Physica D 449 (2023) 133749

Contents lists available at ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd

Distinguishing between regular and chaotic orbits of flows by the weighted Birkhoff average

Nathan Duignan [a,*], James D. Meiss [b]

[a] Department of Mathematics and Statistics, University of Sydney, NSW, 2006, Australia
[b] Department of Applied Mathematics, University of Colorado, Boulder, CO, 80309-0526, USA

ARTICLE INFO

Article history:
Received 21 December 2022
Received in revised form 15 March 2023
Accepted 7 April 2023
Available online 17 April 2023
Communicated by A. Celletti

Keywords:
Birkhoff ergodic theorem
Chaos
Integrability
Dynamical systems
Hamiltonian
Flow

ABSTRACT

This paper investigates the utility of the weighted Birkhoff average (WBA) for distinguishing between regular and chaotic orbits of flows, extending previous results that applied the WBA to maps. It is shown that the WBA can be super-convergent for flows when the dynamics and phase space function are smooth, and the dynamics is conjugate to a rigid rotation with Diophantine rotation vector. The dependence of the accuracy of the average on orbit length and width of the weight function width are investigated. In practice, the average achieves machine precision of the rotation frequency of quasiperiodic orbits for an integration time of $\mathcal{O}(10^3)$ periods. The contrasting, relatively slow convergence for chaotic trajectories allows an efficient discrimination criterion. Three example systems are studied: a two-wave Hamiltonian system, a quasiperiodically forced, dissipative system that has a strange attractor with no positive Lyapunov exponents, and a model for magnetic field line flow.

1. Introduction

Integrability is associated with quasiperiodic dynamics and chaos with sensitive dependence on initial conditions. This contradistinction is especially relevant for smooth Hamiltonian systems: when such a flow is *integrable* the orbits are confined to tori on which the dynamics is conjugate to a rigid rotation. When a Hamiltonian system is smoothly perturbed away from integrability, some of these tori persist—according to KAM theory—and some are replaced by isolated periodic orbits, islands, or chaotic regions [1]. Typically as a perturbation grows the proportion of chaotic orbits increases and more of the tori are destroyed.

Motivated by broad applications to dynamical systems, including fluid flow, the n-body problem, and toroidal magnetic confinement, there has been a concentrated effort to distinguish between chaotic regions of phase space and those with regular dynamics. Invariant tori in Hamiltonian systems can be computed as limits of periodic orbits [2,3] or by the iterative, parameterization method [4]. In these methods, one fixes a frequency vector and attempts to find an invariant set on which the dynamics has this frequency.

More generally—even when the system is not Hamiltonian—one may try to detect when a given orbit is chaotic. By definition, a dynamical system is chaotic on a compact invariant set when it is transitive and exhibits "sensitive dependence on initial conditions" [5,6]. Often such dynamics are (nonuniformly) hyperbolic, meaning that the maximal Lyapunov exponent is positive [7]. For a flow

$$x_t = \varphi_t(x_0), \tag{1}$$

on a phase space X, the exponent is

$$\lambda(x_0, v_0) = \limsup_{T\to\infty} \frac{1}{T} \log \|D\varphi_t(x_0)v_0\|. \tag{2}$$

Here $D\varphi_t(x_0)$ is the Jacobian matrix—the derivative of the flow map at time t with respect to x_0, an initial condition—and v_0 is the initial deviation vector. A positive Lyapunov exponent implies that the length of the infinitesimal deviation grows exponentially in T, at least asymptotically. The most common approach for distinguishing chaos from regularity is to numerically compute (2); however, accurate computation of λ is difficult because convergence is typically as slow as $\frac{\log(T)}{T}$ [8]. Computation of (2) is also expensive because it is necessary to integrate both the trajectory and the linearized dynamics to obtain the Jacobian $D\varphi_t(x_0)$.

There are a number of techniques that have been used to improve the efficiency of estimates for exponential divergence. These include methods based on (2) such as the Fast Lyapunov Indicator (FLI) [9,10], which uses a large value of

$$\mathrm{FLI}(x_0, v_0) = \sup_{t<T} \log \|D\varphi_t(x_0)v_0\| \tag{3}$$

* Corresponding author.
E-mail addresses: Nathan.Duignan@sydney.edu.au (N. Duignan), James.Meiss@colorado.edu (J.D. Meiss).

https://doi.org/10.1016/j.physd.2023.133749
0167-2789/

as an indicator for chaos, where one normalizes the initial deviation, $\|v_0\| = 1$. A related idea, the Mean Exponential Growth factor of Nearby Orbits (MEGNO) [8,11], uses—instead of the supremum in (3)—the average of this log-length along an orbit. Further techniques include computing Greene's residue [2], Slater's method [12], the 0–1 test [13], SALI and GALI [14], expansion entropy [15], and converse KAM theory [16,17]. Unlike directly computing the Lyapunov spectrum, we remark that several of these methods do not require the integration of the full linearized dynamics, but only of a single vector. Consequently, it can be more efficient to use "jet transport" techniques to propagate a neighborhood of an initial condition.

In this paper, we explore an alternative technique to distinguish between chaotic and regular orbits of a flow based on the Weighted Birkhoff Average (WBA) [18–20]. This method permits one to accurately and efficiently compute the average of a function $h : X \to \mathbb{R}$, when the orbit is regular. In particular, h can be chosen to give the rotation vector for a regular orbit on an invariant torus, so that it can also provide a distinction between resonant and quasiperiodic dynamics. The technique is analogous to "frequency analysis" [21,22], which uses a windowed Fourier transform to compute rotation numbers. However, for the WBA, the choice of a smooth window or weight function allows for more rapid convergence. There are several other techniques specifically for maps that enable accurate computation of rotation numbers. For 1D invariant circles, an order-based technique such as Slater's method [12], or the topological method of [23,24] can be used. More generally, an averaging technique, that is distinct from WBA, was demonstrated in [25,26]. These techniques are specific to maps and yield theoretical convergence results that are slower than those for weighted Birkhoff averaging.

Indeed, as we recall in Section 2, the method proposed in [18] uses a C^∞ weight function, which has been shown by [27] to lead to super-polynomial convergence of the average for maps. We will generalize these results for flows in Section 3. In Theorem 3, we establish the super-polynomial convergence of the WBA to the space average provided that the flow is quasiperiodic on an n-torus with Diophantine rotation vector. Theorem 4 extends this, using the results of [28], to give weaker criteria for super-polynomial convergence.

