The critical renormalization fixed point for commuting pairs of area-preserving maps

Gianni Arioli [1,2] and Hans Koch [3]

Abstract. We prove the existence of the critical fixed point (F, G) for MacKay's renormalization operator for pairs of maps of the plane. The maps F and G commute, are area-preserving, reversible, real analytic, and they satisfy a twist condition.

1. Introduction

We consider the fixed point problem for the following operator $\mathfrak{R}$, acting on pairs of area-preserving maps $\mathcal{P} = (F, G)$ of the plane:

$$\mathfrak{R}(\mathcal{P}) = (\widetilde{F}, \widetilde{G}), \qquad \widetilde{F} = \Lambda^{-1} G \Lambda, \qquad \widetilde{G} = \Lambda^{-1} F G \Lambda. \tag{1.1}$$

Here, Λ is a linear scaling $(x, z) \mapsto (\lambda x, \mu z)$, depending on $\mathcal{P}$, with λ and μ defined by the condition $\widetilde{G}(0, 0) = (-1, -1)$. Our main result is the following.

Theorem 1.1. *The transformation $\mathfrak{R}$ has a fixed point (F, G), with associated scalings*

$$\lambda = -0.7067956691\ldots, \qquad \mu = -0.3260633966\ldots. \tag{1.2}$$

The maps F and G are area-preserving, real analytic, and nonlinear. In addition, they satisfy the following, on non-empty open subsets of their domains. G is reversible with respect to the involution $S(x, z) = (-x, z)$, in the sense that $SGS = G^{-1}$. The same holds for F. Furthermore, F and G commute.

Our proof of this theorem is computer-assisted.

This fixed point problem has a rather long history [1–7]. The fixed point described in Theorem 1.1, is known as the "critical" fixed point of $\mathfrak{R}$. Its existence was conjectured in [2], based on a numerical investigations; and some rigorous partial results were obtained in [7]. A related fixed point problem for Hamiltonians was solved recently in [13,14].

The motivation behind these studies is to describe the breakup of golden invariant circles in one-parameter families of maps, such as the standard family

$$(x, z) \mapsto (x + w, w), \qquad w = z - \beta \sin(2\pi x). \tag{1.3}$$

For $\beta = 0$, the map (1.3) has a smooth invariant circles, including one (at $z = \vartheta^{-1}$) whose rotation number is the inverse of the golden mean $\vartheta = \frac{1}{2}\sqrt{5} + \frac{1}{2}$. By KAM theory, the same holds for small $\beta > 0$. The golden circle is observed to persist as β is increased, up to some

---

[1] Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano.

[2] This work was supported in part by MIUR project "Equazioni alle derivate parziali e disuguaglianze funzionali".

[3] Department of Mathematics, University of Texas at Austin, Austin, TX 78712

value β_∞, where it starts to break up. The transition is characterized by several numerical quantities that are universal, in the sense that the exact same values are observed in a large class of one-parameter families of cylinder maps $\beta \mapsto G_\beta$. In particular, the critical map in the family (the map G_β for $\beta = \beta_\infty$) has a non-smooth golden invariant circle. This circle, and the entire orbit structure nearby, is invariant under a scaling $\Lambda = \mathrm{diag}(\lambda, \mu)$, with λ and μ being the values (1.2).

The maps (1.3) commute with integer translations $\mathcal{F}_n(x,z) = (x+n, z)$, so they define maps on the cylinder $\mathbb{T} \times \mathbb{R}$. In a more general situation, the cylinder can be $\mathbb{R}^2/\mathcal{F}$, where $\mathcal{F}$ is a $\mathbb{Z}$-action on $\mathbb{R}^2$, generated by some diffeomorphism $\mathcal{F}_1$. A map on this cylinder can be described by a map $G : \mathbb{R}^2 \to \mathbb{R}^2$ that commutes with $\mathcal{F}_1$. If we think of $\mathcal{F}_1$ as a full rotation of the cylinder, then the equation for a periodic point (x,z) of G, with rotation number $\frac{p}{q}$, is $G^q(x,z) = \mathcal{F}_p(x,z)$. We refer to such a point as being $\frac{p}{q}$-periodic for the pair (F,G), where F is the inverse of $\mathcal{F}_1$. Then a $\frac{p}{q}$-periodic point for (F,G) is just a fixed point for F^pG^q. At the same time it is a $\frac{q-p}{p}$-periodic point for (G, FG), since $F^pG^q = G^{q-p}(FG)^p$. Now consider the continued fraction approximants $r_1 = \frac{1}{2}$, $r_2 = \frac{2}{3}$, $r_3 = \frac{3}{5}$, $r_4 = \frac{5}{8}$, … for the inverse golden mean ϑ^{-1}, and let $n > 1$. Then a r_n-periodic point for (F,G) is a r_{n-1}-periodic point for (G, FG). Furthermore, taking $n \to \infty$ yields an analogous statement about quasiperiodic points with rotation number ϑ^{-1}.

Coming back to the scaling properties of critical cylinder maps, one observation is the following: After a suitable change of coordinates, a sequence of r_n-periodic points (x_n, z_n) accumulates at the origin (lying on the golden circle), with asymptotic ratios $x_{n+1}/x_n \to \lambda$ and $z_{n+1}/z_n \to \mu$. This motivates a rescaling of the map $(F,G) \mapsto (G, FG)$, as in (1.1), and it suggests that, under iteration of $\mathfrak{R}$, a critical area-preserving cylinder map converges to the critical fixed point. In this "renormalization picture", the value β_∞ marks the point where a family $\beta \mapsto \mathcal{P}_\beta$ crosses the (codimension 1) stable manifold of the fixed point. Another universal quantity that can be observed in such families is a value $\delta \approx 1.62795$. It describes the geometric accumulation, at β_∞, of bifurcation points β_n involving r_n-periodic orbits. This number δ is expected to be the expanding eigenvalue of the derivative $D\mathfrak{R}$ at the fixed point. Estimates on δ require a study of this derivative, which we have not done here; but it could be the subject for future work.

Notice that $\mathfrak{R}$ preserves the commuting property of pairs. However, the constraint $FG = GF$ is highly impractical to work with. Thus, we drop it for the time being. Another problem is reversibility: it is not preserved by $\mathfrak{R}$. What does preserve reversibility are "palindromic" compositions like GFG. Thus, we start by considering the transformation

$$\mathfrak{N}(G) = \Lambda^{-1} F G_1 \Lambda\,, \qquad G_1 = \Lambda^{-1} F G \Lambda\,, \qquad F = \Lambda^{-1} G \Lambda\,. \tag{1.4}$$

Here, λ and μ are defined by the equation $G(-\lambda, -\mu) = (-\lambda^2, -\mu^2)$. This guarantees that the normalization condition $G(0,0) = (-1,-1)$ is satisfied again by $\mathfrak{N}(G)$, as well as by G_1. Clearly, if G is a fixed point for $\mathfrak{N}$, satisfying $G_1 = G$, then the corresponding pair (F,G) is a fixed point of $\mathfrak{R}$. Interestingly, the property $G_1 = G$ holds almost automatically: Let G be a reversible fixed point of $\mathfrak{N}$. Then, using that $\Lambda^{-1}G = F\Lambda^{-1}$, we find that $J = G_1^{-1}G$ satisfies

$$J = (\Lambda^{-1} G^{-1} F^{-1} \Lambda)(\Lambda^{-2} G F G \Lambda^2) = \Lambda^{-1} G^{-1} \Lambda^{-1} F G \Lambda^2 = \Lambda^{-1} J^{-1} \Lambda\,. \tag{1.5}$$

In addition, J leaves the origin invariant. Assuming J is analytic near the origin, a simple power series argument, using that $|\mu|$ is not an integer power of $|\lambda|$, shows that the conjugacy (1.5) implies $J = \pm\mathrm{I}$. By proving that J is analytic and different from $-\mathrm{I}$, we can conclude that $G_1 = G$. Hence the pair (F, G) is a fixed point of $\mathfrak{R}$. In addition, the identity $G_1 = G$ implies that G_1 is reversible, which in turn implies that F and G commute. The complete argument is a bit more involved, since the domains of these maps need to be considered.

