

Intuition behind (ii b.1–2). Due to the super-exponential decay of $s_n(\kappa;0)$ in (3.92) and the strong induction hypothesis $|\delta_n| \leq C_0|\epsilon| g\, \aleph^{2n}$ on the constant part of p_n, the flow of the remainder $\bar{p}_n = p_n - \delta_n$ reads roughly

$$\bar{p}_{n+1} \approx (1 - \bar{p}_n \gamma_n)^{-1} \bar{p}_n, \tag{3.97}$$

by (3.73). Hence, the *a priori* bound $|(1-\bar p_n\gamma_n)^{-1}|\leq 1+C|\epsilon|$ yields a sequence diverging in n, with very little hope of proving bounds such as (3.67)—see (3.89). However, the support of γ_k is highly concentrated on the annulus $\aleph^{k+1}b\leq|\kappa|\leq\aleph^k b$. Iterating for $n\geq k$ steps, with κ on the latter interval,

$$\bar p_{n+1}(\kappa)\approx(1-\bar p_1(\kappa)\gamma_k(\kappa))^{-1}\bar p_1(\kappa)=(1+\mathcal{O}(\epsilon))\,\bar p_1(\kappa).$$

That is, $\bar p_n$ remains close to $\bar p_1$, which enables proving (3.67) through (3.89).

In fact, our argument is different still: since $\chi_n(\aleph\,\kappa)=\chi_{n-1}(\kappa)$ and $G(\aleph\,\kappa;0)\approx\aleph^{-1}G(\kappa;0)$, we have $\gamma_n(\aleph\,\kappa)\approx\aleph^{-1}\gamma_{n-1}(\kappa)$ for $n\geq 2$. Inserting this into (3.97), we notice that the *approximate scaling invariance*

$$\bar p_{n+1}(\aleph\,\kappa)\approx\aleph\,\bar p_n(\kappa)$$

is consistent with the flow. This is what the bounds (3.78)–(3.79) reflect.

(ii c) The constant part δ_{n+1} of the diagonal. Recall that γ may be viewed as a function of δ_n by the induction hypotheses; the identity $\delta_n=\Delta_n(\gamma)$ is bijective on J_n. The flow produces a *near-identity analytic* function $\delta_{n+1}=\delta_n+d_n(\delta_n)$ of δ_n on the disk I_n, such that, for ϵ small enough,

$$\delta_{n+1}(I_n)\supset I_{n+1}. \tag{3.98}$$

The analyticity of the map $\delta_n\mapsto d_n$ can be read off (3.75) and the expression of r_n. As far as estimates are concerned,

$$|d_n|\leq\|r_n(0)\tilde\pi_n(0)\|_{n;-n}\leq CC_n^2C_\Gamma\,|\epsilon|^2g\begin{cases}g^2 & \text{if } n=0,\\ \aleph^n e^{-2c\mu^n} & \text{if } n\geq 1,\end{cases}$$

in the complex neighbourhood $2I_n$ of I_n of radius $\frac12|I_n|$, where $|I_n|$ is the diameter of the disk I_n. Consequently, a Cauchy estimate yields the bound

$$\sup_{\delta_n\in I_n}|\partial d_n/\partial\delta_n|\leq\frac{\sup_{\delta_n\in 2I_n}|d_n|}{\frac12|I_n|}\leq\tfrac12 \tag{3.99}$$

on the Lipschitz constant of d_n on I_n, provided $|\epsilon|\leq\epsilon_{n+1}$ with

$$\epsilon_{n+1}^{-1}\geq 2C_0^{-1}CC_n^2C_\Gamma\begin{cases}g & \text{if } n=0,\\ \aleph^{-n}e^{-2c\mu^n} & \text{if } n\geq 1.\end{cases} \tag{3.100}$$

In this case also

$$|d_n|\leq\tfrac12|I_n|-\tfrac12|I_{n+1}| \tag{3.101}$$

holds, which validates (3.98), considering how the boundary of I_n is transformed under δ_{n+1}.

Notice that (3.99) implies

$$\big|\delta_{n+1}(x)-\delta_{n+1}(y)\big|\geq\tfrac12|x-y|\qquad(x,y\in I_n), \tag{3.102}$$

meaning that $\delta_n\mapsto\delta_{n+1}$ is *one-to-one*. By continuity and (3.98), there exists a closed set $\tilde J_{n+1}\subset I_n$ that is bijectively and analytically mapped onto I_{n+1}: $\tilde J_{n+1}:=\delta_{n+1}^{-1}(I_{n+1})$. We can backtrack with the aid of the map Δ_n, obtaining a closed subset $J_{n+1}\subset I_\gamma$ (see (3.77)) that is bijectively and analytically mapped onto I_{n+1} by the map $\Delta_{n+1}:=\delta_{n+1}\circ\Delta_n$:

$$J_{n+1}:=\Delta_{n+1}^{-1}(I_{n+1}).$$

It follows immediately that

$$J_{n+1} \subset J_n.$$

(iii) Large values of n and the limit $g \to 0$. Suppose C_n is independent of g, which is the case for C_0. The recursive conditions (3.86), (3.88), (3.95), and (3.96) can be summarized in bounds of the form

$$C_{n+1} \geq K_n(g)\, C_n \quad \text{and} \quad C_{n+1} \geq L_n(g)|\epsilon|^{1/4}\, C_n^2.$$

Choosing $b := g/3$ (due to (3.95)) and $\aleph := \min\left(\frac{1}{8}, b^2\right)$ (due to (3.96); see also (3.87)), which is allowed, we may bound $K_n(g)$ and $L_n(g)$ *uniformly in* g: $\sup_{0<g<g_0} K_n(g) \leq K_n$ and $\sup_{0<g<g_0} L_n(g) \leq L_n$. This follows from the fact that $\aleph^{-1}e^{-c\mu} \to 0$ as $\aleph \to 0$. Moreover, $L_n \leq L$ for each n, such that we may choose

$$C_{n+1} := \max\left(K_n,\ L|\epsilon|^{1/4}C_n\right) C_n.$$

The numbers $K_n > 1$ converge to unity so fast that the number $K := \prod_{n=0}^{\infty} K_n > 1$ is finite. Now choose ϵ so small that $L|\epsilon|^{1/4}KC_0 \leq 1$. In particular, $C_1 = K_0C_0$, and inductively $C_n = K_0 \cdots K_{n-1}C_0 \leq KC_0$. We conclude that the sequences (C_n) and (ϵ_n) (see (3.84) and (3.100)) converge to positive numbers.

(iv) Fine-tuning the Lyapunov exponent γ. The maps δ_n are relatively expansive; (3.102) holds, while the target I_n contracts by a factor of $\aleph^2 < \frac{1}{2}$ at each step. Thus, demanding $\Delta_n(J_n) = I_n$ at each step for the map $\Delta_n = \delta_n \circ \cdots \circ \delta_0$ amounts to

$$|x - y| \leq 2^n \big|\Delta_n(x) - \Delta_n(y)\big| \leq Cg\,(2\aleph^2)^n \qquad (x, y \in J_n),$$

or $\lim_{n\to\infty}|J_n| = 0$. Because the J_n's form an ever decreasing chain of closed disks, their intersection consists of precisely one point:

$$\{\gamma\} := \bigcap_{n=0}^{\infty} J_n \subset I_\gamma.$$

The value of γ is an analytic function of ϵ, because the sequence $\Delta_n^{-1}(0)$ converges uniformly to γ with respect to ϵ. For real values of ϵ, Δ_n sends real numbers to real numbers, making γ real. □

4. Proof of Theorem 1

Let us summarize what we have learned thus far. The solution $X^0(z)$ to the equations of motion in the uncoupled case was found. In the coupled case we resolved KAM-type small denominator issues, which contributed the $t \to -\infty$ ($z = 0$) asymptotic $X_0(\theta)$ of the general solution $X(z, \theta)$, as well as the linearization $X_1(\theta) \equiv \partial_z X(0, \theta)$.

