On the renormalization of systems with shearless tori

Hans Koch [1]

Abstract. Renormalization has been successful in explaining universal scaling phenomena in systems near criticality. In the best understood cases, some nondegeneracy property is involved that propagates under "coarse-graining". This includes ferromagnetism in statistical mechanic, the negative Schwarzian derivative for maps on an interval, or the twist condition for maps on a cylinder. But similar phenomena have been observed during the breakup of shearless invariant tori in the nontwist standard map. This suggests that renormalization arguments should apply here as well. Our goal was to investigate this question by numerically implementing a renormalization transformation. This turned out to be more challenging than expected. We describe some partial result, problems that were encountered, and a new approach to the renormalization of Hamiltonians. Similar ideas may be useful in other renormalization-related problems. We are not assuming that the reader is familiar with the breakup of invariant tori or renormalization. The necessary concepts will be introduced as needed.

1. Introduction

Critical phenomena in low-dimensional dynamical systems typically involve sequences of bifurcations that accumulate, marking a transition to a new type of asymptotic behavior. This involves asymptotic scaling, both in phase space and parameter space. Another common observation is universality: The asymptotic scaling seems independent of the system being considered, within a large "universality class" of systems. The method of renormalization aims to explain such universality.

Some degree of universality seems natural: cascades of bifurcations of a given type are special enough that the number of ways to produce them should be limited. A renormalization-based argument is roughly the following.

In the case of parameterized families of maps, the construction of long orbits of a given type involves applying a transformation $\mathfrak{R}$ to a map. $\mathfrak{R}$ involves compositions and can be viewed as "coarse-graining" in time. What one ends up observing for such orbits, in a fixed family of maps, depends partly on the family, and partly on the transformation $\mathfrak{R}$. The universal properties are those that can be attributed to $\mathfrak{R}$ and are independent of the family.

For maps with constraints, $\mathfrak{R}$ may act on generating functions. For flows, $\mathfrak{R}$ may act on vector fields, or on Hamiltonians in the case of Hamiltonian flows. For simplicity, we restrict to a real analytic setting in all cases.

A classic example of such a "renormalization transformation" $\mathfrak{R}$ will be described at the beginning of Section 2. The main ideas are simple, and they can be applied to other settings by analogy. No prior knowledge of renormalization is necessary to follow the arguments presented in this paper.

We started by considering symplectic maps on the cylinder $\mathbb{T} \times \mathbb{R}$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Near an invariant circle with irrational rotation number α, such a map G is conjugate to a small perturbation of $\Psi : \left[\begin{smallmatrix} x \\ w \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} x+\psi(w) \\ w \end{smallmatrix}\right]$ that agrees with Ψ on the invariant circle $w = \bar{w}$ with rotation number $\psi(\bar{w}) = \alpha$. The corresponding circle for G is said to have twist if

[1] Department of Mathematics, The University of Texas at Austin, Austin, TX 78712.

$\psi'(\bar w) \neq 0$. At the opposite extreme, if ψ has a local minimum or maximum at $\bar w$, then the circle is said to be shearless.

Twist represents a monotonicity or nondegeneracy property. Such properties seem conducive to critical phenomena. But they may not always be necessary. The observations that prompted the work described here are numerical studies on the breakup of shearless invariant tori for the standard nontwist map. We refer to [9,14] for a description and references. The data suggest that renormalization should apply here as well.

A technical problem in the shearless case is that common types of generating functions cannot be used. And less common types interfere with reversibility. Another problem is that the asymptotic behavior is quite complex: The relevant orbit of $\mathfrak{R}$ seems to be a period of length 12, as opposed to a fixed point in the twist case. This makes numerical investigations difficult and slow. We were aware of this risk but found it worth a try.

Our first goal was to find a suitable transformation $\mathfrak{R}$ that can be studied numerically. We started with symplectic maps, since the Hamiltonian approach is much more involved. The next goal was to find a "critical" approximate orbit for $\mathfrak{R}$. Subsequent steps would depend on our findings.

In parallel to this, we also computed long shearless obits for the standard nontwist map at high precision. The hope was that accurate numerics will help in finding a convenient version of $\mathfrak{R}$ for maps. This hope did not quite materialize. But the data revealed the existence of rotation in the asymptotic scaling, about a point known as the indicator point. This observation seems to be new. A possible source of rotating asymptotics could be infinitely nested meanders at the point of breakup.

Our maps-based approach is described in Section 2. The chosen transformation $\mathfrak{R}$ is of a standard type for problems that involve irrational rotations. As a specific map in the twist (shearless) case we discuss the (nontwist) standard map. In the chosen coordinates, both maps G are reversible with respect to the involution $S_1 = \mathrm{diag}(-1, 1)$. That is, $S_1 G S_1 = G^{-1}$. Unfortunately, trying to take advantage of reversibility causes problems in the shearless case. After having considered various alternatives, we reluctantly decided to renormalize Hamiltonians instead.

The Hamiltonians considered here are functions on $\Omega = \mathbb{T}^2 \times \mathbb{R}^2$. They define a Hamiltonian flow on Ω in the usual way. Here, reversibility H with respect to S_1 is a simple symmetry: $H \circ S_1 = -H$. Restricting the flow to a surface of fixed energy, and then considering the first-return map to a suitable transversal section, yields a symplectic map of the type described above. So the notion of twist and shearlessness are relevant in Hamiltonian systems as well. For a discussion of shearless KAM tori we refer to [22].

Our renormalization of Hamiltonians $H : \Omega \to \mathbb{R}$ involves a re-normalization group that is infinite dimensional. This makes it technically much more complex. The approach is described in Section 3. Our findings are based on extensive numerical experiments. We started with the transformation $\mathfrak{R}$ that was introduced in [8] and used successfully in a computer-assisted proof [15,23]. But it is Newton-based and thus involves inverting a linear operator at each step. In numerical experiments, the inversion fails quite frequently. This was "solved" in [13] by skipped re-normalization steps that caused problems. Here, we avoid the invertibility problem by using a procedure that normalizes H via a flow $\dot H = \{\phi, H\}$ for a time-dependent Hamiltonian ϕ that depends on H.

A temporary highlight of our analysis was the construction of an approximate period 12 for this version of $\mathfrak{R}$. But we also found an undesired eigenvalue. This led us to implement some improvements to $\mathfrak{R}$. After that, the period 12 vanished and turned into near-period 12 that involved an extra rotation. This came as a surprise. But this rotation may be related to the above-mentioned rotation that we observed in the asymptotic scaling of the nontwist standard map near the indicator point.

2. Renormalization of maps

In Subsection 2.1 we will introduce a renormalization transformation $\mathfrak{R}$ for (pairs of) symplectic maps on $\mathbb{T} \times \mathbb{R}$. Subsection 2.2 then specializes to shearless maps. But first, let us describe how renormalization relates to universality.

Among the best known universal phenomena in dynamics is the period-doubling cascade in one-parameter families of quadratic-like maps on an interval [4,19]. To simplify the description, consider the family of maps $g_s : [-1, 1] \to [-1, 1]$ for $0 < s < 1$, defined by $g_s(x) = 1 - sx^2$. The observation is that there exists an infinite sequence of parameter values $0 < s_0 < s_1 < s_2 < \ldots < 1$ with the following property. As s is increased past s_n, a stable periodic orbit of length 2^n turns unstable but gives rise to a nearby stable period 2^{n+1}. The bifurcations accumulate with an asymptotic ratio $\delta = \lim_n \left(\frac{s_{n-1}-s_n}{s_n-s_{n+1}}\right)^{1/n}$. The same is observed for nearby families; and remarkably, the asymptotic ratio is always the same: $\delta = 4.669\ldots$.