Another constraint is that the map G has to preserve area. We deal with this in the usual way, by representing G in terms of a generating function g,

$$G(x,z) = (y,w)\,, \qquad z = -g_1(x,y)\,, \quad w = g_2(x,y)\,, \tag{1.6}$$

where $g_j = \partial_j g$. In other words, the fixed point problem $\mathfrak{N}(G) = G$ is translated into a fixed point problem $\mathcal{N}(g) = g$. After solving the latter, and verifying that $g_{1,2} > 1$ on the relevant domain, we obtain the desired fixed point of $\mathfrak{N}$ by solving equation (1.6). The twist property $g_{1,2} > 0$ guarantees that the solution is unique. Furthermore, since the one-form $wdy - zdx$ is the differential of g, and thus closed, the resulting map G is area-preserving. For simplicity, we reconstruct G and define F on domains that are (each) a union of two overlapping rectangles, satisfying

$$\Lambda\overline{D_G} \subset D_G\,, \qquad \Lambda\overline{D_F} \subset D_G\,, \qquad G\Lambda\overline{D_G} \subset D_F\,. \tag{1.7}$$

Except for the closures, these are the minimal conditions for $\mathfrak{R}(\mathcal{P})$ to be defined on the domain of $\mathcal{P}$, regardless of the order in which F and G are composed. In [6], a fixed point of $\mathfrak{R}$ that satisfies (1.7) is said to have the *extension property.* When combined with additional properties, it is possible to prove a number of interesting facts about cylinder maps that are attracted to such a fixed point under iteration of $\mathfrak{R}$. This includes the existence of a golden invariant circle. For further details we refer to [6]. We did not attempt to verify any of these additional properties, but our computer programs should be well suited for such (and other) investigations.

A related fixed point problem was considered in [7], namely $F = \Lambda^{-3}GFG\Lambda^3$ and $G = \Lambda^{-3}GFGFG\Lambda^3$. This is a palindromic modification of the equation $\mathfrak{R}^3(\mathcal{P}) = \mathcal{P}$. It was proved in [7] that the corresponding fixed point equation for a reduced generating function (corresponding our g_1) has a solution; and the bounds obtained for λ and μ are compatible with (1.2). What was left open is the question of whether the corresponding maps F and G commute (which would yield a fixed point or period 3 for $\mathfrak{R}$, assuming that F and G have proper domains), and whether they are area-preserving.

As indicated earlier, there are analogues of the transformation (1.1), that act on Hamiltonians. An overview of the work in this area, which goes back to [8, 9], can be found in [10, 11]. One such transformation, that preserves analyticity, was proved to have a critical fixed point [13], with a non-smooth invariant torus [14]. Naturally, the scaling constants for Hamiltonians agree with (1.1). As one would expect, there is a connection between $\mathfrak{R}$ and its Hamiltonian analogue. However, this connection [12] is purely formal, due to unknown domains, and a direct analysis of $\mathfrak{R}$ seemed more promising (and interesting) than trying to make this connection rigorous.

One problem with $\mathfrak{R}$ has always been the need to work with commuting pairs; but our argument following (1.5) shows that this issue is less serious than it seems. It should be possible to extend our methods to obtain results on the derivative of $\mathfrak{R}$ at the critical fixed point. An analogous analysis for Hamiltonians seems currently out of reach, due to the complexity of the transformation involved.

Our results on the transformation $\mathcal{N}$ for generating functions, and on a related contraction $\mathcal{M}$, can be found in Sections 4 and 5, respectively. The relevant function spaces, and some basic estimates, are given in Section 3. In Section 6, we discuss an implicit equation that arises in the definition of $\mathcal{N}$. A description of our computer-assisted proof can be found in Section 7. We start by defining the transformation $\mathcal{N}$.

2. Generating functions

The transformation $\mathcal{N}$ for generating functions is formally $\mathcal{N} = \Psi^{-1}\mathfrak{N}\Psi$, where Ψ is the map that assigns an area-preserving map G to a generating function g, via equation (1.6). Our aim here is to give an explicit but formal description of $\mathcal{N}$, and of its derivative.

We start with some simple facts about generating functions and use the opportunity to introduce some notation. As can be verified using (1.6), the generating function $f \star g$ for a composed map FG is given by

$$(f \star g)(x,y) = g(x,\mathcal{V}) + f(\mathcal{V},y)\,, \tag{2.1}$$

with the "midpoint" $\mathcal{V} = \mathcal{V}(x,y)$ determined by the equation

$$\frac{d}{d\mathcal{V}}\bigl[g(x,\mathcal{V}) + f(\mathcal{V},y)\bigr] = 0\,. \tag{2.2}$$

A conjugacy $F = \Lambda^{-1}G\Lambda$ by a scaling $\Lambda = \operatorname{diag}(\lambda,\mu)$ translates into

$$f = (\lambda\mu)^{-1} g \circ \ell\,, \qquad \ell(x,y) = (\lambda x, \lambda y)\,. \tag{2.3}$$

For the generating function g^{-1} of the inverse map G^{-1} we get $g^{-1}(x,y) = -g(y,x)$. From this, one sees that the generating function $\mathbb{S}g$ for $SG^{-1}S$ is given by

$$(\mathbb{S}g)(x,y) = g(-y,-x)\,. \tag{2.4}$$

A function g that is invariant under $\mathbb{S}$ will be called *symmetric*.

Now we consider the generating function analogue of each step in the definition of

$$\mathfrak{N}(G) = \Lambda^{-2}GFG\Lambda^2\,, \qquad F = \Lambda^{-1}G\Lambda\,. \tag{2.5}$$

We assume that G is reversible and satisfies the normalization $G(0,0) = (-1,-1)$. Thus, the corresponding generating function has to be symmetric and satisfy

$$g_1(0,-1) = 0 \qquad g_2(0,-1) = -1\,. \tag{2.6}$$

The generating function f of F is given by (2.3), once we have determined the scaling constants λ and μ. We would like λ and μ to yield the sequence

$$(0,0)\xmapsto{\Lambda^2}(0,0)\xmapsto{G}(-1,-1)\xmapsto{F}(-\lambda,-\mu)\xmapsto{G}(-\lambda^2,-\mu^2)\xmapsto{\Lambda^{-2}}(-1,-1)\,, \tag{2.7}$$

so that $\mathfrak{N}(G)$ is again properly normalized. Notice that the map $G_1=\Lambda^{-1}FG\Lambda$ then satisfies $G_1(0,0)=(-1,-1)$ as well. The first and last condition (arrow) in (2.7) hold for any choice of λ and μ. The second is just the normalization of G, and the third follows from the fourth by the definition of F. So we determine λ and μ by the fourth condition, or equivalently, by the equation

$$g_1(-\lambda,-\lambda^2)=\mu\,,\qquad g_2(-\lambda,-\lambda^2)=-\mu^2\,. \tag{2.8}$$