We can now solve (1.11), and thus find the unstable manifold $\mathcal{W}_\lambda^u$ also "far away" from the torus $\mathcal{T}_\lambda$, by a Contraction Mapping argument.

To begin with, we single out the uncoupled part X^0 of the complete solution X;

$$X = X^0 + \widetilde{X} \quad \text{with} \quad \widetilde{X}|_{\epsilon=0} \equiv 0.$$

As $\mathcal{L}^2X^0 = (\gamma^2 \sin\Phi^0, 0)$, (1.11) now becomes $\mathcal{L}^2\widetilde{X} = -(\gamma^2 \sin\Phi^0, 0) + \Omega(X^0 + \widetilde{X})$. In other words, the map $\widetilde{X}$ has to satisfy

$$\mathcal{K}\widetilde{X} = \widetilde{W}(\widetilde{X}), \tag{4.1}$$

where we define the linear operator

$$\mathcal{K} := \begin{pmatrix} L & 0 \\ 0 & \mathcal{L}^2 \end{pmatrix} \quad \text{with} \quad L := \mathcal{L}^2 - \gamma^2 \cos \Phi^0 \tag{4.2}$$

and the nonlinear operator $\widetilde{W}$ through the expression

$$\widetilde{W}(\widetilde{X}) := (-\gamma^2 \sin \Phi^0 - \gamma^2 (\cos \Phi^0)\widetilde{\Phi}, 0) + \Omega(X^0 + \widetilde{X}). \tag{4.3}$$

Throughout the rest of the work, we shall refer to different parts of the Taylor expansion of a suitable function $h(z,\theta)$ around $z = 0$ using the notation

$$h_k(\theta) := \frac{\partial_z^k h(0,\theta)}{k!}, \quad h_{\leq k}(z,\theta) := \sum_{j=0}^{k} z^k h_k(\theta) \quad \text{and} \quad \delta_k h := h - h_{\leq k-1}.$$

Observe that $X_0 = \widetilde{X}_0$ and $X_1 = (4,0) + \widetilde{X}_1$ exist. Setting

$$\widetilde{X}(z,\theta) \equiv X_{\leq 1}(z,\theta) - (4,0)z + Z(z,\theta), \tag{4.4}$$

we may transform equation (4.1) into the equation

$$\mathcal{K}Z = W(Z) \tag{4.5}$$

for $Z = \delta_2 \widetilde{X}$, where we define W through

$$W(Z) := \delta_2 \left[\widetilde{W}(\widetilde{X}) + \begin{pmatrix} \gamma^2 (\cos \Phi^0) \widetilde{\Phi}_{\leq 1} \\ 0 \end{pmatrix} \right], \tag{4.6}$$

taking now (4.4) as the *definition of* $\widetilde{X}$.

Let us consider the complex Banach space $\mathcal{A}$ of (bounded) analytic functions Z on the compact set

$$\Pi_\tau := \left\{ (z,\theta) \;\middle|\; \mathfrak{Re}\,(z,\theta) \in [-1-\tau, 1+\tau] \times \mathbb{T}^d,\ \mathfrak{Im}\,(z,\theta) \in [-\tau,\tau]^{d+1} \right\},$$

$\tau \geq 0$, equipped with the supremum norm, and its closed subspace

$$\mathcal{A}_1 := \{ Z \in \mathcal{A} \mid Z_{\leq 1} = 0 \}. \tag{4.7}$$

For future use, let us also define the closed origin-centered balls

$$B(R) := \{ Z \in \mathcal{A} \mid \|Z\|_\infty \leq R \} \quad \text{and} \quad B_1(R) := B(R) \cap \mathcal{A}_1.$$

Any element of $\mathcal{A}$ extends analytically to $\Pi_{\tau'}$ for some $\tau' > \tau$, allowing uniform estimates on its derivatives *on* Π_τ.

Remark 4.1. Whereas equation (4.1) is plagued by small denominators, equation (4.5) is not. This is so due to the decomposition (4.4) which separates the previously solved "KAM-asymptotics" $X_{\leq 1}$ from $\widetilde{X}$ and enables reducing (4.1) to (4.5) on the space $\mathcal{A}_1$, which one could well call the small-denominator-free subspace of $\mathcal{A}$.

4.1. **Existence and uniqueness of Z.** Postponing the proofs until the end of this section, we make two observations, important in demonstrating that (4.5) is solvable.

Lemma 4.2. *With sufficiently small R, τ, and ϵ (depending on the analyticity region of f), the operator $W : \mathcal{A} \to \mathcal{A}_1$ maps the ball $B(R)$ in $\mathcal{A}$ into a ball $B_1(R')$ in $\mathcal{A}_1$ with $R' = Cg^2(R^2 + |\epsilon|)$, and $W|_{\mathcal{A}_1}$ is Lipschitz continuous on $B_1(R)$ with a Lipschitz constant proportional to $g^2(R+|\epsilon|)$. If the restriction of $Z \in \mathcal{A}$ to a* real *neighbourhood of $[-1,1] \times \mathbb{T}^d$ has the* real *range $\mathbb{R} \times \mathbb{R}^d$ and ϵ is real, then the same is true of $W(Z)$.*

Lemma 4.3. *If $0 < \tau < 1$, the linear operator $\mathcal{K} : \mathcal{A}_1 \to \mathcal{A}_1$ has a bounded inverse $\mathcal{K}^{-1} \in \mathcal{L}(\mathcal{A}_1)$ obeying $\|\mathcal{K}^{-1}\|_{\mathcal{L}(\mathcal{A}_1)} \leq C\gamma^{-2}\tau^{-1}(1-\tau^2)^{-2}$. It preserves analyticity in ϵ. If the restriction of $Z \in \mathcal{A}$ to a* real *neighbourhood of $[-1,1] \times \mathbb{T}^d$ has the* real *range $\mathbb{R} \times \mathbb{R}^d$, the same is true of $\mathcal{K}^{-1}Z$.*

We have developed enough machinery to extract a solution from (4.5):

Theorem 4.4. *For sufficiently small R, $\epsilon_0 < R/2$, and τ (depending on the analyticity regions of f and $X_{\leq 1}$), equation (4.5) has a unique solution $Z \in B_1(R)$. It is continuous on D, analytic in ϵ, and bounded uniformly by $C|\epsilon|$. The restriction $Z|_{[-1,1]\times\mathbb{T}^d}$ takes values in $\mathbb{R} \times \mathbb{R}^d$, provided ϵ is real.*

Proof. We know by Lemmata 4.2 and 4.3 that $\mathcal{K}^{-1}W$ maps $B_1(R)$ into itself. We may furthermore choose ϵ_0 and R such that the operator $\mathcal{K}^{-1}W$ becomes contractive on $B_1(R)$. The Banach Fixed Point Theorem implies that $\mathcal{K}^{-1}W$ has a unique fixed point Z in the ball $B_1(R)$.