In this case, universality is due to the (nontrivial) fact that the composition operator $g \mapsto g^2$ for equivalence classes of even functions has a real analytic fixed point. To be more precise, a representative g_* in this class is a fixed point of a transformation $\mathfrak{R}$ of the form $\mathfrak{R}(g) = \Lambda_g^{-1} g^2 \Lambda_g$. In fact, it is possible to choose the change of coordinates Λ_g linear, with $\Lambda_g(x) = g(1)x$. Then $\mathfrak{R}$ can be proved to be hyperbolic near g_*, in a space $\mathcal{A}$ of even real-analytic functions; with $DG(g_*)$ having no spectrum outside the open unit disk, except for simple eigenvalue $\delta = 4.669\ldots$.

The universality class in this case includes all one-parameter families $s \mapsto g_s$ of maps in $\mathcal{A}$ that cross the stable manifold of $\mathfrak{R}$ for g_* transversally. Say the crossing point is at $s = s_\infty$. Then, under iteration by $\mathfrak{R}$ and proper parameter rescaling, about s_∞ with asymptotic ratio δ, the family converges to a universal family, namely the unstable manifold for $\mathfrak{R}$ at g_*. As far as universality is concerned, this explains the relevant phenomena. A more detailed description can be found in [19] and many other reviews.

Observe that the transformation $\mathfrak{R}$ consists of two parts: a composition $g \mapsto g^2$, followed by a re-normalization via a change of coordinates Λ_g. The composition part represents the arithmetic map $n \mapsto 2n$ on the set of positive integers.

The transformation $\mathfrak{R}$ defined below admits an analogous decomposition. The composition part represents the Gauss map $\alpha \mapsto \alpha^{-1} - \lfloor \alpha^{-1} \rfloor$, where $\lfloor \alpha^{-1} \rfloor$ denotes the integer part of α^{-1}. In this context, α is a rotation number in the interval $(0, 1)$.

2.1. Rotation numbers

For concreteness, let us first consider the well-known Chirikov standard map. Modulo a trivial conjugacy, the standard map is $G = \Phi\Psi\Phi$, where Ψ and Φ are given by

$$\begin{bmatrix} x \\ w \end{bmatrix} \overset{\Psi}{\longmapsto} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} x + \psi(w) \\ w \end{bmatrix}, \qquad \begin{bmatrix} x \\ w \end{bmatrix} \overset{\Phi}{\longmapsto} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} x \\ w - \varphi(x) \end{bmatrix}, \tag{2.1}$$

with $\psi(w) = w$ and $\varphi(x) = \frac{b}{2}\sin(2\pi x)$.

In the trivial case $b = 0$, where $G = \Psi$, the cylinder $\mathbb{T} \times \mathbb{R}$ is foliated by invariant circles. On the circle $\mathbb{T} \times \{w\}$ we have a rotation $y = x + w$ with rotation number $\alpha = w$. And G satisfies the so-called twist condition $\frac{\partial y}{\partial w} \neq 0$. By standard results in KAM theory, most of these invariant circles persist for values $b \neq 0$ close to zero, with the closeness condition depending on the arithmetic properties of α.

For the purpose of renormalization, a map G on $\mathbb{T} \times \mathbb{R}$ will be regarded as a map on $\mathbb{R} \times \mathbb{R}$ that commutes with the translation $F_\circ$ defined below. For reference later on, we note that $F_\circ = E^2$, where

$$\begin{bmatrix} x \\ w \end{bmatrix} \overset{E}{\longmapsto} \begin{bmatrix} x - 1/2 \\ -w \end{bmatrix}, \qquad \begin{bmatrix} x \\ w \end{bmatrix} \overset{F_\circ}{\longmapsto} \begin{bmatrix} x - 1 \\ w \end{bmatrix}. \tag{2.2}$$

The objects of interest here are periodic and quasiperiodic orbits of G. A point $\zeta \in \mathbb{R}^2$ is periodic with rotation number p/q, if $F_\circ^p G^q \zeta = \zeta$. For an irrational number α between 0 and 1, if $p_n/q_n \to \alpha$ and $F_\circ^{p_n} G^{q_n} \zeta \to \zeta$, then ζ is said to have rotation number α. These definitions extend readily to other pairs of commuting maps (F, G).

The composed maps $F^{p_n} G^{q_n}$ play a role similar to the composed maps g^{2^n} in the period-doubling scenario. Scaled versions of these maps can be generated by starting with a commuting pair $P = (F, G)$ and then iterating the following transformation $\mathfrak{R}$.

$$\mathfrak{R}(P) = \tilde{P}, \qquad \tilde{P} = \left(\Lambda_P^{-1} \hat{F} \Lambda_P \,,\, \Lambda_P^{-1} \hat{G} \Lambda_P\right), \quad \hat{F} = G, \quad \hat{G} = F G^c, \tag{2.3}$$

where $c = \lfloor \alpha^{-1} \rfloor$. Here $\Lambda_P(x, w) = (-\alpha x, \mathfrak{a}(w))$, where $\mathfrak{a}$ is some affine function that can be chosen to depend on P, say via some normalization condition. We note that, if ζ has rotation number α for P, then $\tilde{\zeta} = \Lambda_P^{-1}\zeta$ has rotation number $\tilde{\alpha} = \alpha^{-1} - c$ for $\tilde{P} = \mathfrak{R}(P)$. Here, $\alpha \mapsto \tilde{\alpha}$ is the Gauss map.

For simplicity, we restrict now to the inverse golden mean α, defined as the solution of $\alpha^{-1} - 1 = \alpha$ in $(0, 1)$. Then we can use $c = 1$ in the definition (2.3) at each step of $\mathfrak{R}$. In this case, the numerator p_n in the n-th continued fractions approximant p_n/q_n of α is the n-th Fibonacci number, and $q_n = p_{n+1}$. For completeness, we note that the pair $(F_\circ, G_\circ)$ with $G_\circ : \left[\begin{smallmatrix} x \\ w \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} x+\alpha \\ w \end{smallmatrix}\right]$ is a fixed point of $\mathfrak{R}$, with $\mathfrak{a}(w) = \mu\omega$ and arbitrary $\mu \neq 0$.

As mentioned above, if b is sufficiently close to 0, then the standard map $G = \Phi\Psi\Phi$ given by (2.1) has a smooth invariant circle with rotation number α. As has been described in many studies, this circle breaks up as b passes some critical value. This happens via a cascade of bifurcations of p_n/q_n orbits; but the details are not really relevant here. The conjecture was that this breakup is universal and governed by a fixed point P_* of the transformation $\mathfrak{R}$. For a numerical study and other supporting evidence we refer to [2].

A proof for the existence of this fixed point was given in [21], using computer-assisted methods. The proof also yields highly accurate rigorous bounds on the critical scalings. Numerical computations carried out in connection with [23] also yield values for other critical exponents, such as the expanding eigenvalue of $D\mathfrak{R}(P_*)$. They are consistent with other known data, leaving no doubt that renormalization fulfills its promise in the shearless case.

Remark 1. Our goal here was to start a similar renormalization analysis in the shearless case, if only numerically. Based on numerical findings in [13,14], we expected that $\mathfrak{R}$ has a periodic orbit of length 6 or 12.

2.2. Shearless maps

We note that twist and shearless invariant tori can appear in the same dynamical system. Shearless tori play an important role in plasma physics [9]. They appear to be extremely stable, sometimes separating two seemingly chaotic regions in phase space.

A shearless analogue of Chirikov's standard map is the standard nontwist map. Up to a trivial conjugacy, this map is again of the form $G = \Phi\Psi\Phi$, with Ψ and Φ as in (2.1). But instead of $\psi(w) = w$ we now have $\psi(w) = a - aw^2$, where a is an extra parameter. The twist of $G : \left[\begin{smallmatrix} x \\ w \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} y \\ z \end{smallmatrix}\right]$ is given by $\frac{\partial y}{\partial w} = 2a(\varphi(x) - w)$, so the twist condition $\frac{\partial y}{\partial w} \neq 0$ is violated along the curve $w = \varphi(x)$.