Next, consider the composed map $\widehat{H}=GFG$. Applying the identity (2.1) twice, we see that the generating function $\widehat{h}$ of $\widehat{H}$ is given by

$$\widehat{h}(x,y)=g(x,\mathcal{V})+f(\mathcal{V},\mathcal{W})+g(\mathcal{W},y)\,, \tag{2.9}$$

with $\mathcal{V}=\mathcal{V}(x,y)$ and $\mathcal{W}=\mathcal{W}(x,y)$ making the right hand side of (2.9) stationary:

$$g_2(x,\mathcal{V})+f_1(\mathcal{V},\mathcal{W})=0\,,\qquad f_2(\mathcal{V},\mathcal{W})+g_1(\mathcal{W},y)=0\,. \tag{2.10}$$

A simple calculation, using that both f and g are symmetric, shows that the second equality in (2.10) follows from the first, if $\mathcal{W}=-\mathbb{S}\mathcal{V}$. This identity also ensures that $\widehat{h}$ is symmetric. Thus, it suffices to solve

$$g_2(x,\mathcal{V})+f_1\bigl(\mathcal{V},-\mathbb{S}\mathcal{V}\bigr)=0\,. \tag{2.11}$$

Once this equation is solved, and $\widehat{h}$ is defined via (2.9), the remaining step $\widetilde{G}=\Lambda^{-2}\widehat{H}\Lambda^2$ translates to

$$\widetilde{g}(x,y)=(\lambda\mu)^{-2}\widehat{h}(\lambda^2x,\lambda^2y)\,. \tag{2.12}$$

Here, we have used (2.3) again. Given that $\widetilde{G}=\mathfrak{N}(G)$, the map $g\mapsto\widetilde{g}$ is the desired transformation $\mathcal{N}$.

We will also need to estimate the derivative of $\mathcal{N}$, so let us now compute $\widetilde{g}'=D\mathcal{N}(g)g'$. This is easier than one might think. We assume that both g and g' are symmetric. Using the symmetry of g, the equation (2.8) for λ can be written as $K(g,\lambda)=0$, where

$$K(g,\lambda)=g_1(\lambda^2,\lambda)-g_2(\lambda^2,\lambda)^2\,. \tag{2.13}$$

Setting $DK(g,\lambda)(g',\lambda')=0$ and solving for λ', we find that

$$\lambda'=-\bigl(2\lambda g_{1,1}+g_{1,2}-2g_2[2\lambda g_{2,1}+g_{2,2}]\bigr)^{-1}\bigl[g_1'-2g_2g_2'\bigr]\,, \tag{2.14}$$

where all functions are being evaluated at (λ^2,λ). Since $\mu=-g_2(\lambda^2,\lambda)$, the corresponding variation of μ is

$$\mu'=-g_2'-[2\lambda g_{2,1}+g_{2,2}]\lambda'\,. \tag{2.15}$$

Then the variation of $f = (\lambda\mu)^{-1} g \circ \ell$ is given by

$$f' = (\lambda\mu)^{-1} g' \circ \ell - (\lambda'/\lambda + \mu'/\mu) f + (\lambda'/\lambda)\mathcal{D} f\,, \tag{2.16}$$

where $\mathcal{D}$ denotes the generator of dilations, $(\mathcal{D}f)(x,y) = x f_1(x,y) + y f_2(x,y)$. The variation $\widehat{h}'$ of the function $\widehat{h}$ in the composition (2.9) is simply

$$\widehat{h}'(x,y) = g'(x,\mathcal{V}) + f'(\mathcal{V},\mathcal{W}) + g'(\mathcal{W},y)\,, \tag{2.17}$$

since the right hand side of (2.9) is stationary with respect to variations of $\mathcal{V}$ and $\mathcal{W}$. The last step in the definition of $\mathcal{N}$ is the scaling $\widetilde{g} = (\lambda\mu)^{-2}\widehat{h} \circ \ell^2$. Its variation is analogous to (2.16), so the function $\widetilde{g}' = D\mathcal{N}(g)g'$ is given by

$$\widetilde{g}' = (\lambda\mu)^{-2}\widehat{h}' \circ \ell^2 - 2(\lambda'/\lambda + \mu'/\mu)\widetilde{g} + 2(\lambda'/\lambda)\mathcal{D}\widetilde{g}\,. \tag{2.18}$$

Notice that the basic steps involved in the construction of $\mathcal{N}(g)$ and $D\mathcal{N}(g)g'$ are derivatives, composition of functions, and the solution of implicit equations.

3. Function spaces

In order to control the steps described in the last section, we first have to choose appropriate domains and function spaces. Since we need good approximations for analytic functions, our preference is to use Taylor series, and domains that are disks (in each variable).

The equation (2.4) shows that the generating function g for a reversible map G is an even function of $x + y$. Thus, it is natural to change variables to

$$t = x + y\,, \qquad s = x - y\,. \tag{3.1}$$

However, using a domain of the type $|t - t_0| < \rho_t$ for the variable t poses problems. Expanding about $t_0 = 0$ is essentially useless, even numerically. And using a disk about $t_0 \neq 0$ is not a workable option, since we need a reasonable subspace of even functions. Here, and in what follows, we call a function *even* if it is an even function of t. A possible way out is to write our functions as $P + tQ$, with P and Q even; then expand P and Q in powers of $u = t^2 - t_0^2$ and $v = s - s_0$. But it turns out that the resulting domains are too borderline for a successful analysis of $\mathcal{N}$. What improves the situation drastically is a choice of variables of the form

$$u = \left[t^2 - t_0^2\right] + b[s - s_0]\,, \qquad v = s - s_0\,, \tag{3.2}$$

with b substantially different from 0. Specific values for the parameters t_0, s_0, and b will be given later.

The corresponding function spaces are chosen as follows. Given a pair of positive real numbers $\rho = (\rho_u, \rho_v)$, denote by D_ρ the set of points $(u,v) \in \mathbb{C}^2$ such that $|u| < \rho_u$ and

$|v| < \rho_v$. Define $\mathcal{A}_\rho^\diamond$ to be the space of all analytic functions $P : D_\rho \to \mathbb{C}$, that extend continuously to the boundary of D_ρ, equipped with the norm

$$\|P\|_\rho = \sum_{m,n} |P_{m,n}| \rho_u^m \rho_v^n , \qquad P(u,v) = \sum_{m,n} P_{m,n} u^m v^n . \tag{3.3}$$

Clearly, $\mathcal{A}_\rho^\diamond$ is a Banach algebra, that is, $\|PQ\|_\rho \le \|P\|_\rho \|Q\|_\rho$. If $\mathcal{A}$ is any complex Banach algebra with unit, then for $U, V \in \mathcal{A}$ we define $P(U,V) = \sum_{m,n} P_{m,n} U^m V^n$, provided that the series converges in $\mathcal{A}$.

Before extending the above to include non-even functions, we give here the bound that is used (in our programs) to estimate the various derivatives that appear in the construction of $\mathcal{N}(g)$ and $D\mathcal{N}(g)g'$. Given positive real numbers $\sigma < \tau$, and a non-negative integer k, define

$$W_k(\sigma,\tau) = \max_{m \ge k} W_{k,m}(\sigma,\tau) , \qquad W_{k,m}(\sigma,\tau) = \frac{m!}{(m-k)!} \left(\frac{\sigma}{\tau}\right)^m \sigma^{-k} . \tag{3.4}$$

Proposition 3.1. *Let $r = (r_u, r_v)$, with $0 < r_u < \rho_u$ and $0 < r_v \le \rho_v$. If $P \in \mathcal{A}_\rho^\diamond$ and $k \ge 0$, then $\partial_u^k P \in \mathcal{A}_r^\diamond$ and $\|\partial_u^k P\|_r \le W_k(r_u, \rho_u) \|P\|_\rho$.*

An analogous bound holds of course for derivatives with respect v.