The theorem also implies that Z is analytic in ϵ. Namely, Lemma 4.3 says that $\mathcal{K}^{-1}$ preserves such a property. Furthermore, the ϵ-dependence of W comes solely from γ, X_0, X_1, and Ω, making it analytic. Hence, the uniformly convergent sequence $((\mathcal{K}^{-1}W)^k(0))_{k\in\mathbb{N}}$ reveals the analyticity of the limit Z. The latter is also $\mathbb{R}\times\mathbb{R}^d$-valued on $[-1,1] \times \mathbb{T}^d$ if ϵ is real. Finally,

$$\|Z\|_\infty \leq \|(\mathcal{K}^{-1}W)(Z) - (\mathcal{K}^{-1}W)(0)\|_\infty + \|(\mathcal{K}^{-1}W)(0)\|_\infty \leq L\|Z\|_\infty + C|\epsilon|$$

yields $\|Z\|_\infty \leq C|\epsilon|/(1-L)$. Here $(\mathcal{K}^{-1}W)(0)$ was bounded using R' of Lemma 4.2 at $R = 0$. □

4.2. **Putting it all together.** To reach the statement of Theorem 1 about X^u, we glue together the pieces provided by Theorems 2.1, 3.3, and 4.4.

Assuming $\langle \Psi_0 \rangle = 0$, we have constructed analytic maps γ and

$$X(z,\theta) = X_0(\theta) + zX_1(\theta) + \delta_2 X(z,\theta) \quad \text{with} \quad \delta_2 X = Z + \delta_2 X^0$$

that solve (1.11) in a complex neighbourhood of $[-1,1] \times \mathbb{T}^d$ and satisfy the *physical constraint* $\Phi_1|_{\epsilon=0} = 4$. Recall now (1.16). Since (1.18) is not automatically satisfied, we are required to pinpoint specific values of α and β so as to fulfill $X_{\alpha,\beta}(1,0) = (\pi,0)$. To this end, we utilize the Implicit Function Theorem.

Consider the implicit equation $\mathfrak{X}(\epsilon, g; \alpha, \beta) := X(\alpha,\beta) + (0,\beta) - (\pi, 0) = 0$. Both $\mathfrak{X}$ and $\frac{\partial \mathfrak{X}}{\partial(\alpha,\beta)}$ are continuous, and we get from $X = (\Phi^0, 0) + \mathcal{O}(\epsilon)$ that

$$\mathfrak{X}(0, g; 1, 0) = 0 \quad \text{and} \quad \det\left(\frac{\partial \mathfrak{X}(\epsilon, g; \alpha, \beta)}{\partial(\alpha,\beta)}\right) = \frac{4}{1+\alpha^2} + \mathcal{O}(\epsilon)$$

for $(\epsilon, g) \in D$ and for whichever values of α and β the map $\mathfrak{X}$ is well-defined. Hence, if we choose ϵ_0 small enough, there exist unique continuous functions α and β on D, analytic with respect to ϵ, such that $\alpha(0,g) = 1$, $\beta(0,g) = 0$, and

$$\mathfrak{X}(\epsilon, g; \alpha(\epsilon, g), \beta(\epsilon, g)) = 0.$$

Moreover, $\alpha(\epsilon, g) \in \mathbb{R}$ and $\beta(\epsilon, g) \in \mathbb{R}^d$ for ϵ real, as $\mathfrak{X}$ is then real-valued. A good reference here is [Chi96].

4.3. **Proofs of Lemmata 4.2 and 4.3.** We conclude the section by presenting the proofs of Lemmata 4.2 and 4.3 used in the proof of Theorem 4.4.

Proof of Lemma 4.2. Given $Z \in \mathcal{A}$ with $\|Z\|_\infty \le R$, we study $W(Z)$—defined in (4.6), and clearly an element of $\mathcal{A}_1$. Notice that in the relation (4.4), expressing $\widetilde{X}$ in terms of Z, the maps X_0 and X_1 were previously determined and are independent of Z. Furthermore, taking advantage of (4.4) and Theorems 2.1 and 3.3, we deduce

$$\|\widetilde{X}\|_\infty \le C(|\epsilon| + R). \tag{4.8}$$

With the aid of (1.10), cast equation (4.3) as

$$\widetilde{W}(\widetilde{X}) := (g^2 \sin(\Phi^0 + \widetilde{\Phi}) - \gamma^2 \sin \Phi^0 - \gamma^2 \cos(\Phi^0)\widetilde{\Phi}, 0) + \lambda\, \widetilde{\Omega}(X^0 + \widetilde{X}).$$

Recall that f is analytic on the strip $|\Im\mathfrak{m}\,\phi|, |\Im\mathfrak{m}\,\psi| \le \eta$. Also, $\Im\mathfrak{m}\,\Phi^0(z) = \mathcal{O}(\tau)$ on Π_τ, when $\tau \ll 1$. Hence, owing to (4.8), our function $\widetilde{\Omega}(X^0 + \widetilde{X})$ is well-defined for λ and R sufficiently small and the strip Π_τ about $[-1,1] \times \mathbb{T}^d$ narrow enough.

Since $\sin(\Phi^0 + \widetilde{\Phi}) = \sin \Phi^0 + \cos(\Phi^0)\widetilde{\Phi} + \mathcal{O}(\widetilde{\Phi}^2)$, in a neighbourhood of Π_τ

$$\|\widetilde{W}(\widetilde{X})\|_\infty \le |g^2 - \gamma^2|\, \|\sin \Phi^0 + \cos(\Phi^0)\widetilde{\Phi}\|_\infty + Cg^2\|\widetilde{\Phi}\|_\infty^2 + |\lambda|\, \|\widetilde{\Omega}(X^0 + \widetilde{X})\|_\infty.$$

The factor $g^2 - \gamma^2$ is the reason we chose to subtract $\gamma^2 \cos(\Phi^0)\widetilde{\Phi}$ from both sides in equation (4.1). Namely, $|g^2 - \gamma^2| = |2g + \gamma - g||g - \gamma| \le Cg^2|\epsilon|$. Terms proportional to $\widetilde{\Phi}$ are dominated by (4.8). Thus, for ϵ and R small (independently of g and each other),

$$\|W(Z)\|_\infty \le Cg^2(R^2 + |\epsilon|).$$

In order to obtain the Lipschitz continuity of $W|_{\mathcal{A}_1}$, it suffices to show that $Z \overset{(4.4)}{\mapsto} \widetilde{X} \mapsto \widetilde{W}(\widetilde{X})$ is Lipschitz, as neither $(\widetilde{W}(\widetilde{X}))_{\le 1}$ nor $\widetilde{X}_{\le 1}$ depend on $Z =: \delta_2 \widetilde{X}$. To that end, we use the Mean Value Theorem, see [Cha85], and conclude that for some $Z =: \delta_2\widetilde{X}$ on the line segment between two points $Z' =: \delta_2\widetilde{X}'$ and $Z'' =: \delta_2\widetilde{X}''$

$$\|\widetilde{W}(\widetilde{X}') - \widetilde{W}(\widetilde{X}'')\|_\infty \le \|D\widetilde{W}(\widetilde{X})\|\, \|Z' - Z''\|_\infty.$$

The derivative is bounded by $Cg^2(R + |\epsilon|)$ given (4.8), in particular when $\|Z\|_\infty \le R$.

From its explicit expression, one immediately recognizes that W preserves the class of functions whose restriction to $[-1,1] \times \mathbb{T}^d$ has the *real* range $\mathbb{R} \times \mathbb{R}^d$, if ϵ is real. □

Proof of Lemma 4.3. $\mathcal{L}$ maps $\mathcal{A}_1$ into itself, and $\mathcal{K}$ in (4.2) inherits this feature.

Let us start with the "pendulum part" of $\mathcal{K}$, and solve

$$Lf = g$$

resorting to the method of characteristics; we write $(z,\theta) = (\zeta e^{\gamma t}, \vartheta + \omega t)$ in order to obtain an ordinary differential equation (ODE). Recalling the identity (1.7), we see that

$$\left(\partial_t^2 - \gamma^2 \cos\Phi^0(\zeta e^{\gamma t})\right) f(\zeta e^{\gamma t}, \vartheta + \omega t) = g(\zeta e^{\gamma t}, \vartheta + \omega t), \tag{4.9}$$

and our task reduces to studying $L_t := \partial_t^2 - \gamma^2 \cos\Phi^0(\zeta e^{\gamma t})$. Since a translation in t and ϑ eliminates ζ, we can just as well set $\zeta = 1$.