Both the standard twist map and the standard nontwist map are reversible with respect to the involution $S_1 = \mathrm{diag}(-1, 1)$, meaning that $S_1 G S_1 = G^{-1}$, This can be seen easily from (2.1): Both Ψ and Φ are reversible, so the product $G = \Phi\Psi\Phi$, being palindromic, is reversible as well.

Reversibility has played an important role in the renormalization analysis of various types of dynamical systems. A nontrivial task at the beginning of such an analysis is to find a suitable orbit under $\mathfrak{R}$, say a fixed point or a periodic orbit. If there is evidence that the phenomenon under investigation is exhibited in families of reversible maps, then the restriction to reversible maps greatly simplifies the problem of finding an orbit for $\mathfrak{R}$. In this case, the scaling Λ_P in (2.3) should be chosen to preserve reversibility.

A distinguishing feature of the standard nontwist map $G = \Phi\Psi\Phi$ is that the function $\psi(x) = a - ax^2$ that appears in (2.1) is even. As a result, Ψ commutes with the map E defined in (2.2). Since Φ commutes with E as well, so does G. So the standard nontwist map is reversible not only with respect to S_1, but with respect to $S_1 E$ as well. This second type of reversibility is qualitatively different: S_1 is orientation-reversing and has a fixed line ($x = 0$), while $S_1 E$ is orientation-preserving and only has a single fixed point. The fixed point $\zeta_\circ = \left[\begin{smallmatrix} 1/4 \\ 0 \end{smallmatrix}\right]$ of $S_1 E$ is known as the "indicator point" for G.

The extra parameter a plays an important role for the standard nontwist map. In fact, the breakup of shearless golden tori is a codimension 2 phenomenon, meaning that it is generic only in two-parameter families. A detailed bifurcation analysis is carried out in [6,14]. In a nutshell, they determine periodic orbits with rotation number p_k/q_k that start on a symmetry line like $x = 0$, for a close to the expected bifurcation point. Typically, there are two values of b, or none, where such an orbit exists. If there exists two, then varying (a, b) leads to a bifurcation point, where the two merge into a single orbit. This is

the "shearless" orbit. After the bifurcation, the orbit disappears. The resulting sequence $k \mapsto (a_k, b_k)$ of bifurcation points is observed to accumulate at a critical point (a_∞, b_∞). A transformation like (2.3) is discussed in [7,12,14] as a possible renormalization scheme. Assuming that the standard nontwist family accumulates at a two-dimensional invariant manifold under proper rescaling, the observed sequence of residues at the above-mentioned obits suggest that $\mathfrak{R}$ has a periodic orbit of length 12.

Approximate values of various "universal" scaling quantities, based on measurements on specific families (typically the standard nontwist family) are given in several papers, including [6]. They tend to be quite imprecise. One reason is that it takes at least 12 Fibonacci-magnification steps (one step is the analogue of $\mathfrak{R}$) between two measurements before they are expected to compare. Given how quickly the sequence $k \mapsto p_{12k}$ grows, more than one or two comparisons are not practical. And $k = 2$ can hardly be considered near-asymptotic. In [14] they find asymptotic scaling about a S_1-symmetry point $(0, w_*)$ with $w_* = -0.222621076\ldots$. The comparison of scaled orbits in Figure 2.26 of [14] makes it look very likely that the critical shearless circle is asymptotically self-similar about this point. The x-scaling factor per magnification step seems to be close to α, and the w-scaling factor is found to be near 0.603.

The data given in [13] are inaccurate as well. But the scope of [13] includes universality: The numerical data describe a map $\mathfrak{R}$ on an infinite-dimensional space, and not just orbits of a single two-parameter family.

To improve some data for the standard nontwist family, we considered computing shearless p_n/q_n orbits without having to follow lengthy sequences of bifurcations. Here, we used that, if a shearless orbit is unique, then it has to pass thorough the indicator point $\zeta_\circ = \left[\begin{smallmatrix}1/4\\0\end{smallmatrix}\right]$. See also [10]. For large n, a p_n/q_n orbit passing through $\zeta_\circ$ represents an approximation of the shearless golden circle. So we computed these orbits, for n ranging from 24 to 48, in increments of 3. To find the parameter values (a_n, b_n) we started with a guess (a, b). Then a Nelder-Mead algorithm was used to adjust (a, b) until the error $\|F^{p_n} G^{q_n} \zeta - \zeta\|$ was around 10^{-25}. For $n = 48$ this took several month; but computers are patient. After that, a Newton method was used to reduce the error to 10^{-150} or less. All this was done at high precision, with up to 2^{11} binary digits. Based on these computations, the expected accumulation values are $a_\infty = 0.686049108000003860\ldots$ and $b_\infty = 0.742493549155193278\ldots$

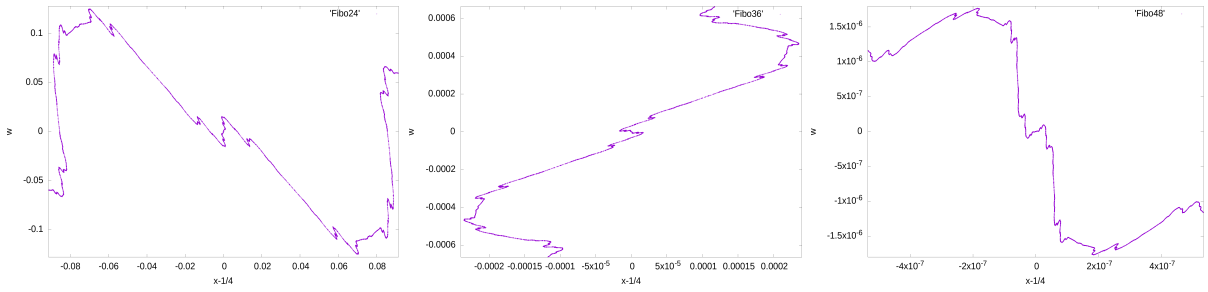

Figure 1. shearless orbits through $\zeta_\circ$ for rotation numbers $\frac{p_{24}}{q_{24}}$, $\frac{p_{36}}{q_{36}}$, and $\frac{p_{48}}{q_{48}}$.

Figure 1 shows parts of three p_n/q_n orbits through the indicator point $\zeta_\circ = \left(\frac{1}{4}, 0\right)$,

computed for $n = 24, 36, 48$. Clearly the same general pattern repeats, but the scaling looks much less simple than what we see near $(0, w_*)$. In particular, self-similar scaling repeats at best after 24 steps. We determined a scaling matrix for three repeating patters via least squares fit. The average dilation is roughly 0.609 per step. But there is also a rotation involved: The amount of rotation over 12 steps seems to be roughly 0.386π modulo π.

This behavior is not what we expected, based on the approximate period 6 that was found in [13]. Some additional evidence for a rotation is described in Subsection 3.5. Our sequence of parameter $n \mapsto (a_n, b_n)$ shows no discernible pattern, except that it seems to converge; presumably to the same limit (a_∞, b_∞) as the (different) sequence found in [14] via bifurcations and orbit re-connections. This raises the question of how the behavior near the indicator point contributes to the breakup of shearless golden circles. One possibility might be an infinite nesting of meanders. Such a nesting will involve a change in angle between one level and the next.