Proof. With P as in (3.3) we have

$$(\partial_u^k P)(u,v) = \sum_{m \ge k} \sum_{n \ge 0} P_{m,n} \frac{m!}{(m-k)!} u^{m-k} v^n , \tag{3.5}$$

and thus

$$\begin{aligned} \|\partial_u^k P\|_r &\le \sum_{m \ge k} \sum_{n \ge 0} |P_{m,n}| \frac{m!}{(m-k)!} r_u^{m-k} r_v^n \\ &\le \sum_{m \ge k} \sum_{n \ge 0} |P_{m,n}| \rho_u^m \rho_v^n \, W_{k,m}(r_u, \rho_u) \le \|P\|_\rho \, W_k(r_u, \rho_u) , \end{aligned} \tag{3.6}$$

as claimed. **QED**

Consider now a fixed choice of the parameters t_0, s_0, and b. Denote by $\mathcal{D}_\rho$ the set of points $(x,y) \in \mathbb{C}^2$, for which $(\boldsymbol{u}(x,y), \boldsymbol{v}(x,y))$ belongs to D_ρ, where $\boldsymbol{u}$ and $\boldsymbol{v}$ denote the functions $(x,y) \mapsto u$ and $(x,y) \mapsto v$, respectively, defined by the change of variables (3.2) and (3.1). Any function $R : \mathcal{D}_\rho \to \mathbb{C}$ can be written as

$$R = P(\boldsymbol{u}, \boldsymbol{v}) + \boldsymbol{t} Q(\boldsymbol{u}, \boldsymbol{v}) , \tag{3.7}$$

where P and Q are functions on D_ρ, and $\boldsymbol{t}(x,y) = x + y$. We define $\mathcal{A}_\rho$ to be the Banach space of all functions (3.7), with $P, Q \in \mathcal{A}_\rho^\diamond$, equipped with the norm

$$\|R\|_\rho = \|P\|_\rho + \rho_t \|Q\|_\rho , \qquad \rho_t = \left[t_0^2 + \rho_u + |b| \rho_v\right]^{1/2} . \tag{3.8}$$

A function $R \in \mathcal{A}_\rho$ will be called *real*, if both P and Q take real values for real arguments. Notice that

$$\boldsymbol{t}^2 = t_0^2 + \boldsymbol{u} - b\boldsymbol{v}\,, \qquad \|\boldsymbol{t}^2\|_\rho = \rho_t^2 = \|\boldsymbol{t}\|_\rho^2\,. \tag{3.9}$$

From this, it follows readily that $\mathcal{A}_\rho$ is a Banach algebra. The subspace of even functions $R = P(\boldsymbol{u}, \boldsymbol{v})$ will be denoted by $\mathcal{A}_\rho^e$. Clearly, $\mathcal{A}_\rho^e$ is isometrically isomorphic to $\mathcal{A}_\rho^\diamond$.

The spaces $\mathcal{A}_\rho$ are convenient for estimating composed maps. More generally, let $\mathcal{A}$ be any commutative Banach algebra over $\mathbb{C}$, with unit $\mathbf{1}$. Let $X, Y \in \mathcal{A}$ and $T = X + Y$. Given $R \in \mathcal{A}_\rho$ as in (3.7), define $R(X,Y) = P(U,V) + TQ(U,V)$, where $V = X - Y - s_0\mathbf{1}$ and $U = T^2 - t_0^2\mathbf{1} + bV$.

Proposition 3.2. *Using the above definitions, assume that $\|T\| \le \rho_t$, $\|U\| \le \rho_u$ and $\|V\| \le \rho_v$. If R belongs to $\mathcal{A}_\rho$, then $R(U,V)$ belongs to $\mathcal{A}$, and $\|R(U,V)\| \le \|R\|_\rho$. Furthermore, the map $(X,Y) \mapsto R(U,V)$ is analytic, on any open domain in $\mathcal{A} \times \mathcal{A}$ where the assumptions above are satisfied.*

The proof of this proposition is a straightforward exercise in power series. When applied with $\mathcal{A} = \mathcal{A}_\varrho$, it can be used to estimate the composed maps appearing in the midpoint equation (2.11). And in the case $\mathcal{A} = \mathbb{C}$, it implies e.g. that $|R(x,y)| \le \|R\|_\rho$, for all $(x,y) \in \mathcal{D}_\rho$.

The spaces $\mathcal{A}_\rho$ are also convenient for estimating linear operators. In particular, the operator norm of a continuous linear map $\mathbb{L} : \mathcal{A}_\rho^e \to \mathcal{A}_\rho^e$ is given by

$$\|\mathbb{L}\|_\rho = \sup_{m,n} \|\mathbb{L}E_{m,n}\|_\rho\,, \qquad E_{m,n} = c_{m,n}\boldsymbol{u}^m\boldsymbol{v}^n\,. \tag{3.10}$$

Here, $c_{m,n} = \|\boldsymbol{u}^m\boldsymbol{v}^n\|_\rho^{-1}$, so that each of the functions $E_{m,n}$ has norm one.

4. The fixed points of $\mathcal{N}$ and $\mathfrak{R}$

In this section, we describe our main results concerning the transformation $\mathcal{N}$. These results are then used to give a proof of Theorem 1.1 and properties (1.7).

Our domain for the transformation $\mathcal{N}$ is a ball in the space $\mathcal{A}_\varrho^e$, for the parameters

$$t_0 = \tfrac{51}{128}\,, \quad s_0 = \tfrac{307}{256}\,, \quad b = 3\,, \quad \varrho_u = \tfrac{7}{4}\,, \quad \varrho_v = \tfrac{3}{4}\,. \tag{4.1}$$

Different parameter values, and thus variables u and v, are used in our representation of the functions f, $\mathcal{V}$, and $\widehat{h}$. We refer to our computer programs [15] for such details. The values in (4.1) are considered fixed from now on, unless specified otherwise. We remark that these values have not been fine-tuned, despite the appearance. They are binary fractions that are close to our simple (in fact our first) decimal guesses. The same holds for the other parameters values given below.

Theorem 4.1. *The transformation $\mathcal{N}$ has a locally unique real fixed point g in $\mathcal{A}_\varrho^e$, is analytic near this fixed point, and has a compact derivative. The scaling constants λ and μ, associated with the fixed point g via equation (1.6), satisfy the bounds (1.2).*

This theorem will be proved in the next section, by reducing the fixed point problem for $\mathcal{N}$ to a fixed point problem for a contraction $\mathcal{M}$.

Our reconstruction of the maps G and $F = \Lambda^{-1}G\Lambda$ from the generating function g involves the rectangle $R = \{(x,z) \in \mathbb{R}^2 : x_0 < x < x_1 \text{ and } z_0 < z < z_1\}$, with

$$x_0 = -\tfrac{107}{256}\,, \qquad x_1 = \tfrac{9}{8}\,, \qquad z_0 = -6\,, \qquad z_1 = \tfrac{251}{128}\,. \tag{4.2}$$

R contains the origin, and in particular, $\Lambda^2 R \subset R$. We define $D_F = \Lambda^{-1}R_- \cup R_-$ and $D_G = R \cup \Lambda R_+$, where $R_- \approx R \approx R_+$, but $\overline{R_-} \subset R$ and $\overline{R} \subset R_+$, to ensure that the first two of the conditions (1.7) hold. We will not specify the rectangles $R_\pm$ here, except for saying that their corners lie within 2^{-50} of the corners of R.