We proceed in the Fourier language. The function f solves equation (4.9) if and only if for all $q \in \mathbb{Z}^d$ the functions $u(t) := e^{iq\cdot\omega t}\hat{f}(e^{\gamma t}, q)$ and $v(t) := e^{iq\cdot\omega t}\hat{g}(e^{\gamma t}, q)$ satisfy

$$L_t u = v.$$

Noticing that $\cos\Phi^0(e^{\gamma t}) = 2\tanh^2\gamma t - 1$, we see that L_t has got the zero mode

$$u_1(t) := (\cosh\gamma t)^{-1},$$

i.e., $L_t u_1 = 0$. Since $L_t u = 0$ is a linear second order ODE, there exists precisely one other zero mode u_2 of L_t that is linearly independent of u_1. Because $u_1(t) \neq 0$ for any $t \in \mathbb{R}$, u_2 may be found by a standard procedure:

$$u_2(t) := u_1(t) \int \frac{dt}{u_1^2(t)} = \frac{t}{2\cosh\gamma t} + \frac{\sinh\gamma t}{2\gamma},$$

omitting any additive constant emerging from the integral. Let us express the linear homogeneous equation $L_t u = 0$ as the first order system $\dot{U} = AU$ with $U := (u, \dot{u})^T$ and $A(t) := \left(\begin{smallmatrix} 0 & 1 \\ \gamma^2\cos\Phi^0(e^{\gamma t}) & 0 \end{smallmatrix}\right)$. Then $w := \left(\begin{smallmatrix} u_1 & u_2 \\ \dot{u}_1 & \dot{u}_2 \end{smallmatrix}\right)$ is a fundamental matrix solution of the system (*i.e.*, $\dot{w} = Aw$) with $\det w = 1$ and thus

$$w^{-1} = \begin{pmatrix} \dot{u}_2 & -u_2 \\ -\dot{u}_1 & u_1 \end{pmatrix} \quad \text{and} \quad w(t)w^{-1}(s) = \begin{pmatrix} * & u_2(t)u_1(s) - u_1(t)u_2(s) \\ * & * \end{pmatrix}.$$

In terms of a first order system, the complete equation $L_t u = v$ reads $\dot{U} = AU + V$, $V := (0, v)^T$. Varying constants,

$$U(t) = w(t)\left(w^{-1}(t_0)U(t_0) + \int_{t_0}^t w^{-1}(s)V(s)\,ds\right).$$

Next, we take $t_0 \to -\infty$. In that limit $u(t_0) = \mathcal{O}(e^{2\gamma t_0})$, such that

$$u(t) = \int_{-\infty}^t \left[u_2(t)u_1(s) - u_1(t)u_2(s)\right] v(s)\,ds.$$

Equivalently,

$$\hat{f}(e^{\gamma t}, q) = \int_{-\infty}^0 \widetilde{K}_\Phi(s; e^{\gamma t})\, \hat{g}(e^{\gamma t}e^{\gamma s}, q)\, e^{iq\cdot\omega s}\,ds \tag{4.10}$$

in terms of the kernel

$$\widetilde{K}_\Phi(s; z) := \mathcal{W}_{\Phi 2}(z)\mathcal{W}_{\Phi 1}(ze^{\gamma s}) - \mathcal{W}_{\Phi 1}(z)\mathcal{W}_{\Phi 2}(ze^{\gamma s}),$$

defined (by analytic continuation) on $\{(s,z) \in \mathbb{R}\times\mathbb{C} \,|\, z \notin \{\pm i, \pm ie^{-\gamma s}\}\}$, where

$$\mathcal{W}_{\Phi 1} := 2P \quad \text{and} \quad \mathcal{W}_{\Phi 2} := \gamma^{-1}P\,\ln + \frac{1}{4}\gamma^{-1}Q,$$

and

$$P(z) := (z^2+1)^{-1}z \quad \text{and} \quad Q(z) := z^{-1}(z^2-1). \tag{4.11}$$

This is so, because $\mathcal{W}_{\Phi j}(e^{\gamma t}) \equiv u_j(t)$.

In a complex strip $|\mathfrak{Im}\, z| \le \tau < 1$, the inequality $|z^2+1| \ge 1-\tau^2$ yields

$$|\widetilde{K}_\Phi(s;z)| \le C(1-\tau^2)^{-2}\gamma^{-1}e^{\gamma|s|}, \quad s \le 0. \tag{4.12}$$

Since $\hat{f}(0,q) = \hat{g}(0,q) = 0$, we find that (4.10) remains true if 0 replaces $e^{\gamma t}$. Inserting all this into the Fourier series of $f(z,\theta)$ leads to

$$f(z,\theta) = \int_{-\infty}^0 \widetilde{K}_\Phi(s;z)\, g(ze^{\gamma s}, \theta+\omega s)\, ds, \quad (z,\theta) \in [-1,1]\times\mathbb{T}^d. \tag{4.13}$$

Here Fubini's Theorem was used, taking advantage of the bound (4.12). Indeed, we may express $g(z,\theta) = z^2 h(z,\theta)$, where h is analytic in the same region as g. Since $|\hat{h}(z,q)| \le \sup_{\Pi_{\tau'}}|h(z,\theta)|e^{-\tau'|q|} \le Ce^{-\tau'|q|}$ for some $\tau' > \tau$, we have on Π_τ that

$$\sum_{q\in\mathbb{Z}^d}\int_{-\infty}^0 \left|\widetilde{K}_\Phi(s;z)\,\hat{g}(ze^{\gamma s},q)\,e^{iq\cdot(\theta+\omega s)}\right| ds \le C(1-\tau^2)^{-2}\gamma^{-2}|z|^2\sum_{q\in\mathbb{Z}^d} e^{-(\tau'-\tau)|q|} < \infty.$$

Following the line of reasoning above, solving the "rotator part"

$$\mathcal{L}^2 f = g$$

amounts to integrating $\ddot{u} = v$ and results in an expression like (4.13) with the kernel

$$\widetilde{K}_\Psi(s;z) := \mathcal{W}_{\Psi 2}(z)\mathcal{W}_{\Psi 1}(ze^{\gamma s}) - \mathcal{W}_{\Psi 1}(z)\mathcal{W}_{\Psi 2}(ze^{\gamma s}) \equiv -s,$$

introducing

$$\mathcal{W}_{\Psi 1} := 1 \quad \text{and} \quad \mathcal{W}_{\Psi 2} := \gamma^{-1}\ln.$$

For each index $n \in \mathbb{N}\cup\{\infty\}$ define now

$$I_n(z,\theta) := \int_{-n}^0 \widetilde{K}(s;z)\, Z(ze^{\gamma s},\theta+\omega s)\, ds \quad \text{with} \quad \widetilde{K} := \begin{pmatrix} \widetilde{K}_\Phi & 0 \\ 0 & \widetilde{K}_\Psi \end{pmatrix},$$

where $(z,\theta) \in \Pi_\tau$ and $Z \in \mathcal{A}_1$ are arbitrary. Also denote

$$\widetilde{\mathcal{K}}Z := I_\infty.$$

Since the integrand here is an analytic function of (z,θ) on the compact region Π_τ and continuous in $s \in [-n,0]$, it follows from an exercise in function theory that I_n *with* $n < \infty$ is analytic on Π_τ; see p. 123 of [Ahl66]. As an element of $\mathcal{A}_1$, $Z(z,\theta)$ has the representation $z^2\widetilde{Z}(z,\theta)$, where $\widetilde{Z}$ is analytic on Π_τ. Accordingly, (4.12) implies

$$\left|\widetilde{\mathcal{K}}Z(z,\theta) - I_n(z,\theta)\right| \le C\gamma^{-1}\int_{-\infty}^{-n} e^{-\gamma s}\,|Z(ze^{\gamma s},\theta+\omega s)|\, ds \le C|z|^2\frac{\|\widetilde{Z}\|_\infty}{\gamma^2\, e^{\gamma n}}, \tag{4.14}$$

showing that $I_n \to \widetilde{\mathcal{K}}Z$ uniformly on Π_τ as $n\to\infty$. Hence, also $\widetilde{\mathcal{K}}Z$ is analytic on the latter region. Moreover, $I_n(z,\theta) = \mathcal{O}(z^2)$ as $z \to 0$, which by virtue of (4.14) yields $\widetilde{\mathcal{K}} : \mathcal{A}_1 \to \mathcal{A}_1$.