In Section 4 we use the distinction between convergence rates for regular and chaotic orbits to give a criterion for detecting chaos. The method is analogous to that used in [29,30] for maps. In that paper, the WBA was shown to be a more efficient test for chaos than the FLI and 0–1 test for Chirikov's area-preserving standard map. Finally, in Section 5, we apply this test to three examples.

The first application, in Section 5.1, is to the two-wave model, perhaps the simplest nonintegrable, $1\frac{1}{2}$ degree-of-freedom Hamiltonian system. The model was also studied in [17,31] using converse KAM theory to detect chaos, and consequently gives a contrast between the two methods. We also investigate the dependence of the accuracy of the WBA on the choice of orbit length and weight function width.

In Section 5.2 we investigate the properties of the WBA for a quasiperiodically forced pendulum that can have geometrically strange attractors with no positive Lyapunov exponents [32]. Since the general definition of chaos requires only the topological form of sensitive dependence on initial conditions and not exponential divergence [15], orbits of this quasiperiodic system with zero exponents may still be chaotic [33]. As we will show, even though methods based on Lyapunov exponents would fail, the WBA can still provide an efficient indicator of chaos.

The final application, in Section 5.3, is to magnetic field line flow. Integrable magnetic field-line configurations are desirable in the design of plasma confinement devices. For example, the tokamak is designed to have a set of nested, axisymmetric tori that are tangent to the magnetic field and correspond to isopressure surfaces [34]. However, integrability can be destroyed by instabilities or non-axisymmetric perturbations that can give rise to magnetic islands and chaotic regions [35]. Using a model introduced by [36], we will show that the weighted Birkhoff average can rapidly and accurately measure the extent of these regions. Moreover, in the study of plasma stability it is important to know the rotation number, or rotational transform, on each magnetic surface. We will demonstrate that the weighted Birkhoff average efficiently computes the rotational transform.

2. Weighted Birkhoff averaging for flows

2.1. Birkhoff average

Suppose that $(X, \mathcal{B}, \mu)$ is a compact probability space with Borel σ-algebra $\mathcal{B}$ [37], and that the probability measure μ is invariant under a flow (1). The flow is *ergodic* with respect to μ if, whenever $S \in \mathcal{B}$ is invariant, i.e., $\varphi_t(S) = S$, $\forall t \in \mathbb{R}$, then $\mu(S) = 0$ or 1. Thus invariant sets are either of zero measure, such as periodic orbits, or of full measure, such as chaotic orbits. Birkhoff's ergodic theorem [38], states that if a flow is ergodic then for μ-almost every $x \in X$, the time average of a function h,

$$B_T(h)(x) = \frac{1}{T}\int_0^T h(\varphi_t(x))dt, \tag{4}$$

converges to its space average

$$\langle h\rangle \equiv \int_X h d\mu \tag{5}$$

as $T \to \infty$.

Theorem 1 (*Birkhoff Ergodic Theorem [38,39]*)**.** *Suppose $(X, \mathcal{B}, \mu)$ is a probability space, $\varphi_t : X \to X$ is a measure-preserving, ergodic flow, and $h \in L^1(X, \mu)$. Then the time average exists and*

$$\lim_{T\to\infty} B_T(h)(x) = \langle h\rangle$$

for μ-almost every $x \in X$.

The Birkhoff ergodic theorem for a map $F : X \to X$ can be obtained upon replacing the integration in (4) by a sum,

$$B_N(h)(x) = \frac{1}{N}\sum_{j=0}^{N-1} h(F^j(x)).$$

Indeed, in the literature, Theorem 1 is almost exclusively stated and proven for maps [40–42]. However, as pointed out in [43], there is a neat trick to obtain the continuous case from the map case.

This is based on the result that if $\varphi_t(x)$ is an ergodic flow then for each $\tau \in \mathbb{R}$, except for a countable subset, the map, φ_τ, is also ergodic [44]. For such a value of τ, define $\tilde{h}(x) = \int_0^\tau h(\varphi_t(x))dt$ and the iterated time-τ map $\varphi_\tau^j = \varphi_{j\tau}$ for each $j \in \mathbb{N}$. Then

$$\int_0^T h(\varphi_t(x))dt = \sum_{j=1}^{\lfloor T/\tau\rfloor} \tilde{h}(\varphi_\tau^j(x)) + \int_{\lfloor T/\tau\rfloor}^T h(\varphi_t(x))dt,$$

since $\varphi_\tau^0(x) = x$. Since h is assumed to be $L^1(X, \mu)$, then $\lim_{T\to\infty} \frac{1}{T}\int_{\lfloor T/\tau\rfloor}^T h(\varphi_t(x))dt = 0$. By the Birkhoff ergodic theorem for maps,

it then follows that

$$\begin{aligned}\lim_{T\to\infty}\frac{1}{T}\int_0^T h(\varphi_t(x))dt &= \frac{1}{\tau}\lim_{T\to\infty}\frac{1}{T/\tau}\sum_{j=1}^{\lfloor T/\tau\rfloor}\tilde{h}(\varphi_\tau^j(x))\\ &= \frac{1}{\tau}\int_X \tilde{h}d\mu = \frac{1}{\tau}\int_X\int_0^\tau h(\varphi_t(x))d\mu\\ &= \frac{1}{\tau}\int_0^\tau\int_X h(x)d\mu\\ &= \int_X hd\mu.\end{aligned}$$

Here the penultimate equality is a consequence of Fubini's theorem and the fact that μ is invariant under φ_t.

Even though the convergence in Theorem 1 is guaranteed, it can be arbitrarily slow depending on the choice of h [45]. Furthermore for almost all $h \in L^1(X,\mu)$, it has been demonstrated that for maps the convergence is at most $\mathcal{O}(1/N)$ [46], and for flows on a Lebesgue space is at most $\mathcal{O}(1/T)$ for almost all $x \in X$ [28]. The only exceptions are when h is almost everywhere constant.

2.2. Weighted Birkhoff average

A weighted Birkhoff average is analogous to (4), with the addition of a weight function $g : [0,1] \to [0,\infty)$, that is normalized:

$$\|g\|_1 \equiv \int_0^1 g(s)ds = 1. \tag{6}$$

For any $g \in \mathcal{G}$, the weighted Birkhoff average is defined by

$$WB_T(h)(x) = \frac{1}{T}\int_0^T g(\tfrac{t}{T})h(\varphi_t(x))dt. \tag{7}$$

In particular, choosing $g(s) = 1$ gives the Birkhoff average (4).