2.3. Renormalization of shearless maps

The standard way of implementing renormalization for (exact) symplectic maps is by working with generating functions. This takes care of the nonlinear constraint $\det(DG) = 1$. Still, to get an idea of what to expect, we started with a brute force approach, using maps on $\mathbb{R}^2$ whose components are represented as independent Taylor series in two variables. The idea was to iterate the transformation (2.3) a few times, starting with the standard nontwist pair $(F_\circ, G)$. This attempt failed rather badly, despite the use of high Taylor degrees and high precision arithmetic. The explicit translations by -1 and α are unmanageably large. A milder version of this problem appears when working with generating functions. In this case, it helps to implement $\mathfrak{R}^3$ directly.

So it was back to generating functions. For twist maps $G\left[\begin{smallmatrix} x \\ w \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} y \\ z \end{smallmatrix}\right]$, the generating function $g = g(x, y)$ has been used successfully in [3,15]. But for shearless maps such as the standard nontwist map, only $g = g(x, z)$ or $g = g(w, y)$ are potentially usable. We have also considered using the so-called primitive function $p = p(x, w)$, but without success. The issue with using one of $g = g(x, z)$ or $g = g(w, y)$ for G is that G^{-1} involves the other. So reversibility is a nontrivial condition, and not just a symmetry of the generating function. One can also try to use both $g = g(x, z)$ and $g = g(w, y)$, with one determined from the other by reversibility. But then there is no guarantee that, after renormalization, both describe the same map. Other seemingly useful ideas turned out to have serious issues as well.

Reluctantly, we abandoned for now the transformation (2.3) for shearless maps. What made this decision particularly hard was our impression that the reversibility of the standard nontwist map G with respect to $S_1 E$ should play an important role. After a translation by $-\zeta_\circ$ that puts the indicator point at the origin, the standard nontwist map is still of the form (2.1), but with $\varphi(x) = \frac{b}{2}\cos(2\pi x)$ in place of $\varphi(x) = \frac{b}{2}\sin(2\pi x)$. Furthermore, the translated standard nontwist map is reversible with respect to $-\mathrm{I}$. So any linear scaling preserves reversibility, including rotations. Unfortunately, we did not find a good way to exploit reversibility with respect to $-\mathrm{I}$.

3. Renormalizing Hamiltonians

An alternative to renormalizing symplectic maps is to renormalize Hamiltonians. Formally, the two procedures can be related [11]. In any case, Hamiltonian flows yield symplectic first-return maps to properly chosen sections.

Consider a flow on $\Omega = \mathbb{T}^2 \times \mathbb{R}^2$ generated by a Hamiltonian $H : \Omega \to \mathbb{R}$. Using angular variables $q \in \mathbb{T}^2$, and conjugate action variables $p \in \mathbb{R}^2$ defined near the origin, the flow generated by H is given by $\dot q = \nabla_p H$ and $\dot p = -\nabla_q H$. A function G on Ω evolves via the Poisson bracket

$$\dot G = \{G, H\} \stackrel{\text{def}}{=} (\nabla_q G) \cdot \nabla_p H - (\nabla_p G) \cdot \nabla_p H \,. \tag{3.1}$$

Recall that the renormalization transformation (2.3) for pairs consists of two parts: a composition $P \mapsto \hat P$, and a re-normalization $\hat P \mapsto \tilde P$. The powers in $\hat F = F^0 G^1$ and $\hat G = F^1 G^c$ are precisely the entries of the matrix T given below. In abbreviated notation, the composition part is $\hat P = P^T$. For Hamiltonians, the composition part is of the form

$$\hat H = H \circ \mathcal{T} \,, \qquad \mathcal{T}(q,p) = \big(Tq, T^{-1}p\big) \,, \qquad T = \begin{bmatrix} 0 & 1 \\ 1 & c \end{bmatrix} . \tag{3.2}$$

So the Hamiltonian analogue of the fixed point equation $P^T \simeq P$ is the fixed point equation $H \circ \mathcal{T} \simeq H$. Both equations seem "natural" and of general interest.

To make the equation $H \circ \mathcal{T} \simeq H$ more concrete, we need to consider dynamic equivalences between Hamiltonians. One of them is a momentum scaling Σ_μ defined by

$$(\Sigma_\mu H)(q,p) = \mu^{-1} H(q, \mu p) \,, \qquad \mu \neq 0 \,. \tag{3.3}$$

It preserves the Poisson bracket (3.1) in the sense that $\{\Sigma_\mu G, \Sigma_\mu H\} = \Sigma_\mu \{G, H\}$. A much larger class of equivalences are changes of variables $H \mapsto H \circ \mathcal{U}$ with $\mathcal{U}$ symplectic and homotopic to the identity. They preserve the Poisson bracket as well. A third type of equivalence is a time scaling $H \mapsto \tau H$, which "only" preserves orbits. Combining the composition $H \mapsto \hat H$ with a re-normalization $\hat H \to \tilde H$ yields a transformation

$$\mathfrak{R}(H) = \tilde H \,, \quad \tilde H = H' \circ \mathcal{U}_{H'} \,, \qquad H' = \tau \Sigma_\mu \hat H \,, \quad \hat H = H \circ \mathcal{T} \,. \tag{3.4}$$

Here τ and μ can depend on H. Our choices will be described below.

For completeness, let us mention that $\mathfrak{R}$ admits a family of trivial orbits. They can be used to prove KAM-type theorems via renormalization [8,17]. In a perturbative setting, $\mathfrak{R}$ can be extended from Hamiltonians to more general vector fields [18,20].

In what follows, we restrict to $c = 1$ in (3.2). This singles out the inverse golden mean α as rotation number. The eigenvalues of T are α^{-1} and $-\alpha$, with eigenvectors ω and Ω, respectively, where

$$T = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} , \qquad \omega = \begin{bmatrix} \alpha \\ 1 \end{bmatrix} , \qquad \Omega = \begin{bmatrix} -1 \\ \alpha \end{bmatrix} . \tag{3.5}$$

To further simplify the analysis, we restrict to Hamiltonians of the form

$$H(q,p) = \omega \cdot p + h\big(q, \Omega \cdot p\big)\,, \quad h(q,z) = \sum_{\nu,k} h_{\nu,k} \cos(\nu \cdot q) z^k\,, \tag{3.6}$$

where the sum ranges over frequencies $\nu \in \mathbb{Z}^2$ and powers $k \in \mathbb{N}$. In fact, due to parity, we can restrict to $\nu_2 \geq 0$. This ansatz was used already in [5]. Pick $\omega' \in \mathbb{R}^2$ parallel to ω, satisfying $\omega' \cdot \omega = 1$. Then the variable $\omega' \cdot q$ is in essence time, since it evolves linearly with velocity 1.

The choice of having the term $\omega \cdot p$ in (3.6) with a fixed coefficient 1 is part of a normalization condition. For Hamiltonians of the form (3.6), this normalization is preserved under $\mathfrak{R}$ if we choose $\tau = \alpha^{-1}$. So this choice will be fixed from now on. In the shearless case, we require that h include a nonzero term $h_{0,3} z^3$. Then the scaling μ in the definition (3.4) is determined by the normalization condition $h_{0,3} = \frac{1}{2}$.

This leaves us with a choice for the canonical change of variables $\mathcal{U}_{H'}$. This part is both nontrivial and important. To be more specific, let us try to define $\mathfrak{R}$ on a space $\mathcal{A}$ of analytic Hamiltonians (3.6) with finite norm

$$\|h\| = \sum_{\nu,k} |h_{\nu,k}| e^{\rho_0 |\omega \cdot \nu|} \cosh(\rho_1 \Omega \cdot \nu) \rho_2^k\,, \tag{3.7}$$

with $\rho_0, \rho_1, \rho_2 > 0$ fixed. (Later on, we also consider the choice $\rho_0 = 0$.) To simplify the description, let us include in $\mathcal{A}$ constant multiples of the linear Hamiltonian $(q,p) \mapsto \omega \cdot p$. If H belongs to $\mathcal{A}$ then H is analytic in a strip $|\operatorname{Im} \omega' \cdot q| < \rho_0$. But in general, $H' = \alpha^{-1} \Sigma_\mu (H \circ \mathcal{T})$ is analytic only in the narrower strip $|\operatorname{Im} \omega' \cdot q| < \alpha \rho_0$.