We showed above that if $Z \in \mathcal{A}_1$ and $\mathcal{K}Z = Z'$ (thus $Z' \in \mathcal{A}_1$), then $Z = \widetilde{\mathcal{K}}Z'$ holds on $[-1,1]\times\mathbb{T}^d \subset \Pi_\tau$. But each side of the latter equation is analytic on Π_τ and hence agree there, meaning that $\widetilde{\mathcal{K}}$ is the left inverse of $\mathcal{K}$: $\widetilde{\mathcal{K}}\mathcal{K} = \mathbb{1}_{\mathcal{A}_1}$. A direct computation shows that it is also the right inverse. In other words,

$$\widetilde{\mathcal{K}} = \mathcal{K}^{-1} \text{ on } \mathcal{A}_1.$$

$K(s;z) \in \mathbb{R}$, provided $z \in \mathbb{R}$. Thus, should the restriction $Z|_{[-1,1]\times\mathbb{T}^d}$ be real-valued, so is $(\mathcal{K}^{-1}Z)|_{[-1,1]\times\mathbb{T}^d}$.

The integrals I_n also depend analytically on γ. Thus, according to Theorem 3.3, they are analytic functions on the domain $|\epsilon| < \epsilon_0$. Since $|\gamma - g| < Cg|\epsilon|$, the trivia $\gamma > \frac{1}{2}g > 0$ and (4.14) guarantee that the convergence $I_n \to \mathcal{K}^{-1}Z$ takes place uniformly on compact subsets of D defined in (1.15) (g bounded away from zero).

It remains to be checked that $\mathcal{K}^{-1}$ is bounded. For $Z \in \mathcal{A}_1$, $Z(z,\theta) = \sum_{k=2}^{\infty} \frac{1}{k!} Z_k(\theta)\, z^k$ converges in the disk $\bar{\mathbb{D}}(0,\tau) := \{z \in \mathbb{C} \mid |z| \le \tau\}$. Using the Cauchy inequalities $|Z_k(\theta)| \le k!\, \tau^{-k}\, \|Z\|_\infty$ we deduce the bound

$$|Z(z,\theta)| \le 2(|z|/\tau)^2 \|Z\|_\infty \quad \text{if} \quad z \in \bar{\mathbb{D}}(0,\tau/2)$$

In Π_τ, $|z| \le R$ for a certain $R = 1 + \mathcal{O}(\tau)$, such that $ze^{\gamma s} \in \bar{\mathbb{D}}(0,\tau/2)$ whenever $s \le S := -\gamma^{-1}\ln(2R/\tau))$. The bound (4.12) for $\widetilde{K}_\Phi$ applies to $\widetilde{K}_\Psi$ as well. Summarizing,

$$\|\mathcal{K}^{-1}Z\|_\infty \le \frac{C\|Z\|_\infty}{\gamma(1-\tau^2)^2}\left(\int_S^0 e^{-\gamma s}\,ds + \int_{-\infty}^S \frac{e^{\gamma s}R^2}{\tau^2(1+\tau)^2}\,ds\right) \le \frac{C\|Z\|_\infty}{\gamma^2\tau(1-\tau^2)^2},$$

which finishes the proof. ☐

5. Analytic Continuation of the Solution

Here we present the proof of Theorem 2. In the notation of Section 4, the *existing* map $Z = \delta_2\widetilde{X}$ solves (4.5) and, by virtue of $\widetilde{W}$'s analyticity, admits the representation

$$\begin{aligned}\delta_2\widetilde{X} = \mathcal{K}^{-1}\delta_2\Bigg[\begin{pmatrix}\gamma^2\cos\Phi^0 & 0\\ 0 & 0\end{pmatrix}\widetilde{X}_{\le 1} + \sum_{k=0}^{\infty} w^{(k)}\big(\widetilde{X}_{\le 1}\big)^{\otimes k}\Bigg] + \\ + \mathcal{K}^{-1}\sum_{k=1}^{\infty}\Big[w^{(k)}\big(\widetilde{X}_{\le 1} + \delta_2\widetilde{X}\big)^{\otimes k} - w^{(k)}\big(\widetilde{X}_{\le 1}\big)^{\otimes k}\Big]\end{aligned} \tag{5.1}$$

on the set Π_τ, taking ϵ small enough, and denoting

$$w^{(k)} := \frac{1}{k!}D^k\widetilde{\widetilde{W}}(0) \tag{5.2}$$

and a repeated argument of such a symmetric k-linear operator by

$$(x)^{\otimes k} := \underbrace{(x,\ldots,x)}_{k \text{ times}},$$

for the sake of brevity. Observe that we have omitted a δ_2 in front of the square brackets on the second line of (5.1) as redundant.

Equation (5.1) may be viewed as a recursion relation for $\delta_2\widetilde{X}$. It is crucial that

$$w^{(0)},\, w^{(1)} = \mathcal{O}(\epsilon g^2), \tag{5.3}$$

when $(\epsilon, g) \in D$; see (1.15). Namely, any given order $\delta_2 \widetilde{X}^\ell$ in the convergent expansion

$$\delta_2 \widetilde{X} = \sum_{\ell=1}^{\infty} \epsilon^\ell \, \delta_2 \widetilde{X}^\ell$$

is then completely determined by $\widetilde{X}_{\leq 1}$ and the *lower orders* $\delta_2 \widetilde{X}^l$ $(1 \leq l \leq \ell - 1)$ through the right-hand side of (5.1). Moreover, since $\widetilde{X}_{\leq 1} = \mathcal{O}(\epsilon)$, only *finitely many* terms in the sum over the index k are involved. Together these facts imply that only *finitely many* recursive steps using (5.1) are needed to completely describe any given order $\delta_2 \widetilde{X}^\ell$ in terms of $\widetilde{X}_{\leq 1}$ alone and that, at each such step, only *finitely many* terms from the k-sum contribute.

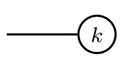
It is important to understand that $\widetilde{X}_{\leq 1}$ is a predetermined function. As we shall see, the recursion procedure will then provide the analytic continuation of each $X^{u,\ell} = \widetilde{X}^\ell_{\leq 1} + \delta_2 \widetilde{X}^\ell$ $(\ell \geq 1)$ to the large region $\mathbb{U}_{\tau,\vartheta} \times \{|\Im\mathfrak{m}\, \theta| \leq \sigma\}$ of Theorem 2.