By a judicious choice of the weight g, the convergence of $WB_T(h)$ to the space average can be accelerated, for certain orbits; i.e., as we demonstrate below in Theorem 3, when the dynamics is conjugate to a sufficiently incommensurable rotation. In particular, we will consider the space of bump functions whose value and first $m-1$ derivatives vanish on the boundary,

$$\begin{aligned}\mathcal{G}_m = \big\{ g \in C^m([0,1],\mathbb{R}^+) \,\big|\, \|g\|_1 = 1,\\ g^{(i)}(0) = g^{(i)}(1) = 0,\ i = 0,1,\dots,m-1\big\}.\end{aligned} \tag{8}$$

For example, the smooth bump function

$$g(s) = \begin{cases} Ce^{-[s(1-s)]^{-1}} & s \in (0,1)\\ 0 & s = 0,1\end{cases}, \tag{9}$$

is in $\mathcal{G}_\infty$ and was adopted by [19,20,47] in their studies of maps. Here, we set the normalization constant $C \approx 142.2503758$ to satisfy (6).

Whenever $g \in \mathcal{G}_1$ it can be shown, using a result of [48], that the weighted Birkhoff average $WB_T(h)(x)$ converges to the space average $\langle h\rangle$ for any $h \in L^1(X,\mu)$. Of course this applies to the case (9) as well.

Proposition 2. *Under the hypotheses of Theorem 1, then whenever $g \in \mathcal{G}_1$*

$$\lim_{T\to\infty} WB_T(h) = \langle h\rangle,$$

for μ-almost every $x \in X$.

Proof. The proof relies on the summation criteria due to Silverman [48]. Specifically, suppose that $k(T,t)$ is a function defined for $T \in \mathbb{R}^+$ and $0 \le t \le T$, that is integrable for fixed T and satisfies the three criteria:

$$\lim_{T\to\infty}\int_0^T k(T,t)dt = 1; \tag{10a}$$

$$\lim_{T\to\infty} k(T,t) = 0 \text{ uniformly in } t \text{ on } [0,q] \text{ for each } q \in \mathbb{R}^+; \tag{10b}$$

$$\int_0^T |k(T,t)|dt < A \text{ for some } A > 0 \text{ whenever } T \in \mathbb{R}^+. \tag{10c}$$

Then Silverman [48, Thm. 1] shows that whenever $u : \mathbb{R}^+ \to \mathbb{R}$ satisfies $\lim_{T\to\infty} u(T) = \bar{u} < \infty$

$$\lim_{T\to\infty}\int_0^T k(T,t)u(t)dt = \bar{u}.$$

We will take $u(T) = B_T(h)(x)$ given by (4). Using integration by parts on (7), and the assumptions that $g \in C^1([0,1],\mathbb{R}^+)$ and $g(0) = 0 = g(1)$, then gives

$$\begin{aligned}WB_T(h)(x) &= \frac{1}{T}\int_0^T g(\tfrac{t}{T})h(\varphi_t(x))\,dt\\ &= \frac{1}{T}\left[g(\tfrac{t}{T})\int_0^t h(\varphi_s(x))\,ds\right]_0^T\\ &\quad - \frac{1}{T^2}\int_0^T g'(\tfrac{t}{T})\int_0^t h(\varphi_s(x))\,ds\,dt\\ &= \int_0^T k(T,t)B_t(h)(x)\,dt.\end{aligned} \tag{11}$$

where we set

$$k(T,t) = -\frac{t}{T^2}g'(\tfrac{t}{T}). \tag{12}$$

We will check Silverman's criteria for (12). Again using integration by parts we have

$$-\int_0^T \frac{t}{T^2}g'(\tfrac{t}{T})\,dt = -\left[\frac{t}{T}g(\tfrac{t}{T})\right]_0^T + \frac{1}{T}\int_0^T g(\tfrac{t}{T})dt = 1,$$

since $\|g\|_1 = 1$. It follows that k is integrable for each $T \in \mathbb{R}^+$ and that (10a) holds.

Now, for any $q > 0$, $0 \le t \le q < T$ we have that

$$|k(T,t)| = \frac{t}{T^2}|g'(\tfrac{t}{T})| \le \frac{q}{T^2}\|g'\|_\infty,$$

where $\|g'\|_\infty < \infty$ because g is continuously differentiable on a compact interval. It follows that $\lim_{T\to\infty} k(T,t) = 0$ and thus Eq. (10b) holds.

Finally, for all $T \in \mathbb{R}^+$, (12) gives

$$\int_0^T |k(T,t)|dt \le \|g'\|_\infty T^{-2}\int_0^T t\,dt = \tfrac{1}{2}\|g'\|_\infty.$$

Thus (10c) holds with $A = \frac{1}{2}\|g'\|_\infty$.

Hence, the choice (12) satisfies the criteria (10). Since by Birkhoff's ergodic theorem, $B_T(h)(x) \to \langle h\rangle$ for μ-almost all x, then [48, Thm. 1] applied to (11), gives $WB_T(h)(x) \to \langle h\rangle$. □

3. Super-convergence for flows

It was shown in [20,27] that if a smooth map F has a quasiperiodic orbit $\{F^t(x)\}$ with Diophantine rotation vector and h and g are C^∞, then the weighted average (7) is *super-polynomial convergent*: $WB_T(h)$ converges to $\langle h\rangle$ faster than any power of T. Super-convergence is especially useful for the case of Hamiltonian flows or symplectic maps, where regular orbits lie on invariant tori, and KAM theory implies the structural stability of those with Diophantine rotation vectors.

In this section we will extend the map result to the case of flows. In addition, following [28], we will show that super-polynomial convergence also holds under weaker hypotheses on the ergodic flow and h.

Note that if the weight function g has only finitely many vanishing derivatives at the endpoints, (8), then the convergence rate is T^{-m}. For example the weight $g(t) \propto \sin^2(\pi t)$ for $t \in (0, 1)$, is first order smooth, but not second order, since $g^{(2)}(0^+) \neq g^{(2)}(0^-)$, and this implies that the convergence is $\mathcal{O}(T^{-2})$.

By contrast, the weighted average appears to converge only as $T^{-1/2}$ for chaotic orbits [29,49]. Thus, as we will discuss in Section 4, the convergence rate of the weighted average can provide a useful distinction between regular and chaotic orbits [29,30].

Definition 1 (*Super-Convergent*)**.** A function $f : [0, \infty) \to \mathbb{R}$ with $\lim_{T\to\infty} f(T) = f^* < \infty$ is *super-polynomial* convergent if, for each $m \in \mathbb{N}$, there is a constant $c_m > 0$ such that

$$|f(T) - f^*| \le c_m T^{-m},$$

for all $T > 0$.