This can regarded as a normalization issue: We should restrict $\mathfrak{R}$ to a subspace $\mathcal{A}_{\mathrm{r}}$ of Hamiltonians H that are "normalized", in the sense that $H' = \alpha^{-1} \Sigma_\mu (H \circ \mathcal{T})$ belongs to $\mathcal{A}$, whenever $H \in \mathcal{A}_{\mathrm{r}}$. Then the domain of $\mathfrak{R}$ can be defined as the set of all normalized Hamiltonians for which there exists a canonical change of variables $\mathcal{U}_{H'}$ such that $\mathfrak{R}(H) = H' \circ \mathcal{U}_{H'}$ is again normalized. Of course we want this domain to be nontrivial.

This highlights an important difference between the renormalization of maps and flows. For maps, the re-normalization step $\hat{P} \mapsto \tilde{P}$ can be chosen within a finite-parameter group of similarity transformations. But for flows, this group is infinite dimensional. One source of this difference can be seen if we associate to a flow its first-return map to some transversal section. All variations in the first-return time drop out in this process. They corresponds roughly to the nonresonant modes described in Section 3.2.

Remark 2. The above-mentioned needs can be met in a perturbative setting [8,17], as well as for near-critical twist Hamiltonians [15]. But the construction of $\mathcal{U}_{H'}$ involves a Nash-Moser type iteration. This is what makes the Hamiltonian approach quite involved.

To be more specific, the canonical transformation $\mathcal{U}_{H'}$ will be obtained as a composition of infinitely many exact symplectic coordinates changes $U : (q,p) \mapsto (q',p')$. When close to the identity, such a change of coordinates can be defined implicitly by a generating function $\phi = \phi(q,p')$, via $q' = q + \nabla_p \phi(q,p')$ and $p' = p - \nabla_q \phi(q,p')$. Here ∇_q and ∇_p denote

the gradients with respect to the first and second argument, respectively. It is convenient to work with $\psi = \partial_x \phi$ instead of ϕ. Here $x = \omega' \cdot q$ and $\partial_x = \omega \cdot \nabla_q$. We restrict to changes of coordinates where ϕ and ψ depend on p' only via $z' = \omega \cdot p'$. Let $\Omega' \in \mathbb{R}^2$ be parallel to Ω, satisfying $\Omega' \cdot \Omega = 1$. Define $y = \Omega' \cdot q$ and $\partial_y = \Omega \cdot \nabla_q$, as well as $z = \Omega \cdot p$ and $\partial_z = \Omega' \cdot \nabla_p$. Then the composition of H with $U = U_\psi$ is given by

$$\begin{aligned}(H \circ U_\psi)(q,p) = H(q',p') &= \omega \cdot p' + h(q',z') \\ &= \omega \cdot p - \psi(q,p') + h\big(q + \mathcal{D}_z\psi(q,z')\Omega, z - \mathcal{D}_y\psi(q,z')\big)\,,\end{aligned} \tag{3.8}$$

where $\mathcal{D}_y = \partial_x^{-1}\partial_y$ and $\mathcal{D}_z = \partial_x^{-1}\partial_z$.

In the remaining part of this Section, we restrict to the shearless case.

3.1. Shearless Hamiltonians

The formulation (3.8) in terms of $\psi = \partial_x \phi$ is particularly useful when symmetries are involved. To give an example that is relevant here, let θ be one of the three vectors in $\Theta = \{(1,0),(0,1),(1,1)\}$. A Hamiltonian H with Fourier-Taylor coefficients $h_{\nu,k}$ is said to be θ-reversible if $h_{\nu,k} = 0$ unless $\theta \cdot \nu + k$ is odd. If both H and ψ is θ-reversible, then so is $H \circ U_\psi$. We will come back to this property below.

The renormalization of near-critical shearless Hamiltonians was first considered in [13]. See also [16]. To simplify notation, let use introduce the coordinate functions $\boldsymbol{q}(q,p) = q$ and $\boldsymbol{p}(q,p) = p$. A trivial orbit for $\mathfrak{R}$ is given by the pair of Hamiltonians

$$K_\pm = \omega \cdot \boldsymbol{p} \pm \tfrac{1}{2}\boldsymbol{z}^3\,, \qquad \boldsymbol{z} = \Omega \cdot \boldsymbol{p}\,, \tag{3.9}$$

in the sense that $\mathfrak{R}(K_\pm) = K_\mp$. This holds for the choice $\mu = \alpha^2$ and $\mathcal{U} = \mathrm{I}$. The coefficient $\pm\frac{1}{2}$ of z^3 represents an arbitrary normalization. $K_\pm$ has an invariant torus $\mathbb{T}^2 \times \{p\}$ for each p. The torus at $p = 0$ is shearless: it has a rotation number α that is maximal for H_- and minimal for H_+.

A nontrivial approximate period 12 for $\mathfrak{R}$ was found in [13]. We describe here a modified version, where all Hamiltonians are even in the torus variable q. Each Hamiltonian H on this orbit is Θ-reversible, meaning θ-reversible for some $\theta \in \Theta$. Notice that, if H is θ-reversible, then $H' = \alpha^{-1}\Sigma_\mu(H \circ \mathcal{T})$ is θ'-reversible, with $\theta' = T^{-1}\theta$. So the change of coordinates $H' \mapsto H' \circ \mathcal{U}_{H'}$ that is part of $\mathfrak{R}$, is defined in terms of generating functions ψ that are θ'-reversible as well. This way, $\mathfrak{R}$ maps θ-reversible Hamiltonians to θ'-reversible Hamiltonians, with $\theta' = T^{-1}\theta$. Notice that T^{-1} permutes the set Θ modulo 2. So any periodic orbit for $\mathfrak{R}$ that consists of Θ-reversible Hamiltonians must have a length that is a multiple of 6.

A map G is said to be reversible with respect to an involution J, if $JGJ = G^{-1}$. The analogue property for a Hamiltonian is much simpler, namely $HJ = -H$. This is our main reason for considering flows now instead of maps. Setting $J_\theta = V_\theta S_2$, where $V_\theta = (\boldsymbol{q} + \pi\theta, \boldsymbol{p})$ and $S_2 = \mathrm{diag}(1,-1)$, reversibility of H with respect to J_θ agrees with θ-reversibility.

Another interesting numerical observation in [13] was that the Hamiltonians on the approximate orbit of length 12 satisfy $\mathfrak{R}^6(H) = H \circ V_{\theta'}$, with $\theta' = T^{-1}\theta$, if H is θ-reversible.

So the approximate period 12 is in effect an approximate period 6. Without this symmetry, or without reversibility, every period 12 for $\mathfrak{R}$ would yield three distinct periods 12, via translations V_θ with $\theta \in \Theta$. Here, we are using that $\mathfrak{R}^6$ commutes with each translation V_θ. (A related fact is that T^6 is congruent to the 2×2 identity matrix modulo 2.) Similar uniqueness-promoting features can be found in other renormalization schemes.