We note that F and G can be constructed directly, via equation (1.6), in domains that contain all points along the normalization chain (2.7). The "direct" domain for G includes the rectangle R, but it misses points from the S-reflected chain, such as $(\lambda, -\mu)$, which would be convenient for proving that $GF = FG$. This is the reason for including the scaled rectangle ΛR, where G will be defined via extension.

Lemma 4.2. *The fixed point g described in Theorem 4.1 has the following additional properties. The equation (1.6) defines a real analytic map G_0 on R_+ , and G_0 maps $\Lambda^2 R_+$ into $\Lambda^{-1}R_-$. Define $G_1 = \Lambda^{-1}F_0G_0\Lambda$ on ΛR_+ , with $F_0 = \Lambda^{-1}G_0\Lambda$ defined on $\Lambda^{-1}R_+$. Then G_1 maps ΛR_+ into $\Lambda^{-1}R_-$. Furthermore, $J = G_1^{-1}G_0$ is well defined near the origin, and different from $-\mathrm{I}$.*

The analyticity of G_0 follows from the analyticity of g via the implicit function theorem. The remaining part of our proof is computer-assisted. For further information and details we refer to Section 7, and to the code of our programs [15].

We note that the formal identity (1.5) only involves compositions along the chain (2.7), if J is evaluated at the origin. Thus, all these compositions are well defined, with (F, G) replaced by (F_0, G_0), if J is restricted to a small open neighborhood of the origin in $\mathbb{C}^2$. For the same reason, $\mathfrak{N}(G_0) = G_0$ near the origin. This can be used to write $J = H_0^{-1}F_0G_0$, where $H_0 = \Lambda G_0^{-1}\Lambda^{-1}$. It is this the expression for J that will be used to verify that $J \neq -\mathrm{I}$ near the origin.

Proof of Theorem 1.1. The map $J = G_1^{-1}G_0$ has the origin as a fixed point. In addition, it is analytic near the origin (as G_0 is analytic), where it satisfies $\Lambda^{-2}J\Lambda^2 = J$. Expanding J in powers of x and y, comparing the coefficients with those of $\Lambda^{-2}J\Lambda^2$, and using that $|\lambda|^3 < |\mu| < |\lambda|^4$, one readily finds that either $J = \mathrm{I}$ or $J = -\mathrm{I}$. The second alternative has been excluded, so $J = \mathrm{I}$. As a consequence, $G_0 = G_1 = \Lambda^{-1}F_0G_0\Lambda$ near the origin.

By analytic continuation, both G_0 and G_1 are restrictions of a single map G that is real analytic on D_G . Furthermore, since $(G_0\Lambda)\Lambda R_+$ and $(G_1\Lambda)R_+$ are both contained in the domain of F_0 , we have $G = \Lambda^{-1}F_0G\Lambda$ on all of D_G . Now we can replace F_0 by $F = \Lambda^{-1}G\Lambda$, which is defined on all of D_F . This shows that (F, G) is a fixed point of $\mathfrak{R}$, with domains that satisfy (1.7).

The reversibility of G follows from the symmetry of the generating function g. Specifically, we have $G_0SG_0S = \mathrm{I}$ near the origin, since $(0,0)$ and $(SG_0S)(0,0) = (1,-1)$ belong to the domain of G_0 . Similarly, $F_0SF_0S = \mathrm{I}$ near $(1,-1)$, since the point $(-1,-1)$ and its image $(\lambda, -\mu)$ under SF_0 belong to the domain of F_0 . This fact will be used in the equation below.

Reversibility in turn implies that F and G commute: Substituting $G = \Lambda^{-1}FG\Lambda$ into the identity $GSGS = \mathrm{I}$, we find that $FGSFGS = \mathrm{I}$ near the origin. Using the restricted domains yields

$$\mathrm{I} = F_0 G_1 S F_0 G_0 S = F_0 G_1 (S F_0 S)(S G_0 S) = F_0 G_1 F_0^{-1} G_0^{-1} . \tag{4.3}$$

This shows that $GFF^{-1}G^{-1} = FGF^{-1}G^{-1}$ near zero, and thus $GF = FG$ in an open neighborhood of $p = (\lambda, -\mu)$. By analytic continuation, the identity $GF = FG$ holds on the component of $D_{FG} \cap D_{GF}$ containing p. The reversibility domains of F and G extend similarly.

QED

5. Replacing $\mathcal{N}$ by a contraction

For practical purposes, we extend our renormalization procedure to generating functions g that need not satisfy the normalization condition (2.6). The following extension $\mathcal{N}'$ is rather ad-hoc, but it is simple and serves the intended purpose. The condition (2.6) can be written as $Ng = (1, 0)$, where $Ng = (g_1(1, 0), g_2(1, 0))$. Consider the projection $\mathbb{P}$,

$$(\mathbb{P}g)(x, y) = g(x, y) + C_u(g)u + C_v(g)v , \tag{5.1}$$

where $C^u_r(g)$ and $C^v_r(g)$ are determined by the condition $N\mathbb{P}g = (1, 0)$. Notice that $\mathbb{P}$ is linear and bounded, on any space $\mathcal{A}^e_\rho$ for which $\mathcal{D}_\rho$ contains the point $(1, 0)$. Now we define $\mathcal{N}' = \mathcal{N}\mathbb{P}$. Since the normalization condition $Ng = (1, 0)$ is preserved by the transformation $\mathcal{N}$, the generating function $\mathcal{N}'(g)$ is always normalized properly. In particular, a fixed point g of $\mathcal{N}'$ is also a fixed point of $\mathcal{N}$.

Next, we convert the fixed point equation $\mathcal{N}'(g) = g$ to a fixed point equation for a map $\mathcal{M}$ that can be expected to be a contraction. For $\mathcal{M}$ we choose a Newton-type map

$$\mathcal{M}(\gamma) = \gamma + \mathcal{N}'(g_0 + M\gamma) - (g_0 + M\gamma) , \qquad \mathcal{N}' = \mathcal{N}\mathbb{P} . \tag{5.2}$$

Here, g_0 is a fixed (normalized) generating function that is an approximate fixed point of $\mathcal{N}$. Then $\gamma = 0$ is almost a fixed point of $\mathcal{M}$. The linear operator M in this definition is taken to be an approximate inverse of $\mathbb{I} - D\mathcal{N}'(g_0)$, so that the derivative

$$D\mathcal{M}(\gamma) = \mathbb{I} - \bigl[\mathbb{I} - D\mathcal{N}'(g_0 + M\gamma)\bigr]M \tag{5.3}$$

of $\mathcal{M}$ is small near $\gamma = 0$. Since $D\mathcal{N}$ is compact, we choose for $A = \mathbb{I} - M$ a finite rank "matrix", in the sense that $AE_j = \sum_i A_{i,j}E_i$, with $A_{i,j} = 0$ for all but finitely many index pairs (i, j). The indices here are pairs of nonnegative integers, and $E_{m,n}$ is the function defined in (3.10). In addition, we ensure that $NA = 0$, which guarantees that $N\mathcal{M} = 0$.