In particular the weighted Birkhoff average super-converges to the space average for flows that are conjugate to a rigid rotation with a sufficiently irrational vector, e.g., one that satisfies a Diophantine property.

Definition 2 (*Diophantine [50]*)**.** A vector $\omega \in \mathbb{R}^d$ is *Diophantine* if there is a $c > 0$ and $\tau \ge d - 1$ such that

$$\omega \in \mathcal{D}_{c,\tau} := \left\{\omega \in \mathbb{R}^d \mid |k \cdot \omega| > c\|k\|_\infty^{-\tau},\ \forall k \in \mathbb{Z}_0^d\right\}, \tag{13}$$

where $\mathbb{Z}_0^d = \mathbb{Z}^d \setminus 0$, and $\|k\|_\infty$ is the sup-norm of k.

Thus, for example, the vector $\omega = (\phi, 1) \in \mathbb{R}^2$, where $\phi = \frac{1}{2}(\sqrt{5} - 1)$ is the (inverse of) the golden mean, is in $\mathcal{D}_{1/\sqrt{5},1}$. More generally, there are bounds for the Diophantine constants for integral bases of an algebraic field of degree d [51,52].[1]

With these definitions, we can restate the result of [27] for quasiperiodic flows.

Theorem 3. *Let $M \simeq \mathbb{T}^d$ be a smooth manifold and $\varphi_t : M \to M$ be a smooth, quasiperiodic flow with invariant probability measure μ. Assume φ_t is C^∞ conjugate to a rigid rotation with a Diophantine rotation vector $\omega \in \mathcal{D}_{c,\tau}$. Suppose that $h \in C^\infty(M, \mathbb{R})$ and $g \in \mathcal{G}_\infty$. Then for each $x \in M$, the weighted Birkhoff average (7) is super-polynomial convergent to the space average. Moreover, the convergence is uniform in x.*

More generally, if $g \in \mathcal{G}_m$ then the convergence of (7) is as T^{-m} provided that $h \in C^l(M, \mathbb{R})$ for some $l > d + m\tau$.

Proof. The proof follows the arguments of [27] with some minor alterations for the flow case. By assumption φ is diffeomorphic to the flow $\varphi_t^\omega(\theta) = \theta + t\omega$ on $\mathbb{T}^d$. Hence, it can be assumed that we have taken coordinates $\theta \in \mathbb{T}^d$ on M so that φ_t is simply φ_t^ω and the invariant measure μ becomes the constant measure $d\theta$, which is preserved by φ_t^ω. Then

$$h(\varphi_t(\theta)) = h(\theta + t\omega).$$

The weighted Birkhoff average (7) then becomes

$$WB_T(h)(\theta) = \frac{1}{T}\int_0^T g(\tfrac{t}{T})h(\theta + t\omega)dt.$$

Since h is L^2, it has a Fourier series

$$h = \sum_{k\in\mathbb{Z}^d} a_k e^{2\pi i k\cdot\theta}.$$

that is almost everywhere convergent [53]. Note that $a_0 = \int_M h d\theta$ and that $a_0 = WB_T(a_0)(\theta_0)$ for any θ_0 and T since a_0 is constant. It follows that

$$\begin{aligned}\left|WB_T(h)(\theta_0) - \int_M h d\theta\right| &= \left|\sum_{k\in\mathbb{Z}_0^d} a_k WB_T\left(e^{2\pi i k\cdot\theta}\right)(\theta_0)\right| \\ &= \frac{1}{T}\left|\sum_{k\in\mathbb{Z}_0^d} a_k e^{2\pi i k\cdot\theta_0}\int_0^T g(\tfrac{t}{T})e^{2\pi i t k\cdot\omega}dt\right| \\ &\le \frac{1}{T}\sum_{k\in\mathbb{Z}_0^d}|a_k|\left|\int_0^T g(\tfrac{t}{T})e^{2\pi i t k\cdot\omega}dt\right|.\end{aligned}$$

Set $s = t/T$ so that $Tds = dt$ and define $\Omega_k = 2\pi T k\cdot\omega$. Then

$$\left|WB_T(h)(\theta_0) - \int_M h d\theta\right| \le \sum_{k\in\mathbb{Z}_0^d}|a_k|\left|\int_0^1 g(s)e^{i\Omega_k s}ds\right|.$$

Integrating by parts $m \le l$ times and noting that the boundary terms vanish by (8), it follows that

$$\begin{aligned}\left|\int_0^1 g(s)e^{i\Omega_k s}ds\right| &= |\Omega_k|^{-m}\left|\int_0^1 g^{(m)}(s)e^{i\Omega_k s}ds\right| \\ &\le |\Omega_k|^{-m}\|g^{(m)}\|_1,\end{aligned}$$

where $\|\ \|_1$ is the L^1 norm.

Now, if $h \in C^l(\mathbb{T}^d, \mathbb{R})$ then $|a_k| = \mathcal{O}(\|k\|^{-l})$, so that there is a constant $c_l > 0$, independent of k, such that

$$|a_k| \le \frac{c_l}{\|k\|^l},$$

for each $k \in \mathbb{Z}_0^d$ [53, Thm 3.3.9]. Thus,

$$\begin{aligned}\left|WB_T(h)(\theta_0) - \int_M h d\theta\right| &\le c_l(2\pi T)^{-m}\sum_{k\in\mathbb{Z}_0^d}\|k\|^{-l}|k\cdot\omega|^{-m}\|g^{(m)}\|_1 \\ &= C^* T^{-m}\sum_{k\in\mathbb{Z}_0^d}\|k\|^{-l}|k\cdot\omega|^{-m},\end{aligned}$$

where we defined $C^* = c_l(2\pi)^{-m}\|g^{(m)}\|_1$, which depends on m. The theorem will follow provided we can show that $\sum_{k\in\mathbb{Z}_0^d}\|k\|^{-l}|k\cdot\omega|^{-m}$ is bounded. Since $\omega \in \mathcal{D}_{c,\tau}$ (13) it follows that

$$\sum_{k\in\mathbb{Z}_0^d}\|k\|^{-l}|k\cdot\omega|^{-m} < c^{-m}\sum_{k\in\mathbb{Z}_0^d}\|k\|^{m\tau - l}.$$

Finally, noting that whenever $l - m\tau > d$, the sum above converges, we have the desired bound. That is, the weighted Birkhoff average converges as T^{-m} provided that $h \in C^l$ with $l > d + m\tau$. □

Remark 1. The weight function, $g(t) \propto \sin^2(\pi t)$ on $[0, 1]$, is in $\mathcal{G}_1$ but not $\mathcal{G}_2$ since then its second derivative is not continuous on the boundary. Nevertheless, the integration-by-parts in the proof of the theorem can proceed up to $m = 2$, since the boundary terms vanish for g and g' and $g^{(2)} \in L^1$. Thus in this case when $h \in C^l$ and $l > d + 2\tau$, the convergence is as T^{-2}. So for example, in [20], $d = 2$, and the rotation vector $\omega = (\sqrt{2} - 1, 1)$ is Diophantine with $\tau = 1$. So the convergence is $m = 2$ whenever $l > 4$. They consider a case with $h \in C^\infty$ and numerically observe a bit faster convergence, as $T^{-2.5}$.