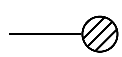
5.1. **Tree expansion.** We next give a pictorial representation of the above recursion. It involves tree diagrams similar to those of Gallavotti, *et al.* (see, *e.g.*, [Gal94b,CG94]), with one difference: there will be no resummations nor cancellations, as the expansion in (5.1) contains no resonances and is instead well converging. This so-called tree expansion is needed for bookkeeping and pedagogical purposes; we simply choose to draw a tree instead of spelling out a formula.

Let us first define the auxiliary functions

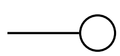
$$h^{(k)} := \begin{cases} w^{(0)} + \left[\begin{pmatrix} \gamma^2 \cos \Phi^0 & 0 \\ 0 & 0 \end{pmatrix} + w^{(1)} \right] \widetilde{X}_{\leq 1} & \text{if } k = 1, \\ w^{(k)} \big(\widetilde{X}_{\leq 1} \big)^{\otimes k} & \text{if } k = 2, 3, \dots, \end{cases}$$

and make the identifications

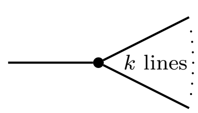


$:= \mathcal{K}^{-1} \delta_2 h^{(k)}$ and  $:= \mathcal{K}^{-1} \delta_2 \sum_{k=0}^{\infty} h^{(k)}.$ (5.4)

Furthermore, let

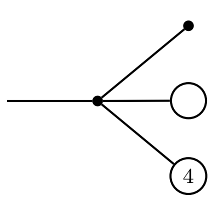
 $:= \delta_2 \widetilde{X},$ $:= \widetilde{X}_{\leq 1},$

and, for $k \geq 1$,



$:= \mathcal{K}^{-1} w^{(k)}.$

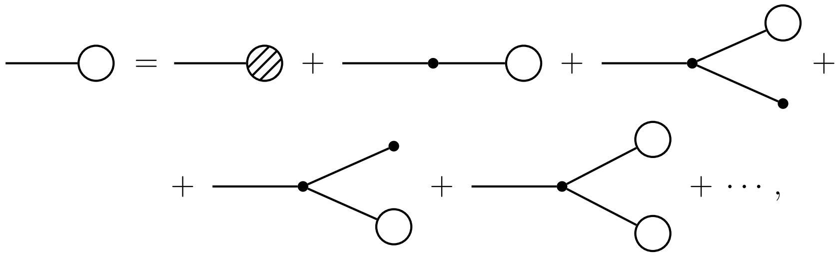
In the diagram representing the k-linear $w^{(k)}$, the k "free" lines to the right of the node stand for the arguments. We say that these lines *enter* the *internal* node, whereas the single line to the left of the node *leaves* it. For instance,



$= \mathcal{K}^{-1} w^{(3)} \big(\widetilde{X}_{\leq 1}, \, \delta_2 \widetilde{X}, \, \mathcal{K}^{-1} \delta_2 h^{(4)} \big).$

Notice that, as $w^{(k)}$ is symmetric, permuting the lines entering a node does not change the resulting function. We emphasize that all of the functions introduced above are analytic on Π_τ and $|\epsilon| < \epsilon_0$.

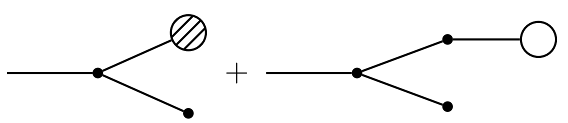
In terms of such *tree diagrams*, or simply *trees*, equation (5.1) reads



(5.5)

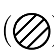
using multilinearity to split the sums $\widetilde{X}_{\leq 1} + \delta_2 \widetilde{X}$ into pieces. Above, the sum after the first tree consists of *all* trees having one internal node and an arbitrary number of *end nodes, at least one of which, however, is a white circle.* This rule encodes the fact that on the second line of (5.1) the summation starts from $k = 1$ and that the contributions with only $\widetilde{X}_{\leq 1}$ in the argument (*i.e.*, trees with only black dots as end nodes) are cancelled.

Using (5.1) recursively now amounts to replacing each of the lines with a white-circled end node by the complete expansion of such a tree above. This is to be understood additively, so that replacing one end node, together with the line leaving it, by a sum of two trees results in a sum of two new trees. For example, such a replacement in the third tree on the right-hand side of (5.5) by the first two trees gives the sum



Before proceeding, we introduce a little bit of terminology. The leftmost line in a tree is called the *root line*, whereas the node it leaves (*i.e.*, the uniquely defined leftmost node) is called the *root*. A line leaving a node v and entering a node v' can always be interpreted as the root line of a *subtree*, the maximal tree consisting of lines and nodes in the original tree with v as its root. We call v a (not necessarily unique) *successor* of v', whereas v' is the unique *predecessor* of v.

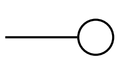
The recursion (5.5) can be repeated on a given tree if it has at least one white circle left. Otherwise, the tree in question must satisfy

(R1′) The tree has only filled circles () and black dots ($\bullet$) as its end nodes,

together with

(R2′) Any internal node has an entering (line that is the root line of a) subtree containing at least one filled circle as an end node.

After all, the recursion can only stop by replacing an existing white circle with a filled one. Continuing *ad infinitum* yields the expansion

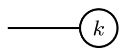
 $$= \sum \big(\text{Trees satisfying (R1′) and (R2′)}\big) = \sum_{\text{trees } T}{}' \, T, \qquad (5.6)$$

where the prime restricts the summation to trees T satisfying (R1′) and (R2′). We point out that each admissible tree appears precisely once in this sum, considering different two trees that can be superposed by a (nontrivial) permutation of subtrees that enter the same node.

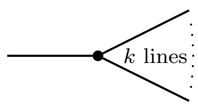
The earlier discussion concerning the description of $\delta_2 \widetilde{X}^\ell$ in terms of a finite sum involving only $\widetilde{X}_{\leq 1}$ translates to the language of trees in a straightforward fashion. First, the second part of (5.3) and $\widetilde{X}_{\leq 1} = \mathcal{O}(\epsilon)$ amount pictorially to

 $= \mathcal{O}(\epsilon)$ and $= \mathcal{O}(\epsilon)$.

Second, $w^{(k)} = \mathcal{O}(g^2)$ and the first part of (5.3) yield

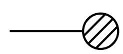
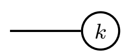
 $= \mathcal{O}(\epsilon^k) \qquad (k = 1, 2, \dots)$

and

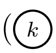


$= \mathcal{O}(1) \qquad (k = 2, 3, \dots)$.

Expanding the filled end nodes

 $= \sum_{k=1}^{\infty}$  , (5.7)

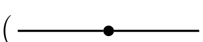
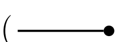
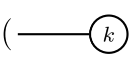
according to (5.4), on the right-hand side of (5.6), we get a new version of the latter by replacing the rules (R1′) and (R2′), respectively, with

(R1) The tree has only numbered circles ( with arbitrary values of k) and black dots (•) as its end nodes,

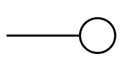
and

(R2) Any internal node has an entering (line that is the root line of a) subtree containing at least one numbered circle as an end node.

Let us define the *degree* of a tree as the positive integer

$\deg T := \#($  $) + \#($  $) + \sum_{k=1}^{\infty} k\,\#($  $)$ (5.8)

for any tree T satisfying (R1) and (R2). By $\#(G)$ we mean the number of occurrences of the graph G in the tree T. That is, the degree of a tree is the number of its end nodes with suitable weights plus the number of nodes with precisely one entering line. Since a tree has finitely many nodes, its degree is well-defined. Then a rearrangement of the sum arising from (5.7) being inserted into (5.6) yields formally

 $= \sum_{l=1}^{\infty} \sum^{*}_{\substack{\text{trees } T \\ \deg T = l}} T,$ (5.9)

where the asterisk reminds us that the rules (R1) and (R2) are being respected.