Given any function $h \in \mathcal{A}^e_\varrho$ and any real number $r > 0$, denote by $B_r(h)$ the closed ball in $\mathcal{A}^e_\varrho$ of radius r, centered at h.

Lemma 5.1. *There exists a normalized real polynomial $g_0 \in \mathcal{A}^e_\varrho$, a bounded linear operator M on $\mathcal{A}_\varrho$ as described above, as well as real numbers $r > 0$ and $R \ge \|M\|_\varrho r$, such* *that the following holds. The transformation $\mathcal{N}$ is well defined, bounded, and analytic, as a map from $B_R(g_0)$ to $\mathcal{A}_\rho$, with $\rho_u = \frac{17}{16}\varrho_u$ and $\rho_v = \frac{17}{16}\varrho_v$. For the corresponding map $\mathcal{M}$,*

$$\|\mathcal{M}(g_0)\|_\varrho \le \varepsilon\,, \qquad \|D\mathcal{M}(\gamma)\|_\varrho \le \kappa\,, \tag{5.4}$$

with $\varepsilon, \kappa > 0$ satisfying $\varepsilon + \kappa r < r$. Here, and in the statement that follows, γ denotes an arbitrary function in $B_r(0)$. The equation (2.8), with $g = g_0 + M\gamma$, determines two locally unique constants λ and μ, and these constants satisfy the bounds (1.2).

Our proof of this lemma is computer-assisted and will be described in in Section 7.

Theorem 4.1 follows as a corollary: By the contraction mapping principle, $\mathcal{M}$ has a unique fixed point γ in the ball $B_r(0)$. The corresponding function $g = g_0 + M\gamma$ belongs to $B_R(g_0)$ and is a fixed point of $\mathcal{N}$. This fixed point is locally unique, as M cannot have an eigenvalue 0, given that $\mathcal{M}$ is a contraction. The compactness of the derivative (near g) follows from the analyticity-improving property of $\mathcal{N}$, since the inclusion map $\mathcal{A}_\rho \to \mathcal{A}_\varrho$ is compact.

6. Implicitly defined quantities

The definition of $\mathcal{N}$ involves a number of implicit equations, such as the equation $g_1(\lambda^2,\lambda) - g_2(\lambda^2,\lambda)^2 = 0$ for λ, or the equation $RX = \mathbf{1}$ for the multiplicative inverse X of a given function R, or equation (2.11) for the midpoint function $\mathcal{V}$. In our computer-assisted proof, implicit equations are always solved by first determining numerically an approximate solution, and then estimating the error. As an example, we discuss the solution of (2.11). The other (simpler) implicit equations are solved similarly.

Let $\boldsymbol{x}(x,y) = x$. Given functions $\phi, \psi, \mathcal{V}$, such that $\psi(\boldsymbol{x}, \mathcal{V}) + \phi(\mathcal{V}, -\mathbb{S}\mathcal{V}) \approx 0$, the goal is to find a function ν such that

$$\mathcal{K}(\mathcal{V}+\nu) \stackrel{\text{def}}{=} \psi(\boldsymbol{x}, \mathcal{V}+\nu) + \phi(\mathcal{V}+\nu, -\mathbb{S}(\mathcal{V}+\nu)) \tag{6.1}$$

is equal to zero. Modulo notation, this is the problem (2.11). The derivative of $\mathcal{K}$ at $\mathcal{V}+\nu$ can be written as

$$\begin{aligned} D\mathcal{K}(\mathcal{V}+\nu)h &= A_\nu h + B_\nu \mathbb{S}h \\ &= D\mathcal{K}(\mathcal{V})h + (A_\nu - A_0)h + (B_\nu - B_0)\mathbb{S}h\,, \end{aligned} \tag{6.2}$$

where

$$\begin{aligned} A_\nu &= \psi_2(\boldsymbol{x}, \mathcal{V}+\nu) + \phi_1(\mathcal{V}+\nu, -\mathbb{S}(\mathcal{V}+\nu))\,, \\ B_\nu &= -\phi_2(\mathcal{V}+\nu, -\mathbb{S}(\mathcal{V}+\nu))\,. \end{aligned} \tag{6.3}$$

A straightforward computation shows that

$$D\mathcal{K}(\mathcal{V})^{-1}h = R^{-1}(\mathbb{S}A_0)h - R^{-1}B_0\mathbb{S}h\,, \tag{6.4}$$

where $R = A_0\mathbb{S}A_0 - B_0\mathbb{S}B_0$ and $R^{-1} = 1/R$.

In the following proposition, $\mathcal{K}$ is considered a map on $\mathcal{A}_\rho$, defined near $\mathcal{V} \in \mathcal{A}_\rho$.

Proposition 6.1. *Let $r > 0$. Assume that the following holds, for all $\nu \in \mathcal{A}_\rho$ of norm r or less. The functions A_ν, B_ν, R^{-1} belong to $\mathcal{A}_\rho$ and satisfy a bound*

$$\|R^{-1}\|_\rho\bigl(\|A_0\|_\rho + \|B_0\|_\rho\bigr)\bigl(\|A_\nu - A_0\|_\rho + \|B_\nu - B_0\|_\rho\bigr) \le \kappa\,, \tag{6.5}$$

with $\kappa < 1$. Furthermore,

$$\bigl\|D\mathcal{K}(\mathcal{V})^{-1}\mathcal{K}(\mathcal{V})\bigr\|_\rho \le \varepsilon < (1-\kappa)r\,. \tag{6.6}$$

Then the equation $\mathcal{K}(\mathcal{V}+\nu) = 0$ has a unique solution $\nu_* \in \mathcal{A}_\rho$ of norm $\le r$.

Proof. Define

$$\mathcal{C}(\nu) = \nu - D\mathcal{K}(\mathcal{V})^{-1}\mathcal{K}(\mathcal{V}+\nu)\,. \tag{6.7}$$

Notice that $\|\mathcal{C}(0)\|_\rho \le \varepsilon$ by the assumption (6.6). We also have

$$\begin{aligned} D\mathcal{C}(\nu) &= \mathrm{I} - D\mathcal{K}(\mathcal{V})^{-1}D\mathcal{K}(\mathcal{V}+\nu) \\ &= -D\mathcal{K}(\mathcal{V})^{-1}\bigl[D\mathcal{K}(\mathcal{V}+\nu) - D\mathcal{K}(\mathcal{V})\bigr] \\ &= -R^{-1}\bigl[(\mathbb{S}A_0) - B_0\mathbb{S}\bigr]\bigl[(A_v - A_0) + (B_v - B_0)\mathbb{S}\bigr]\,, \end{aligned} \tag{6.8}$$

so the inequality (6.5) implies that $\|D\mathcal{C}(\nu)\| \le \kappa$ on the ball $\|\nu\|_\rho \le r$. The assertion now follows from the contraction mapping principle. **QED**

When solving the midpoint equation (2.11), we start with a function g belonging to $\mathcal{A}^{\rm e}_\varrho$, and a function f belonging to another space of this type. In order to verify the hypotheses of Proposition 6.1, we first have to estimate the derivatives $\phi = f_1$ and $\psi = g_2$, as as well as the first partial derivatives of ϕ and ψ. This can be done by using Proposition 3.1. Then we can use Proposition 3.2 to estimate the functions A_ν and B_ν. The inverse $1/R$ is estimated by using the contraction $\xi \mapsto (\mathcal{X}+2\xi)[\mathbf{1} - R\mathcal{X}] - R\xi^2$, whose fixed point ξ_* is the difference between the true inverse and an approximate inverse $\mathcal{X}$. Combining steps and using the chain rule, we see that the function $\mathcal{C}(\nu)$ in (6.7) depends analytically on the pair (f,g), on any open domain where the necessary norm inequalities are satisfied. By uniform convergence, the analytic dependence carries over to the solution ν_*.