An alternative super-polynomial convergence result was obtained by [28], we review it in the Appendix. This result follows under weaker hypothesis on the flow than those in Theorem 3, however, it requires that h has a particular structure.

[1] Note that for the discrete time case, a k-dimensional rotation vector α is Diophantine if the vector $\omega = (\alpha, 1) \in \mathbb{R}^{k+1}$ is Diophantine in the sense of (13).

4. Weighted Birkhoff average as a test for chaos

The Weighted Birkhoff average can be used as a test for chaos by examining the convergence rate of $WB_T(h)(x)$ to $\langle h\rangle$ as $T \to \infty$, for a given function h. As discussed in Section 3, if an orbit is quasiperiodic with Diophantine rotation vector, the weighted Birkhoff average is super-convergent. By contrast, it is observed that the convergence rate for chaotic orbits is much slower.

Following [29,30] we estimate the convergence rate by comparing the values of (7) for successive time intervals of a fixed length. That is, an estimate for the accuracy of (7) for a time T is found by first computing $WB_T(h)(x_0)$, and then $WB_T(h)(x_T)$, the average along a second time T segment that begins at $x_T = \varphi_T(x_0)$. A comparison of these values gives an estimate of the 'error' in the time T average relative to the true average for the initial point x_0.

We consider two primary options to quantify this error. The first,

$$\mathrm{absdig}_T(h)(x_0) \equiv -\log_{10} |WB_T(h)(x_0) - WB_T(h)(x_T)|, \tag{14}$$

we call the *absolute digit accuracy*. This is the measure proposed in [29,30], where it was denoted simply by dig_T. Eq. (14) is the number of decimal digits that two segments of the average have in common. It should be useful when the expected averages are of the same magnitude. A second error quantification is the *relative digit accuracy*,

$$\mathrm{reldig}_T(h)(x_0) \equiv -\log_{10} \frac{|WB_T(h)(x_0) - WB_T(h)(x_T)|}{\frac{1}{2}(|WB_T(h)(x_0)| + |WB_T(h)(x_T)|)}. \tag{15}$$

This measures the number of digits relative to an expected value, estimated as the mean absolute value of the two partial averages, hence the factor of $\frac{1}{2}$ in the denominator. This should be useful when the average of h varies widely in magnitude as x_0 varies. Note, however, that it is not a good quantification if the average is expected to be zero, in which case (14) is more appropriate. Finally, we define the *maximum digit accuracy*

$$\mathrm{dig}_T \equiv \max\{\mathrm{absdig}_T, \mathrm{reldig}_T\}. \tag{16}$$

By comparing the values of these measures for different initial conditions, we can differentiate between orbits for which the averages converge more rapidly than others and hence, classify which orbits are chaotic and which are regular.

5. Applications

In this section we apply the method outlined above to three example flows: the two-wave model [54], a quasiperiodically forced vector field with a strange attractor [32], and a model for magnetic fields investigated in [36].

To numerically integrate these examples, we use the algorithm Vern9, in the Julia package DifferentialEquation.jl [55, 56]. Each integration is computed in double precision, choosing absolute and relative tolerances between 10^{-10} and 10^{-15}, depending on the differential equation. The WBA in (7) was calculated by numerically solving the differential equation

$$\frac{d}{dt}W(t) = g\left(\tfrac{t}{T}\right) h(\varphi_t(x_0)), \tag{17}$$

where g is the bump function (9). Integration of (17) is simply done by adding it to the set of differential equations defining the flow φ_t. Thus (17) is integrated with the same algorithm as the trajectory. Finally we compute the time-T weighted Birkhoff average as $\frac{1}{T}W(T)$. This allows us to adjust the accuracy of the integrator so that the computation of (7) has an accuracy comparable to the trajectory itself; indeed, the Vern9 algorithm will adjust its step-size to attempt to attain the requested accuracy. All computations were run on a laptop with an Intel i7 processor running at 2.8 GHz using 32 GiB of RAM, running Linux version 5.19. Execution time for each produced data set was less than 12 h.

5.1. Two-wave model

Our first example corresponds to the 1D motion of a charged particle in the electric field two longitudinal electrostatic waves [54]. This *two-wave model* was used in [17,31] to illustrate the so-called *converse KAM* method that detects the breakup of tori. Since the destruction of tori is a signal of the onset of chaos, we can use this model to compare the efficiency of the weighted Birkhoff and converse KAM methods as chaos detectors.

Following [54], the two-wave system has the nonautonomous Hamiltonian

$$H(q, p, t) = \tfrac{1}{2}p^2 - \mu\cos(2\pi q) - \mu\cos(2\pi(q - t)), \tag{18}$$

for the position q and momentum p of the particle. Here, without loss of generality, we choose the mass of the particle to be one, the phase velocities of the two waves to be zero and one, respectively, and the wavenumber of the first wave to be one. For simplicity, we follow [17,31] to assume that the wavenumber of the second wave is also one and that the two waves have the same amplitude, μ. Thus this simplified model has only one parameter. By taking q and $t \mod 1$, the extended phase space can be thought of as $(q, p, t) \in \mathbb{T} \times \mathbb{R} \times \mathbb{T}$.