According to the analysis above, the particular graphs appearing in the definition of $\deg T$ are the only possible single-node subgraphs of T proportional to a positive power of ϵ. Since each tree is an analytic function of ϵ, writing again $(\,\cdot\,)^k$ for the kth coefficient of the power series in ϵ, we have

$$T = \sum_{k=\deg T}^{\infty} \epsilon^k\, T^k = \epsilon^{\deg T} \sum_{k=0}^{\infty} \epsilon^k\, T^{k+\deg T}.$$

Hence, only trees with degree *at most* equal to ℓ can contribute to $\delta_2 \widetilde{X}^\ell$:

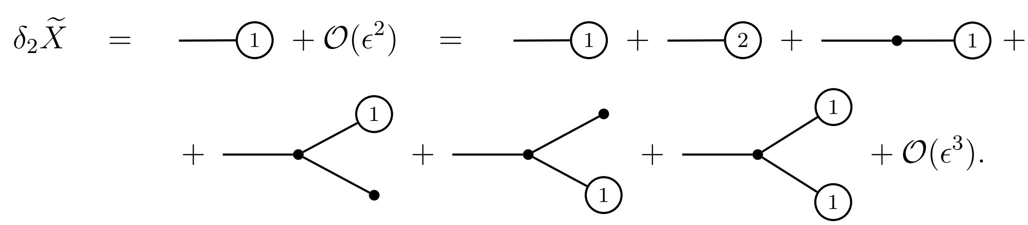
$$\delta_2 \widetilde{X}^\ell = \sum_{l=1}^{\ell} \sideset{}{^*}\sum_{\deg T = l} T^\ell = \Bigg(\sideset{}{^*}\sum_{\deg T \le \ell} T \Bigg)^{\ell} \tag{5.10}$$

or, alternatively,

$$\delta_2 \widetilde{X} = \sideset{}{^*}\sum_{\deg T \le \ell} T + \mathcal{O}(\epsilon^{\ell+1}) \qquad (\epsilon \to 0) \tag{5.11}$$

for *each and every* $\ell = 1, 2, \ldots$. The expansion in (5.9) is in fact just a compact way of writing (5.11). We emphasize that the latter can be derived completely rigorously, for each value of ℓ separately, but resorting to the use of formal series allowed us to treat all orders of $\delta_2 \widetilde{X}$ at once. We call the series (5.9) an *asymptotic expansion* of $\delta_2 \widetilde{X}$; the partial sums $\sum^*_{\deg T \le \ell} T$ need not converge to $\delta_2 \widetilde{X}$ for any fixed ϵ as $\ell \to \infty$, but for a fixed ℓ the error is bounded by an ℓ-dependent constant times $|\epsilon|^{\ell+1}$ on the mutual domain of analyticity, $|\epsilon| < \epsilon_0$.

Example 5.1. The beginning of the asymptotic expansion (5.11) reads



5.2. **Analyticity domain of trees.** As already pointed out, all trees T above are analytic functions of (z, θ, ϵ) on $\Pi_\tau \times \{|\epsilon| < \epsilon_0\}$. Due to the projections δ_2 appearing in (5.4), they also satisfy $T|_{z=0} = \partial_z T|_{z=0} = 0$, *i.e.*, are elements of the space $\mathcal{A}_1$ defined in (4.7). On this space, the inverse of $\mathcal{K} = \left(\begin{smallmatrix} L & 0 \\ 0 & \mathcal{L}^2 \end{smallmatrix}\right)$ (see (4.2)) constructed in the proof of Lemma 4.3 satisfies

$$\mathcal{K}^{-1} h(z, \theta) = \int_{-\infty}^{0} \widetilde{K}(s; z)\, h(z e^{\gamma s}, \theta + \omega s)\, ds. \tag{5.12}$$

Consequently, we will now show that the analyticity domain of a tree in the z-variable is in fact much larger than the neighbourhood of $[-1, 1]$ that is included in Π_τ; namely it includes the wedgelike region

$$\mathbb{U}_{\tau,\vartheta} := \big\{|z| \le \tau\big\} \bigcup \big\{\arg z \in [-\vartheta, \vartheta] \cup [\pi - \vartheta, \pi + \vartheta]\big\} \subset \mathbb{C}$$

(with a new τ and "small" ϑ).

Lemma 5.2 (Analytic continuation of trees)**.** *Without affecting the analyticity domain with respect to* ϵ*, there exist numbers* $0 < \tau < 1$*,* $0 < \vartheta < \pi/2$*, and* $0 < \sigma < \eta$ *such that each tree in the sums* (5.9) *and* (5.11) *extends to an analytic function of* (z, θ) *on* $\mathbb{U}_{\tau,\vartheta} \times \{|\mathfrak{Im}\,\theta| \leq \sigma\}$*.*

Proof. Observe that, as a polynomial, $\widetilde{X}_{\leq 1}$ is an entire function of z. On the other hand, $\Phi^0(z) = 4\arctan z = 2i\big(\log(1-iz) - \log(1+iz)\big)$, implying that $|\mathfrak{Im}\,\Phi^0(z)| \leq \eta$ in $\mathbb{U}_{\tau,\vartheta}$ with τ and ϑ sufficiently small. By Remark 1.3, $f(\Phi^0(z), \theta)$ is analytic, making the maps $h^{(k)}$ and $\widetilde{X}_{\leq 1}$ analytic on $\mathbb{U}_{\tau,\vartheta} \times \{|\mathfrak{Im}\,\theta| \leq \sigma\}$ for some $0 < \sigma < \eta$, where η is determined by f and σ by $\widetilde{X}_{\leq 1}$ (ultimately by f and ω).

Suppose $h = \delta_2 h$ is a map analytic on $\mathbb{U}_{\tau,\vartheta} \times \{|\mathfrak{Im}\,\theta| \leq \sigma\}$. Then the integrand in (5.12) is analytic in a neighbourhood of the latter set. By virtue of Fubini's theorem,

$$\oint_\Gamma \mathcal{K}^{-1} h(\zeta, \theta)\, d\zeta = \int_{-\infty}^0 \oint_\Gamma \widetilde{K}(s;\zeta)\, h(\zeta e^{\gamma s}, \theta + \omega s)\, d\zeta\, ds = 0$$

for any closed contour Γ inside a sufficiently small neighbourhood of $\mathbb{U}_{\tau,\vartheta}$ and enclosing z. Hence, Morera's theorem yields analyticity of $\mathcal{K}^{-1}h$ with respect to z. As always, analyticity with respect to θ follows from an exponentially decaying bound on the Fourier coefficients. Applying this argument at each node of a tree proves the claim. □

Proof of Theorem 2. Since the number of terms in the sum in (5.10) is finite and the functions $\widetilde{X}_{\leq 1}$ and X^0 in $X = X^0 + \widetilde{X}_{\leq 1} + \delta_2 \widetilde{X}$ are analytic on $\mathbb{U}_{\tau,\vartheta} \times \{|\mathfrak{Im}\,\theta| \leq \sigma\}$, the analyticity of X^ℓ follows from Lemma 5.2.