The last argument is based on the fact that the fixed point for the contraction $\mathcal{C}$ can be obtained by iteration, yielding a sequence that converges (geometrically or better) in norm. The same arguments apply to the fixed point problem $\mathcal{M}(\gamma) = \gamma$. Our constructive definition of $\mathcal{M}$ yields either an empty domain (unsatisfied norm inequalities), or else an analytic map.

7. Organization of the programs

What remains to be proved is Lemma 5.1, including (1.2), the domain and range properties described in Lemma 4.2, and a simple bound on $DJ(0,0)$. These are all inequalities. The goal is to reduce inequalities like $\|D\mathcal{M}(\gamma)\|_\varrho \le \kappa$ into several simpler ones, and to continue this reduction, until the inequalities that need to be checked are completely trivial. The computer is used not only to check these trivial inequalities, but also to generate them, using the propositions from this paper, or more basic facts, or definitions.

The basic techniques used in our proof are not new. Thus, we will limit our description to the main structure, and to some novel aspects. The precise definitions, and all other details of the proof, can be found in the source code of our programs. (They are written in the programming language Ada95 [17].) The programs should be organized well enough to be readable without much knowledge of programming.

Since more complex structures are defined in terms of simpler ones, we will start with a description of the lowest level. One of the issues at this level is rounding. To avoid a possible misunderstanding, we would like to stress that the control of roundoff errors is a rather trivial aspect of this type of proofs. The main difficulty is to control objects in infinite dimensions, namely our spaces $\mathcal{A}_\rho$, with a finite amount of information. This requires keeping track of what is relevant at each step of the proof, and discarding unnecessary information.

In what follows, words in `this font` will denote entities (data types, procedures, packages, ...) in our programs.

As mentioned earlier, implicit equations are first "solved" numerically, and then we prove that there exists a true solution nearby. Thus, most procedures are designed to be run either in numeric or rigorous mode, depending on whether the generic type `Scalar` is instantiated with `Numeric` or `Ball`, respectively. Modes are switched withing the program as needed. In numeric mode, the floating point unit is instructed to round to the nearest `Rep` (representable number, in our case 80 bit [16]), while in rigorous mode, we put the unit into round-up mode. This guarantees e.g. that `R1+R2` returns an upper bound on the true sum of `R1` and `R2`. A lower bound can be obtained from `-(-R1-R2)`. This allows for rigorous interval arithmetics.

Our "intervals" are in fact special cases of balls $\mathcal{B}(c,r,b) = (c + r\,\mathcal{U})\mathbf{1} + b\,\mathbf{U}$ in a commutative Banach algebra $\mathcal{A}$ with unit $\mathbf{1}$, where c and $r, b \geq 0$ are representable real numbers, $\mathcal{U}$ is the unit ball in $\mathbb{R}$ or $\mathbb{C}$, and $\mathbf{U}$ is the unit ball in $\mathcal{A}$. The corresponding data type `Ball` is a record `S=(S.C,S.R,S.B)` with components of type `Rep`. Using controlled rounding as described above, it is easy to implement an operation `S1+S2` that returns a `Ball S`, with the property that $\mathcal{B}(\texttt{S})$ contains all sums $s_1 + s_2$, with $s_1 \in \mathcal{B}(\texttt{S1})$ and $s_2 \in \mathcal{B}(\texttt{S2})$. Such low level operations are defined in the Ada package `Balls`. In what follows, we identify a data type like `Ball` with the collection of all representable sets $\mathcal{B}(\texttt{S})$ based on this type.

These sets $\mathcal{B}(\texttt{S})$ are sufficient when working with $\mathcal{A} = \mathbb{R}$ or $\mathcal{A} = \mathbb{C}$. Consider now the space $\mathcal{A} = \mathcal{A}_\rho^\diamond$. In this case, $\mathcal{B}(\texttt{S})$ represents a neighborhood of the constant function $(u, v) \mapsto \texttt{S.C}$. More elaborate subsets of $\mathcal{A}_\rho^\diamond$ are represented by a type `Taylor2`, consisting of a pair `T=(T.R,T.C)`, where `T.R` is a pair of numbers of type `Radius` (nonnegative `Rep`), representing the domain parameter ρ, and where `T.C` is a two-dimensional `array(0..PDeg,0..PDeg)` with components `T.C(M,N)` of type `Ball`. The pair `T` represents the set $\mathcal{B}(\texttt{T}) = \sum \mathcal{B}(\texttt{T.C(M,N)})\mathbf{u}^M\mathbf{v}^N$. This sum ranges over nonnegative integers `M` and `N`, with `M+N` not exceeding `PDeg`. Clearly, we can define a `Taylor2`-sum `T1+T2` with the desired property (analogous to the one described above for balls), in terms of `Ball`-sums

T1.C(M,N)+T2.C(M,N). This and other bounds on operations involving functions in $\mathcal{A}_\rho^\diamond$ are defined in the package Taylors2.

For the quadratic functions $\boldsymbol{u}$ and $\boldsymbol{v}$ defined in Section 3, we use a data type Args, given by a record A=(A.T0,A.S0,A.B,A.A) with components of type Rep. The first three components describe the parameters t_0, s_0, and b. The component A.A is the coefficient a in a more general version $a[t^2 - t_0^2] + [s - s_0]$ of our function $\boldsymbol{v}$, but we only use $a = 0$ here. Some basic operations involving such quadratic functions, and changes of variables, are defined in the packages MiniFuns and MiniFuns.Ops.

Our standard sets in the space $\mathcal{A}_\rho$ are defined by a type Fun, consisting of a quadruplet F=(F.A,F.E,F.P,F.Q), where F.P and F.Q are of type Taylor2. The component F.E is a Boolean parameter; if True, then the set represented by F is $\mathcal{B}(\texttt{F}) = \mathcal{B}(\texttt{F.P}) + \boldsymbol{t}\mathcal{B}(\texttt{F.Q})$, with $\mathcal{B}(\texttt{F.P})$ and $\mathcal{B}(\texttt{F.Q})$ as described above, except for a change of variables $u = \boldsymbol{u}(x, y)$ and $v = \boldsymbol{v}(x, y)$ defined by the Args-type argument F.A of F. We refer to such a Fun as being of "even" type. If F.E is False, then the Ball-type components F.P.C(M,N) and F.Q.C(M,N) use $\mathcal{A}_\rho$ as algebra, and not just the even subspace $\mathcal{A}_\rho^e$. So the functions in the corresponding sets $\mathcal{B}(\texttt{F.P})$ and $\mathcal{B}(\texttt{F.Q})$ need not be even.

This "general" version of Fun appears naturally when composing with a function in $\mathcal{A}_\rho$. For such compositions, we use Proposition 3.2 to estimate the errors. The even type is more convenient for estimating derivatives, since we can use Proposition 3.1 directly, via the chain rule. Thus, once the midpoint equation (2.11) is solved, and we have a "general" Fun for the function $\widehat{h}$ defined in (2.17), we convert this set to even type. The basic operations involving sets in $\mathcal{A}_\rho$ are defined in the package Funs2.