A Poincaré section at $t = 0 \mod 1$ is shown in Fig. 1 for $\mu = 0.03$. When $\mu \ll 1$, most orbits lie on rotational, invariant 2D tori (i.e., tori that are homotopic to the set $p = 0$). For small positive μ, there are two primary elliptic periodic orbits that cross the Poincaré section near $(0, 0)$ and $(0, 1)$ and two hyperbolic periodic orbits crossing the section near $(\frac{1}{2}, 0), (\frac{1}{2}, 1)$. Each of these four orbits correspond to fixed points of the Poincaré map. The section for the range $p \in [-0.2, 0.5]$ shown in Fig. 1 shows only orbits trapped in the stationary wave, near $p = 0$. The 2D tori encircling the primary elliptic orbits are *librational*; they correspond to particles trapped in one of the two electrostatic waves. Also shown in Fig. 1 are other resonant islands; these correspond to orbits trapped near elliptic periodic orbits with rational winding numbers on $\mathbb{T}^2$. The largest seen in the figure is a pair of islands surrounding a period-two orbit near $p = 0.5$ on the section.

5.1.1. Distinguishing regular and chaotic orbits

To demonstrate the difference in convergence of the WBA between chaotic and regular orbits we consider the weighted Birkhoff average of the function $h(q, p, t) = p$. Note that the average of this function for any quasiperiodic orbit will be ρ, the rotation number of the orbit.

$$\rho = \lim_{T\to\infty} \frac{q(T) - q(0)}{T} = \lim_{T\to\infty} \frac{1}{T}\int_0^T p(\tau)d\tau = \langle p\rangle, \tag{19}$$

taking the lift of the coordinate q to $\mathbb{R}$.

Fig. 2 shows the maximum digit accuracy (16) as a function of T for two initial conditions, $(q_0, p_0) = (0, 0.45)$ and $(0, 0.3)$. As can be seen in Fig. 1, the first orbit lies on a librational torus in a period-two island, while the second appears to be chaotic. For the regular orbit, the WBA appears to converge to ten digits by $T = 1000$, and dig_T indicates double precision accuracy of $\langle p\rangle$ by $T = 2000$. Since the lower bound of dig_T increases linearly with T, the convergence appears to be exponential. By contrast, dig_T fluctuates around 2 for the second, chaotic orbit. A similar dichotomy was also seen for maps in [29].

To reinforce the distinction between convergence rates for chaotic and regular orbits, Fig. 3 shows a heat map of the maximum digit accuracy for $T = 1000$ for 501 equally spaced initial

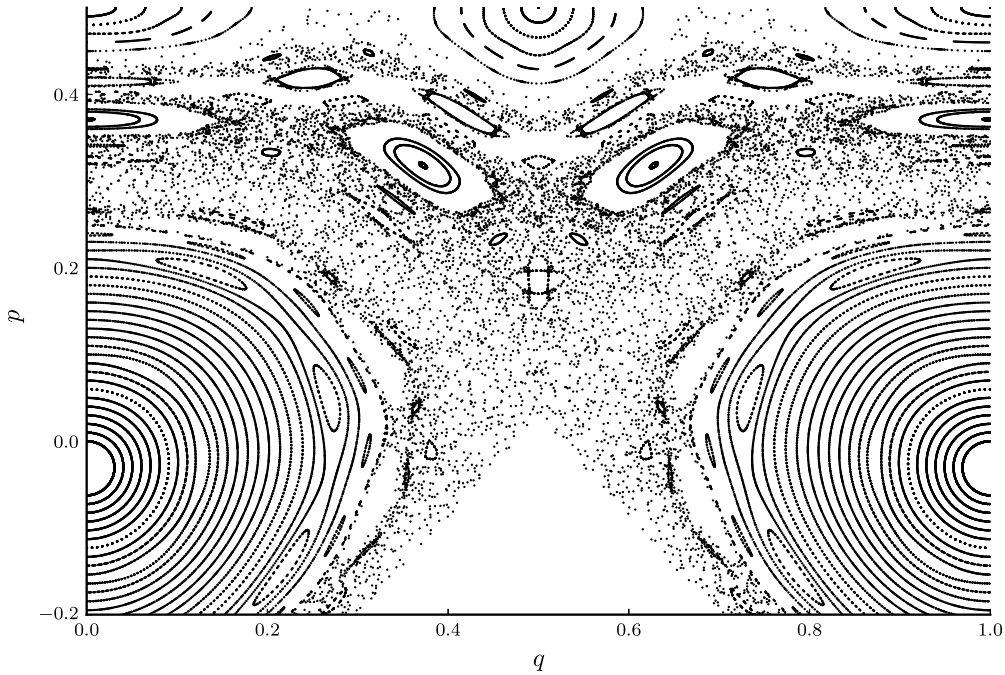


Fig. 1. Poincaré section for (18) at $t = 0 \mod 1$ $\mu = 0.03$.

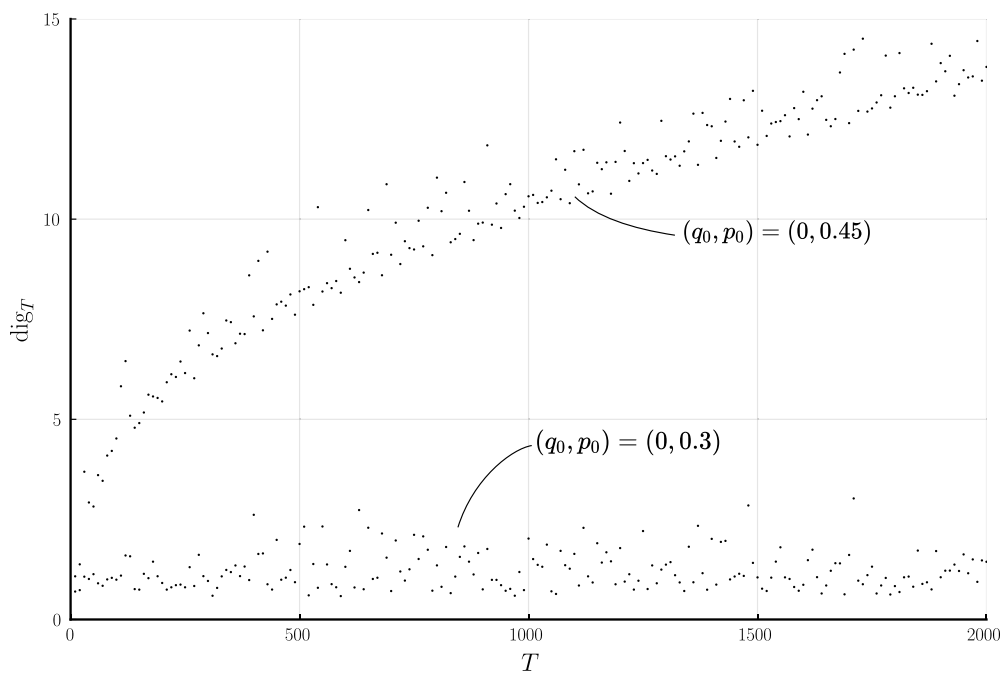


Fig. 2. The maximum digit accuracy dig_T vs T using the function $h(q, p, t) = p$ for two orbits of (18) with $\mu = 0.03$. The orbit with initial condition $(q, p) = (0, 0.45)$ is regular and that with initial condition $(q, p) = (0, 0.3)$ is chaotic.

conditions $(0, p_0, 0)$, $0 \le p_0 \le 0.5$ on the Poincaré section of Fig. 1. Note that the regular orbits in the low-period islands have $\mathrm{dig}_T \gtrsim 10$, while the strongly chaotic orbits outside these islands have $\mathrm{dig}_T \lesssim 3$. Hence, the colors show that there is a clear distinction between the regular and chaotic orbits in Fig. 1.