Based on this observation, our goal was to find a fixed point of the transformation $\mathfrak{R}_6$ defined (after fixing some $\theta' \in \Theta$) by

$$\mathfrak{R}_6(H) = \mathfrak{R}^6(H) \circ V_{\theta'} \,. \tag{3.10}$$

3.2. Resonant normalization

Using the Fourier-Taylor representation (3.6), we define a projection $\mathbb{P}_{\rm r}$ on $\mathcal{A}$ by setting

$$\mathbb{P}_{\rm r} h = \sum_{\nu,k} \varrho_{\nu,k}\, h_{\nu,k} \cos(\nu \cdot \boldsymbol{q}) \boldsymbol{z}^k \,, \qquad h \in \mathcal{A} \,, \tag{3.11}$$

where $\varrho_{\nu,k} = 1$ if $\max\bigl\{\sigma|\Omega \cdot \nu|\,,\, \kappa k\bigr\} \geq |\omega \cdot \nu|$, and $\varrho_{\nu,k} = 0$ otherwise. Here σ and κ are fixed positive constants. The subspace $\mathcal{A}_{\rm r} = \mathbb{P}_{\rm r}\mathcal{A}$ will be referred to as the "resonant" subspace of $\mathcal{A}$. The "nonresonant" subspace is defined as $\mathcal{A}_{\rm n} = \mathbb{P}_{\rm n}\mathcal{A}$ via the complementary projection $\mathbb{P}_{\rm n} = {\rm I} - \mathbb{P}_{\rm r}$.

Recall from (3.4) that $\mathfrak{R}(H) = H' \circ \mathcal{U}_{H'}$, with $H' = \alpha^{-1}\Sigma_\mu(H \circ \mathcal{T})$. Assuming that $0 < \mu < \alpha$, a simple computation shows that the map $H \mapsto H'$, when restricted to the resonant subspace, is bounded and analyticity improving. So the domain of $\mathfrak{R}$ is in essence the set of all resonant Hamiltonians H, with the property that $H' \circ \mathcal{U}_{H'}$ is again resonant, for some canonical change of variables $\mathcal{U}_{H'}$.

In reference [8] and later work, the change of variables $\mathcal{U}_{H'}$ is constructed iteratively: If H_j is resonant up to an error of size ε, then we can determine a generating function ψ_j of order ε, such that

$$H_{j+1} \stackrel{\rm def}{=} H_j \circ U_{\psi_j} \tag{3.12}$$

is resonant up to an error of order roughly ε^2. This is essentially Newton's method. If this works, starting with $H_0 = H'$, then $\mathcal{U}_{H'} = U_{\psi_0} \circ U_{\psi_1} \circ U_{\psi_2} \circ \dots$ is a solution to our normalization problem. Otherwise, if H' is too far from being resonant, one can try to determine an initial ψ_0 by nonperturbative means, such that $H_1 = H' \circ U_{\psi_0}$ is sufficiently close to being resonant. Then the problem for $j \geq 1$ is again perturbative. Fortunately, this worked in [15] for near-critical twist Hamiltonians.

In summary, $\mathfrak{R}(H)$ is obtained from H by first computing $H_0 = \alpha^{-1}(\Sigma_\mu H) \circ \mathcal{T}$ and then eliminating the nonresonant part of H_0 by a sequence of coordinate changes (3.12).

Concerning the perturbative steps, we see from (3.8) that $H \circ U_\psi$ is resonant to first order precisely if

$$\mathbb{P}_{\rm n}[{\rm I} + \mathbb{D}(h)]\psi = \mathbb{P}_{\rm n} h \,, \qquad \mathbb{D}(h) \stackrel{\rm def}{=} [\partial_z h]\mathcal{D}_y - [\partial_y h]\mathcal{D}_z \,. \tag{3.13}$$

Given h, this is a linear equation for ψ. Restricting to $\psi \in \mathcal{A}_{\rm n}$, we can take advantage of the fact that $\mathcal{D}_y$ and $\mathcal{D}_z$ are continuous on $\mathcal{A}_{\rm n}$, with operator norms bounded by σ^{-1} and

$\kappa^{-1}\rho_2^{-1}$. Still, if $\partial_z h$ or $\partial_y h$ is not close to zero in $\mathcal{A}$, then there is no guarantee that the equation (3.13) can be solved.

3.3. Initial numerical experiments

The transformation $\mathfrak{R}$ has been implemented numerically. A function $h \in \mathcal{A}$ is represented as a cosine series $\sum_\nu h_\nu \cos(\nu \cdot \boldsymbol{q})$ with coefficients h_ν that are Taylor series $\sum_k h_{\nu,k} \boldsymbol{z}^k$. Sine series are represented analogously. At the precision that we refer to as `L-W-D`, the frequencies $\nu \in \mathbb{Z}^2$ are restricted to a lattice "parallelogram" of length roughly `2L` in the direction ω and width roughly `2W` in the direction Ω. A function d on this parallelogram defines the projection $\mathbb{P}_{\rm n}$ by specifying the maximum value $d(\nu)$ of k for which $\varrho_{\nu,k} = 1$ in (3.11). For the basic operations like sums, products, derivatives, compositions we use in essence the programs developed in [15] for a computer-assisted proof; but here we simply truncate series instead of carrying out error estimates.

As a first experiment, we tried to renormalize the Hamiltonians adapted from [13]. The step $H \mapsto H_0 = \alpha^{-1}\Sigma_\mu(H \circ \mathcal{T})$ is trivial. But H_0 is far from resonant, so the approach sketched after (3.12) is not even close to being applicable. Our first solution to this problem was to eliminate only a small fraction ε_0 of $\mathbb{P}_{\rm n} H_0$, by using $U_{\varepsilon_0\psi_0}$ instead of U_{ψ_0}. Here ψ_0 is determined as described above. Presumably, the nonresonant part of $H_1 = H_0 \circ U_{\varepsilon_0\psi_0}$ is smaller than that of H_0. Then the process can be repeated by eliminating only a fraction ε_j of $\mathbb{P}_{\rm n} H_j$ via $U_{\varepsilon_j\psi_j}$ for $j = 1, 2, \ldots$. Extensive experimentation and fine-tuning eventually resulted in a successful numerical computation of $\mathfrak{R}(H)$.

The main issue was to invert the operator $\mathrm{I} + \mathbb{P}_{\rm n}\mathbb{D}(h)\mathbb{P}_{\rm n}$ that appears in (3.13), on the nonresonant subspace $\mathcal{A}_{\rm n}$. We had already increased σ and κ far beyond the values used in [13], to 1.625 and 2.250, respectively. Still the spectral radius of $\mathbb{P}_{\rm n}\mathbb{D}(h)\mathbb{P}_{\rm n}$ was often larger than 1. This does not necessarily mean that $\mathrm{I} + \mathbb{P}_{\rm n}\mathbb{D}(h)\mathbb{P}_{\rm n}$ cannot be inverted. But nothing like this had happened in the renormalization analysis of twist Hamiltonians. It may not happen here either, once we are close to a periodic orbit of $\mathfrak{R}$. But there is no way of knowing, if we can barely compute $\mathfrak{R}^n(H)$ for $n = 1$.

Some effort was spent analyzing other two-parameter families of Hamiltonians. For simple Hamiltonians, critical parameter values were found by numerically integrating the orbit passing through the indicator point and searching for a critical golden torus. This strategy can be useful for finding an attracting set for $\mathfrak{R}$. But the non-invertibility problem for $\mathrm{I} + \mathbb{P}_{\rm n}\mathbb{D}(h)\mathbb{P}_{\rm n}$ did not improve and remained the main difficulty.

3.4. A normalization flow

An observation concerning the first experiment described above is that the effort of solving (3.13) exactly seems largely wasted, if only $\varepsilon\psi$ is used afterwards. So maybe there is no need to invert the operator $\mathrm{I} + \mathbb{P}_{\rm n}\mathbb{D}(h)\mathbb{P}_{\rm n}$ exactly. Also, an evolution via a large number of steps $U_{\varepsilon_j\psi_j}$ close to the identity, starts to resemble a time-dependent flow.