We recall that (2.11) is solved by first computing a numerical approximation for the functions $\mathcal{V}$. This is done by the procedure Funs2.Num.NumCompZero. Then RG.MidPoint verifies that this approximate solution satisfies the hypotheses of Proposition 6.1. This yields an upper bound r on the norm of the error, so it suffices to add a ball of radius r to the approximate solution, to obtain a Fun-type set that contains the true midpoint function $\mathcal{V}$.

The above discussion should make clear that we can constructively define a map Renorm, from Fun to Fun $\cup$ {Error}, with the following property: If $g \in \mathcal{B}(\texttt{G1})$, and if Renorm yields a set G2, then $\mathcal{N}(g) \in \mathcal{B}(\texttt{G2})$. In the context of computer-assisted proofs, such a set-map is called a "bound" on the map $\mathcal{N}$. Bounds on maps like $\mathcal{N}$ and $\mathcal{M}$ are defined in the package RG. They use bounds on more basic maps, defined in Funs2, which in turn use bounds defined in Taylors2, etc. If a domain Error occurs along the way (meaning that some condition could not be verified), then the program is simply halted.

Our bound on $\mathcal{M}$ is named Contract. One of the steps in the proof of Lemma 5.1 is to verify that $\|\mathcal{M}(0)\|_\varrho \le \varepsilon$. This is done simply by applying Contract to the set $\{0\}$, and then evaluating the Norm of the resulting set of functions, which yields a set of numbers named Eps. The maximum Sup(Eps) defines our choice of ε in Lemma 5.1. Then r is determined in such a way that $\varepsilon + \kappa r < r$ holds if κ less than KMax $= \frac{35}{64}$.

We note that, even though $\mathcal{M}$ is a contraction, Contract will not map any set from Fun into itself. The reason is that these sets do not carry enough information to exhibit the cancellations that are responsible for the contraction property of $\mathcal{M}$. But the cancellations do occur when our bound DContract on $D\mathcal{M}$ is applied to a basis vector E_j. This allows us to estimate the operator norm of $D\mathcal{M}(\mathcal{G})$, using the formula (3.10). Here, $\mathcal{G}$ is the ball $B_r(0)$ described in Lemma 5.1. The norm $\|D\mathcal{M}(\mathcal{G})E_j\|_\rho$ is estimated explicitly, for finitely many indices $j=(m,n)$, including all those that have $A_{i,j}\neq 0$ for some i. The remaining basis vectors E_j , which correspond to larger values of $|m+n|$, are contracted so strongly by $D\mathcal{M}(\mathcal{G})$ that they can be mapped collectively, in a small number of sets of type `Fun`. As a result, we obtain the desired bound on the norm of $D\mathcal{M}(\mathcal{G})$ in 70 steps of `DContract`, organized by the function `DContractNorm`. The bound is less than `KMax`.

At this point, we have a ball $B_R(g_0)$, described by a record `G` of even type `Fun`, that contains a fixed point of $\mathcal{N}$. The proof of Lemma 5.1 is completed by executing `RG.LambdaMu(G,La,Mu)`, which returns two balls `La` and `Mu` containing the scaling constant λ and μ, respectively, for all functions in $B_R(g_0)$. In particular, λ is estimated by solving the fixed point problem for the map $\lambda\mapsto K(g,\lambda)$ defined by equation (2.13). The bound (1.2) holds for all real values in the balls `La` and `Mu`.

Our proof of Lemma 4.2 is comparatively low-tech. For sets of points in $\mathbb{C}^2$ we use a data type `Point`, which has two components of type `Ball` (with $\mathcal{A}=\mathbb{C}$, but recall that the center of a `Ball` is real). A claim of the type $G_0R_1\subset R_2$ is checked simply by solving equation (1.6) for all points $(x,z)\in R_1$, and then checking that $(y,w)=G_0(x,z)$ belongs to R_2 . In particular, the function `RG.GLambda0` takes a `Point P` as an argument, and returns a `Point Q` that contains $q=G_0\Lambda p$ for all p in the set $\mathcal{B}$(`P`) defined by `P`. Thus, in order to prove that $(G_0\Lambda)\Lambda R_+\subset R_-$, we simply cover ΛR_+ with a finite number of such sets $\mathcal{B}$(`P`) and verify that the sets $\mathcal{B}$(`Q`) returned by `RG.GLambda0` are all contained in $\Lambda^{-1}R_-$. Again, this is done simultaneously for all functions $g\in B_R(g_0)$. The other domain conditions are verified analogously. This task is coordinated by `RG.CheckMaps`. Before that, `CheckJ(G)` verifies that the 2×2 matrix $DJ(0,0)$ is different from $-$I. This is done by multiplying the derivatives of G_0, F_0, and H_0^{-1}, evaluated at the appropriate points. These derivatives can all be expressed in terms of second derivatives of the generating function g. Derivatives are always estimated by using Proposition 3.1, even if only particular values are needed.

For further details, the reader is referred to the source code of these programs [15]. When the program `Verify` is compiled (by an Ada compiler) and then run, the above-mentioned steps are carried out, and the resulting numerical inequalities are verified. This process takes about 28 hours on a current personal computer. The values of the parameters described in Lemma 5.1 are roughly $\varepsilon\approx 1.9*10^{-13}$ and and $R\approx 8.3*10^{-12}$.

References

[1] L.P. Kadanoff, *Scaling for a critical Kolmogorov–Arnold–Moser trajectory.* Phys. Rev. Lett. **47**, 1641–1643 (1981).

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

[3] R.S. MacKay, *Renormalisation Approach to Invariant Circles in Area–Preserving Maps,* Physica D **7**, 283–300 (1983).

[4] J.M. Greene, J-M. Mao, *Higher-order fixed points of the renormalisation operator for invariant circles.* Nonlinearity **3**, 69–78 (1990).

[5] J. Wilbrink, *New fixed point of the renormalisation operator associated with the recurrence of invariant circles in generic Hamiltonian maps.* Nonlinearity **3**, 567–584 (1990).

[6] A. Stirnemann, *Renormalization for Golden Circles.* Comm. Math. Phys. **152**, 369–431 (1993).

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

[8] D.F. Escande, F. Doveil, *Renormalisation Method for Computing the Threshold of the Large Scale Stochastic Instability in Two Degree of Freedom Hamiltonian Systems.* J. Stat. Phys. **26**, 257–284 (1981).

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

[10] C. Chandre and H.R. Jauslin, *Renormalization–group analysis for the transition to chaos in Hamiltonian systems*, Physics Reports **365**, 1–64, (2002).

[11] H. Koch, *Renormalization of vector fields.* In: Holomorphic Dynamics and Renormalization, M. Lyubich and M. Yampolsky (eds.), Fields Institute Communications, AMS (2008) 269–330.

[12] H. Koch, *On the renormalization of Hamiltonian flows, and critical invariant tori.* Discrete Contin. Dynam. Systems A, **8**, 633–646 (2002).

[13] H. Koch, *A Renormalization Group Fixed Point Associated with the Breakup of Golden Invariant Tori.* Discrete Contin. Dynam. Systems **11**, 881–909 (2004).

[14] H. Koch, *Existence of Critical Invariant Tori*, Erg. Theor. Dyn. Syst., **28**, 1879–1894 (2008).

[15] G. Arioli, H. Koch, *The critical renormalization fixed point for commuting pairs of area-preserving maps.* The preprint `mp_arc 09-67` includes the programs.
See also `ftp://ftp.ma.utexas.edu/pub/papers/koch/maps-fixpt/`

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

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