5.1.2. Comparisons of relative and absolute accuracy

Fig. 4 shows the results of computations of the criteria (14) and (15) for the same set of initial conditions as Fig. 3. Panel (a) shows that reldig_T performs poorly for the regular, librational tori around $(0, 0)$. This is expected since $\langle p \rangle = 0$ for these orbits, so that the denominator of (15) is near zero. However, reldig_T nears machine precision for the orbits that are trapped in the period-two island chain, near $p = 0.5$.

Panel(b) shows that absdig_T clearly distinguishes between the regular, island-trapped orbits and the chaotic orbits that were seen in Fig. 1. The initial momenta corresponding to chaotic orbits have $\mathrm{absdig}_T \lesssim 3$, while the regular orbits have $\mathrm{absdig}_T \gtrsim 7$, and there are only five orbits with $4 < \mathrm{dig}_T < 6$. A plot of dig_T, not shown, would be identical to panel (b): for this case $\mathrm{dig}_T = \mathrm{absdig}_T$ for all orbits.

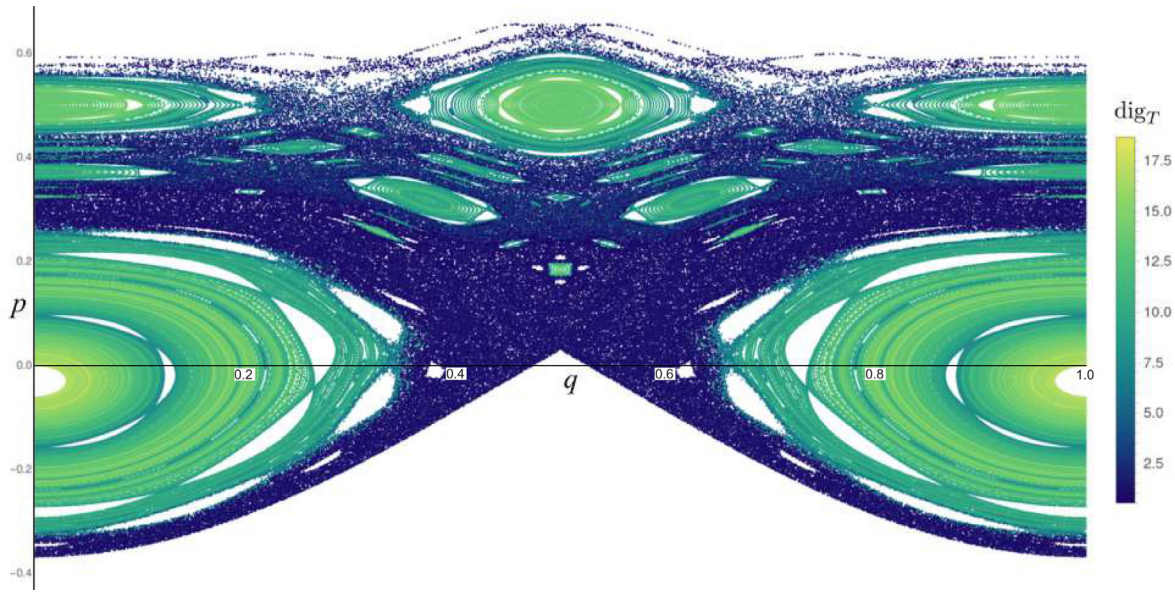


Fig. 3. A Poincaré section of (18) for $\mu = 0.03$. The 500 orbits have initial conditions $(0, p_0, 0)$, $p_0 \in [0, 0.5]$. The colors, as shown in the color bar, represent dig_T for $T = 1000$ and $h = p$. The white regions correspond to points that are not reached by the 500 orbits shown.

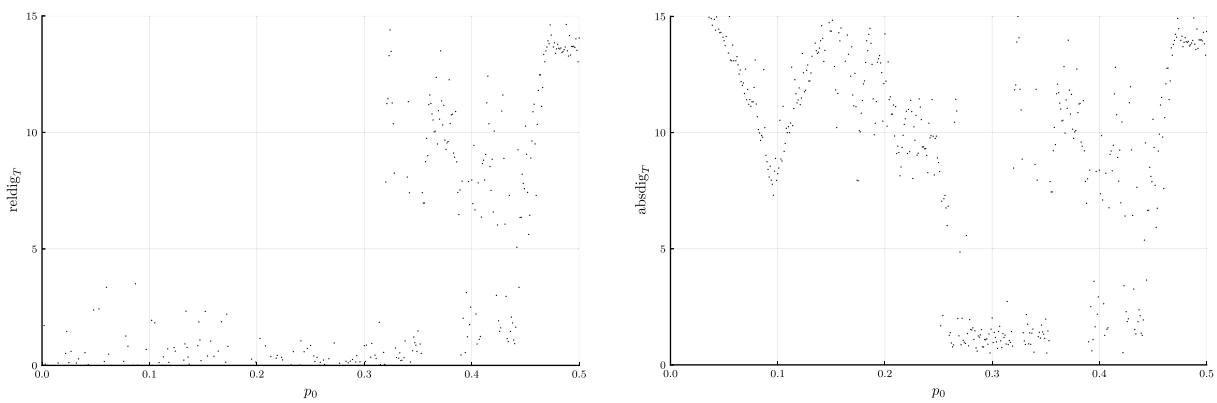


Fig. 4. Relative and absolute digit accuracy as a function of p_0 for initial conditions along the line $q_0 = 0$. In each case $T = 1000$ and $h = p$.