An infinitesimal change of coordinates via ψdt yields a change $\dot{h}dt$ with

$$\dot{h} = -[\mathrm{I} + \mathbb{D}(h)]\psi\,. \tag{3.14}$$

This describes an evolution $\dot{H} = \{\phi, H\}$ of H under the flow given by a time-dependent Hamiltonian $\phi = \partial_x^{-1}\psi$. In some way, we are trying to minimize a function like

$$E_{\scriptscriptstyle H}(\Psi) = \tfrac{1}{2}\|\mathbb{P}_{\rm n}(H \circ U_\Psi)\|^2 \tag{3.15}$$

by a procedure that resembles a gradient flow. Assume that the norm in (3.15) is given by an inner product, and that the projection $\mathbb{P}_{\rm n}$ is orthogonal. Then the gradient flow for $E_{\scriptscriptstyle H}$ at $\Psi = 0$ yields a rate of change $\dot\Psi = \psi$, with

$$\psi = \tau\big[\mathrm{I} + \mathbb{D}(h)^*\big]\mathbb{P}_{\rm n} h\,. \tag{3.16}$$

Here $\tau > 0$ is a suitable constant that defines a time scale. Concerning possible improvements, we note that the scalar τ could be replaced by a suitable positive operator.

A crucial aspect of the flow defined via (3.14) and (3.16) is that does not involve inverting any operator! After implementing this "normalization flow" numerically, the logical next step was to try finding a point near the stable manifold of the expected attractor for $\mathfrak{R}$. We started with a two-parameter family of Hamiltonians $H_{a,x} = H_{0,0} + aH_2 + xH_1$ and tried to determine the values of (a, x) for which $n \mapsto \mathfrak{R}^n(H_{a,x})$ neither diverges nor approaches a trivial orbit, after some controllable number of steps. This was repeated many times, using $\mathfrak{R}^6(H_{a,x})$ as a new starting point $H_{0,0}$ for the next family $H_{a,x} = H_{0,0} + aH_2 + xH_1$. The hope was that the resulting quasi-orbit ends up being shadowed by a true orbit. The Hamiltonians H_1 and H_2 are rough guesses for the two expanding eigenvectors of $D\mathfrak{R}^{12}$ at a fixed point. The values of a were chosen by hand, while for a fixed value of a, the parameter x was adjusted by a golden section algorithm that minimizes a suitable penalty function. The penalty included the inverse number of successful iterations, the norm of the nonresonant part $\mathbb{P}_{\rm n} H$ after each normalization step, and various measures for the amount of effort that was involved.

As a result of these changes and extensive iteration, we finally found an approximate period 2 for the transformation $\mathfrak{R}_6$ given by (3.10). In fact, both points on that period were rough fixed points of $\mathfrak{R}_6$. Strangely, $D\mathfrak{R}_6(H)$ had an eigenvalue λ_3 "near -1", in a sense described below. The fact that λ_3 is not near 1 helped us later to find a numerical fixed point of $\mathfrak{R}_6$ via a Newton method. Getting to this point was not easy, but the result was a Hamiltonian H_0 that differs from $\mathfrak{R}_6(H_0)$ by an an error 10^{-12} in norm. Of course this is purely numerical, and for a degree-truncated version of $\mathfrak{R}_6$. Still, small numerical errors in such computations indicate that the approach is basically sound.

The above-mentioned Hamiltonian H_0 is θ-reversible with $\theta = (0, 1)$. Its 16 leading coefficients are shown in Table 1. Additional data can be found in [25].

ν_1	ν_2	k	$h_{\nu,k}$
2	3	2	$-2.96211204830463497E+00$
0	1	0	$8.70097437356090711E-01$
0	0	3	$5.00000000000000000E-01$
1	1	2	$-2.97870505130652936E-01$
3	4	3	$-2.58446973495392146E-01$
0	1	2	$2.36034099941285126E-01$
1	3	2	$1.20306399275013194E-01$
0	0	1	$-8.92647025112508078E-02$
2	4	3	$-8.02494834569980432E-02$
0	2	1	$-7.41203812831851134E-02$
2	2	3	$6.43466790625083266E-02$
2	1	0	$6.12173500512698560E-02$
1	0	1	$-5.76400738354366048E-02$
1	2	3	$4.81657160846837909E-02$
2	3	4	$-3.10186284729477973E-02$
1	1	4	$2.20086515859728277E-02$

Table 1. Leading 16 coefficients of an approximate fixed point H_0 of $\mathfrak{R}_6$.

Numerically, the two largest eigenvalues λ_1 and λ_2 for $D\mathfrak{R}_6(H_0)$ satisfy $\lambda_1^{1/6} \approx 2.6184$ and $\lambda_2^{1/6} \approx 1.6106$. For the average scaling $\bar{\mu} = (\mu_0\mu_1\cdots\mu_5)^{1/6}$ per renormalization step we find $\bar{\mu} \approx 0.3695$. All these values agree quite nicely with earlier estimates based on bifurcation sequences for the standard nontwist map; see [9,14] and references therein. These computations were carried out at degrees `11-5-9`.

3.5. Numerical Improvements

Unfortunately, $D\mathfrak{R}_6(H_0)$ has another non-contracting eigenvalue λ_3. It is negative, with $|\lambda_3|^{1/6} \approx 1.0680$. Our guess was that we overlooked some symmetry, and that λ_3 approximates an eigenvalue -1 for the un-truncated system. The existence of symmetry-related eigenvalues is quite common in renormalization. But we found no good reason for the existence of an eigenvalue like λ_3. It seemed time to try finding an approximate fixed point of $\mathfrak{R}_6$ at higher degrees. This was done using a continuous homotopy between the restriction of $\mathfrak{R}$ to degrees `11-5-9` and the restriction of $\mathfrak{R}$ to degrees `13-6-9`. But all attempts kept failing at about 2/3 into the homotopy. Returning to `11-5-9`, our next goal was to improve on the choice of the generating function ψ.

The plain gradient flow may be too crude here. Consider the projections of (3.14) and (3.16) onto the nonresonant subspace $\mathcal{A}_{\rm n}$. The projected equation (3.14) is $A\psi = -\mathbb{P}_{\rm n}\dot{h}$, with $A = \mathrm{I} + \mathbb{P}_{\rm n}\mathbb{D}(h)\mathbb{P}_{\rm n}$. And (3.16) is $\psi = B_0\mathbb{P}_{\rm n}h$, with $B_0 = \tau A^*$. We note that, for matrices, $B_0 = \tau A^*$ is a standard choice for an approximate inverse of A. Better inverses $B_1, B_2, \ldots$ can be obtained via the recursion $B_{k+1} = B_k(2\mathrm{I} - P_k)$ with $P_k = AB_k$. This is in fact Newton's method for finding the inverse of a matrix. For details see [1]. Under a simple condition on τ, the sequence $k \mapsto B_k$ converges to the inverse of A, if it exists. Otherwise, the sequence $k \mapsto P_k$ converges to the orthogonal projection onto the range of A. In either case, nothing diverges, and useful information could be gained.

Replacing B_0 in (3.16) by B_k yields $\mathbb{P}_{\rm n}\dot{h} = -P_k\mathbb{P}_{\rm n}h$. This represents a safe and more stable alternative to our original attempt at inverting A. We have tested it numerically for $k = 1, 2, 3$. Somewhat unexpectedly, the results were similar to those obtained for $k = 0$. (As it turned out later, there was an error in our implementation of B_k for $k > 0$.)

In addition to the above, we also implemented a number of improvements to our numerical integration of (3.14). This included keeping the change of variables uniformly small. Less emphasis was put on having $\mathbb{P}_{\rm n}h$ exactly zero at the end.