From the equations of motion, (1.11), a Taylor expansion yields

$$\mathcal{L}^2 \widetilde{X} = -\mathcal{L}^2 X^0 + \Omega(X^0) + \sum_{m=1}^\infty \frac{1}{m!} D^m \Omega(X^0)(\widetilde{X})^{\otimes m},$$

where the trigonometric degree of $D^m\Omega(X^0)$ is N for $\epsilon \neq 0$ but vanishes at $\epsilon = 0$ because X^0 does not depend on θ. For each $k \geq 1$, let n_k stand for the trigonometric degree of $\widetilde{X}^k$. Equating like powers of ϵ in the expansion above, we infer two things. First, $n_1 = N$. Second, we must have, for each $\ell \geq 2$,

$$n_\ell \leq \begin{cases} n_{k_1} + \cdots + n_{k_m} & \text{where } k_1 + \cdots + k_m = \ell, \\ N + n_{k_1} + \cdots + n_{k_m} & \text{where } k_1 + \cdots + k_m = \ell - 1, \end{cases}$$

because the trigonometric degree of a product is at most the sum of the trigonometric degrees of the factors; $e^{iq\cdot\theta}e^{iq\cdot\theta} = e^{i2q\cdot\theta}$ and $e^{iq\cdot\theta}e^{-iq\cdot\theta} = 1$.

Next, assume that $n_k \leq kN$ holds for each $1 \leq k \leq \ell - 1$, recalling that this is the case if $k = 1$. Subsequently, the estimate for n_ℓ above becomes $n_\ell \leq \ell N$. □

References

[AA68] V. I. Arnold and A. Avez, *Ergodic problems of classical mechanics*, Translated from the French by A. Avez, W. A. Benjamin, Inc., New York-Amsterdam, 1968. MR MR0232910 (38 #1233)

[Ahl66] L. V. Ahlfors, *Complex Analysis*, 2nd ed., McGraw-Hill, 1966.

[Arn63] V. I. Arnold, *Proof of a Theorem of A. N. Kolmogorov on the Preservation of Conditionally-periodic Motions under a Small Perturbation of the Hamiltonian*, Russ. Math. Surv **18** (1963), no. 5, 9–36.

[Arn64] ———, *Instability of dynamical systems with many degrees of freedom*, Dokl. Akad. Nauk SSSR **156** (1964), 9–12. MR MR0163026 (29 #329)

[BGK99] J. Bricmont, K. Gawędzki, and A. Kupiainen, *KAM theorem and quantum field theory*, Comm. Math. Phys. **201** (1999), no. 3, 699–727. MR 1685894

[CG94] L. Chierchia and G. Gallavotti, *Drift and diffusion in phase space*, Ann. Inst. H. Poincaré Phys. Théor. **60** (1994), no. 1, 144. MR 1259103

[Cha85] S. B. Chae, *Holomorphy and Calculus in Normed Spaces*, Marcel Dekker, New York, 1985.

[Chi79] B. V. Chirikov, *A universal instability of many-dimensional oscillator systems*, Phys. Rep. **52** (1979), no. 5, 264–379. MR MR536429 (80h:70022)

[Chi96] L. Chierchia, *Lezioni di Analisi matematica 2*, Aracne Editrice, 1996.

[Din99] S. Dineen, *Complex Analysis on Infinite Dimensional Spaces*, Springer-Verlag, London, 1999.

[Eli94] L. H. Eliasson, *Biasymptotic solutions of perturbed integrable Hamiltonian systems*, Bol. Soc. Brasil. Mat. (N.S.) **25** (1994), no. 1, 57–76. MR MR1274762 (95b:58133)

[Eli96] ———, *Absolutely convergent series expansions for quasi periodic motions*, Math. Phys. Electron. J. **2** (1996), Paper 4, 33 pp. (electronic), here the preprint of 1988 is finally published. MR 1399458

[Gal94a] G. Gallavotti, *Twistless KAM tori*, Commun. Math. Phys. **164** (1994), no. 1, 145–156. MR 1288156

[Gal94b] ———, *Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable Hamiltonian systems. A review*, Rev. Math. Phys. **6** (1994), no. 3, 343–411. MR 1305589

[Gen95a] G. Gentile, *A proof of existence of whiskered tori with quasi-flat homoclinic intersections in a class of almost integrable Hamiltonian systems*, Forum Math. **7** (1995), no. 6, 709–753. MR MR1359423 (97a:58164b)

[Gen95b] ———, *Whiskered tori with prefixed frequencies and Lyapunov spectrum*, Dynam. Stability Systems **10** (1995), no. 3, 269–308. MR MR1356323 (97a:58164a)

[GM96] G. Gentile and V. Mastropietro, *Methods for the analysis of the Lindstedt series for KAM tori and renormalizability in classical mechanics. A review with some applications*, Rev. Math. Phys. **8** (1996), no. 3, 393–444. MR MR1388257 (98f:58169)

[Gra74] S. M. Graff, *On the conservation of hyperbolic invariant tori for Hamiltonian systems*, J. Differential Equations **15** (1974), 1–69. MR MR0365626 (51 #1878)

[Kol54] A. N. Kolmogorov, *On Conservation of Conditionally-periodic Motions for a Small Change in Hamilton's Function*, Dokl. Akad. Nauk SSSR **98** (1954), no. 4, 525–530, (Russian).

[Laz03] V. F. Lazutkin, *Splitting of separatrices for the Chirikov standard map*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **300** (2003), no. Teor. Predst. Din. Sist. Spets. Vyp. 8, 25–55, 285, The famous article from 1984. MR 1993024

[LMS03] P. Lochak, J.-P. Marco, and D. Sauzin, *On the splitting of invariant manifolds in multidimensional near-integrable Hamiltonian systems*, Mem. Amer. Math. Soc. **163** (2003), no. 775, viii+145. MR 1964346

[Loc99] P. Lochak, *Arnold diffusion; a compendium of remarks and questions*, Hamiltonian systems with three or more degrees of freedom (S'Agaró, 1995), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., vol. 533, Kluwer Acad. Publ., Dordrecht, 1999, pp. 168–183. MR MR1720892 (2001g:37107)

[Mel63] V. K. Melnikov, *On the stability of a center for time-periodic perturbations*, Trudy Moskov. Mat. Obšč. **12** (1963), 3–52. MR MR0156048 (27 #5981)

[Mos62] J. Moser, *On Invariant Curves of Area-preserving Mappings of an Annulus*, Nachr. Akad. Wiss. Göttingen, Math. Phys. Kl IIa (1962), 1–20.

[Mos66a] ———, *A rapidly convergent iteration method and non-linear partial differential equations. I*, Ann. Scuola Norm. Sup. Pisa (3) **20** (1966), 265–315. MR MR0199523 (33 #7667)

[Mos66b] ———, *A rapidly convergent iteration method and non-linear partial differential equations. II*, Ann. Scuola Norm. Sup. Pisa (3) **20** (1966), 499–535. MR MR0206461 (34 #6280)

[Mos67] ———, *Convergent series expansions for quasi-periodic motions*, Math. Ann. **169** (1967), 136–176. MR MR0208078 (34 #7888)

[Poi93a] H. Poincaré, *New methods of celestial mechanics. Vol. 1*, History of Modern Physics and Astronomy, vol. 13, American Institute of Physics, New York, 1993, Periodic and asymptotic solutions, Translated from the French, Revised reprint of the 1967 English translation, With endnotes by V. I. Arnol′d, Edited and with an introduction by Daniel L. Goroff. MR 1194622

[Poi93b] ———, *New methods of celestial mechanics. Vol. 3*, History of Modern Physics and Astronomy, vol. 13, American Institute of Physics, New York, 1993, Integral invariants and asymptotic properties of certain solutions, Translated from the French, Revised reprint of the 1967 English translation, With endnotes by G. A. Merman, Edited and with an introduction by Daniel L. Goroff. MR 1194624

[Wei] E. W. Weisstein, *Infinite Products*, Eric Weisstein's World of Mathematics, `http://mathworld.wolfram.com/`.

E-mail address: `mikko.stenlund@helsinki.fi`

Department of Mathematics and Statistics, University of Helsinki, P.O. Box 68, FI-00014 University of Helsinki, Finland