In this context, we should mention that we used $\rho_0 = 0$ in the norm (3.7) that defines the space $\mathcal{A}$. This makes the scaling $H \mapsto H' = \alpha^{-1}\Sigma_\mu(H \circ \mathcal{T})$ nonsingular on $\mathcal{A}$. It does not make coordinate changes obsolete, but it eliminates the need to normalize $\mathbb{P}_{\rm n}(H' \circ \mathcal{U}_{H'}) = 0$ exactly. This seems somewhat at odds with the concept of "normalization". We will comment on this below.

To summarize our latest findings: We saw no more periodic orbits after the above-mentioned improvements. The attractor still behaves like a near-period 12, but the previous eigenvalues λ_2 and λ_3 have turned into an expanding complex-conjugate pair. This was unexpected, but it seems consistent with the observation of asymptotic rotation for the nontwist standard map near the indicator point.

4. Concluding remarks

It seems possible that the renormalization transformation (2.3) for commuting maps has indeed a period 12, if one scales about a point $w = w_*$ on the symmetry line $x = 0$. Scaling about the indicator point $\zeta_\circ = \left(\frac{1}{2}, 0\right)$ seems more interesting, since it focuses on a feature that is specific to the shearless case. But any asymptotic scaling about his point cannot be too simple. Based on our computation of shearless p_n/q_n orbits, periodicity with a period shorter than 24 seems unlikely in this case. And scaling appears to involve a rotation. Independently of all this, our main reason for not pursuing this approach was that we did not find a satisfactory representation of $\mathfrak{R}$ with reversible pairs.

One may wonder whether the different asymptotics found near $(0, w_*)$ and near $\zeta_\circ$ are compatible with the general picture of renormalization. We believe that they are, since the associated renormalization transformations zoom in on different parts of the orbit. Something analogous occurs in the renormalization of skew-product maps: a period 3 and a period 6, both describing the same system [24]. The situation is less clear for the renormalization of Hamiltonians, since it is more "global". No part of $\mathbb{T}^2$ is being scaled off to infinity while one zooms in on another part. So it is unclear what kind of $\mathfrak{R}$-orbit for Hamiltonians could describe the two distinct behaviors that are observed for maps near $(0, w_*)$ and near $\zeta_\circ$.

In hindsight, our expectation that the transformation (3.4) for Hamiltonians should have an orbit of period 6 or 12 seems unrealistic. However, it has led to the development of a more robust re-normalization procedure: a flow on a space of Hamiltonians that converts a given Hamiltonian to a resonant normal form. What is puzzling is that the breakup of shearless tori in Hamiltonian flows no longer looks like a codimension 2 phenomenon. We have no good explanation for this.

One may wonder why this was not seen earlier in [13]. A possible reason may be that skipped normalization steps resulted in some frequencies becoming too large and

then getting eliminated simply by truncation. Truncation can turn a potentially relevant eigenvalue into a contracting eigenvalue.

Our normalization flow is based on a notion of "resonant" that is somewhat ad-hoc. (Less so than it may seem, as it works fine in the twist setting.) It seems worth trying to find/use a normalization that is more intrinsic.

Numerically, significant further progress may require significantly faster computer hardware. This certainly applies to the renormalization of Hamiltonians. For commuting pairs, what could change the situation is a suitable representation that takes advantage of all symmetries and constraints, including reversibility.

Acknowledgments
The author would like to thank D. Gaidashev for helpful discussions.

References

[1] A. Ben-Israel, *A note on an iterative method for generalized inversion of matrices*, Mathematics of Computation, **20**, 439–440 (1966).
[2] R.S. MacKay, *Renormalisation in Area Preserving Maps.* Thesis, Princeton (1982). World Scientific, London (1993).
[3] J.-P. Eckmann, H. Koch, P. Wittwer, *A computer-assisted proof of universality for area-preserving maps*, Memoirs of the AMS **47**, 1–121 (1984).
[4] W. de Melo, S. van Strien, *One-dimensional dynamics*, Modern Surveys in Mathematics, Series 3, Volume 25, Springer-Verlag, 1993.
[5] J.J. Abad, H. Koch, P. Wittwer, *A renormalization group for Hamiltonians: numerical results*, Nonlinearity **11**, 1185–1194 (1998).
[6] D. Del-Castillo-Negrete, J.M. Greene, P.J. Morrison, *Area preserving non-twist maps: Periodic orbits and transition to chaos.* Physica D **91**, 1–23 (1996).
[7] D. Del-Castillo-Negrete, J.M. Greene, P.J. Morrison, *Renormalization and transition to chaos in area preserving non-twist maps.* Physica D **100**, 311–329 (1997).
[8] H. Koch, *A renormalization group for Hamiltonians, with applications to KAM tori*, Erg. Theor. Dyn. Syst. **19**, 1–47 (1999).
[9] P.J. Morrison, *Magnetic field lines, Hamiltonian dynamics, and nontwist systems*, Physics of Plasmas **7**, 2279–2289 (2000).
[10] E. Petrisor, *Nontwist area preserving maps with reversing symmetry group*, Int. J. Bifurcat. Chaos **11**, 497–511 (2001).
[11] H. Koch, *On the renormalization of Hamiltonian flows and critical invariant tori*, Discrete Contin. Dynam. Systems **8**, 633–646 (2002).
[12] A. Apte, A. Wurm, P.J. Morrison, *Renormalization and destruction of $1/\gamma^2$ tori in the standard nontwist map.* Chaos **13**, 421–433 (2003).
[13] D. Gaidashev and H. Koch, *Renormalization and shearless invariant tori: numerical results*, Nonlinearity **17**, 1713–1722 (2004).
[14] A.S. Apte, *Numerical studies of the standard nontwist map and a renormalization group framework for breakup of invariant tori*, Ph.D. Thesis, UT Austin (2004).
[15] H. Koch, *A renormalization group fixed point associated with the breakup of golden invariant tori*, Discrete Contin. Dynam. Systems **11**, 881–900 (2004)
[16] D.G. Gaidashev, *Renormalization of isoenergetically degenerate Hamiltonian flows and associated bifurcations of invariant tori*, Discrete Contin. Dynam. Systems **13**, 63–102 (2005).

[17] K. Khanin, J. Lopes Dias, J. Marklof, *Multidimensional continued fractions, dynamic renormalization and KAM theory*, Commun. Math. Phys. **270**, 197–231 (2007).
[18] Koch, S. Kocić, *Renormalization of Vector Fields and Diophantine Invariant Tori*, Erg. Theor. Dyn. Syst. **28**, 1559–1585 (2008).
[19] P. Collet, J.-P. Eckman, *Iterated maps on the interval as dynamical systems* , Modern Birkhäuser Classics, Birkhäuser Boston, MA, 2009.
[20] Koch, S. Kocić, *A renormalization group approach to quasiperiodic motion with Brjuno frequencies*, Erg. Theor. Dyn. Syst. **30**, 1131–1146 (2010).
[21] G. Arioli and H. Koch, *The Critical Renormalization Fixed Point for Commuting Pairs of Area-Preserving Maps*, Commun. Math. Phys. **295**, 415–429 (2010).
[22] A. González-Enriquez, A. Haro, R. de la Llave, *Singularity Theory for Non-Twist KAM Tori*, Memoirs of the AMS, **227**, 115p (2014).
[23] H. Koch, *On hyperbolicity in the renormalization of near-critical area-preserving maps*, Discrete Contin. Dynam. Systems A. **36**, 7029–7056 (2016).
[24] H. Koch, *Asymptotic scaling and universality for skew products with factors in* $\mathrm{SL}(2,\mathbb{R})$, Erg. Theor. Dyn. Syst. (2022), 39pp. doi:10.1017/etds.2022.22
[25] H. Koch, The source code for our programs, and some data, are available at web.ma.utexas.edu/users/koch/papers/shearless-numrg/