For the two-wave model, the results in Fig. 4 indicate a threshold for distinguishing chaotic and regular orbits

$$\mathrm{dig}_{1000} < 5, \qquad \text{(chaos criterion)}, \tag{20}$$

for the function $h = p$.

A curiosity of Fig. 4(b) is the decrease in absdig_T near $p_0 = 0.1$, even though most orbits trapped in the period-one island appear to be regular. As can be seen in Fig. 3, this dip corresponds to the region near a hyperbolic period-six orbit on the section that starts near $(0, 0.1)$. These regular orbits lie just outside the separatrix of this island chain and thus take a long time to complete a full rotation—indeed this time would go to infinity for the integrable case as the orbit nears the separatrix. Moreover, the librational rotation number of these orbits around the elliptic point approaches the rational that corresponds to that of the hyperbolic orbit. Consequently such orbits will explore a smaller fraction of an invariant torus over some finite time than those with initial condition further away. The result is a less-accurate, finite-time approximation of the space average of h.

Since there are typically arbitrarily thin chaotic layers between any pair of KAM tori in generic Hamiltonian systems, it can be very difficult to distinguish orbits in such layers from those that are regular. For example, for the two-wave model shown in Fig. 1, there is a chaotic layer near the separatrix of the period seven island chain that intersects the Poincare section near $(0, 0.2137)$. This layer has thickness at most 2×10^{-3}. The WBA of the function $h = p$ for the initial point $(0.001, 0.2138)$ has $\mathrm{dig}_T \approx 7$ for times up to 5000. By the criterion we use, this orbit would not be classified as chaotic. Nevertheless, this value of dig_T is still distinct from that of the invariant circle through the point $(0.001, 0.2134)$, which has $\mathrm{dig}_{5000} \approx 13$.

The computed rotation number (19) as a function of p_0 is shown in Fig. 5. In the region of librational tori near $(0, 0)$, $\rho = 0$. Since most of these orbits are regular ρ is computed with high accuracy. Indeed, as we saw in Fig. 4, $\mathrm{absdig}_T > 5$ for all $p_0 \in [0, 0.251]$. The rapid fluctuations in ρ as a function of initial condition near $p_0 = 0.3$ reflect the poor convergence of the WBA for these chaotic orbits. Additional regular regions appear for higher period islands around elliptic periodic orbits; these have constant, rational rotation number. The chaotic regions between pairs of neighboring islands give the scattered values between the flat intervals.

5.1.3. *Varying the width of the bump function*

Another choice worth investigating is that of the weight function $g : \mathbb{R} \to [0, \infty)$ in (7). Theorem 3 implies that superconvergence follows whenever g is C^∞ and flat at 0 and 1. In

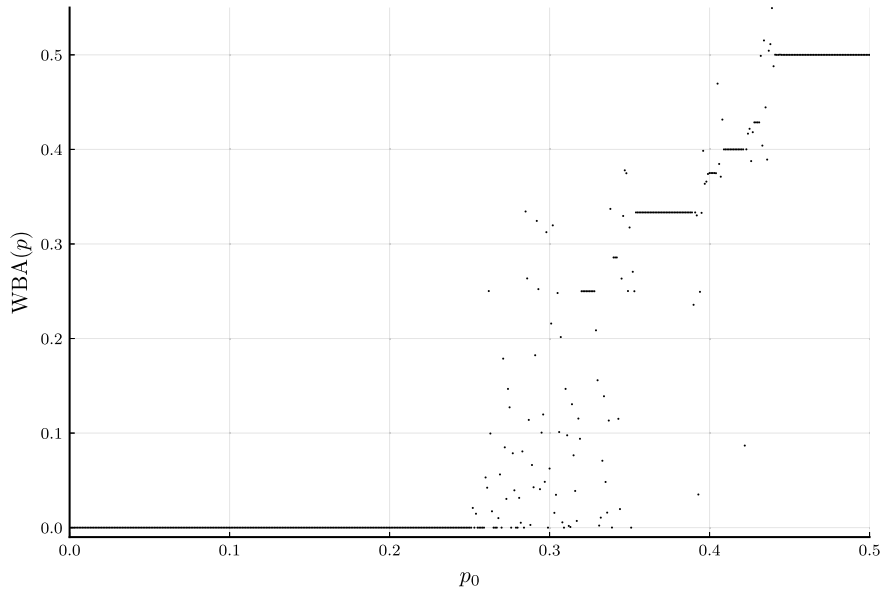


Fig. 5. The computed rotation number $\rho \approx WB_T(p)$ for initial conditions $(0, p_0, 0)$ as a function of p_0. As above this is for (18) with $\mu = 0.03$.

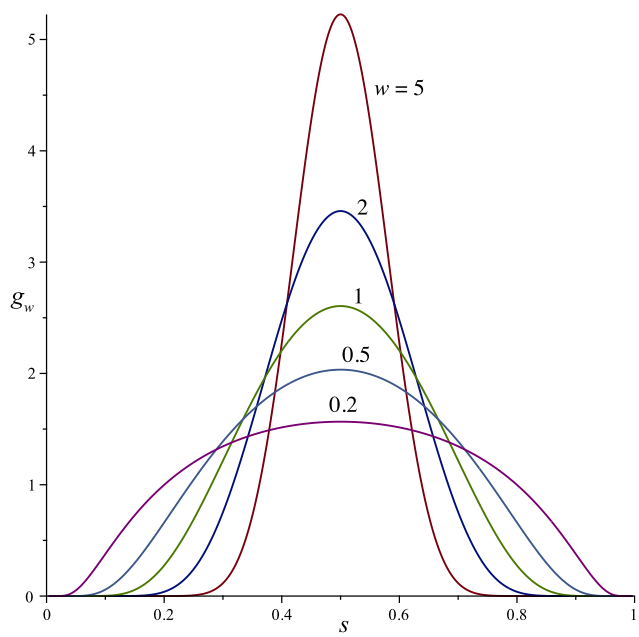


Fig. 6. The weighted bump function (21) for five values of w.

this section we continue to use (9), but vary it slightly by adding a width parameter, $w > 0$:

$$g_w(s) = \begin{cases} C \exp\left(\frac{-w}{s(1-s)}\right), & s \in (0, 1) \\ 0, & s \le 0 \text{ or } s \ge 1. \end{cases} \tag{21}$$

Again, C is chosen so that g_w has the normalization (6). The resulting function is shown for several values of w in Fig. 6. If $w \ll 1$ then g_w is near its maximum over a large fraction of [0, 1], and the average limits to the unweighted, time-T average. If $w \gg 1$ then g_w is essentially zero except for a small interval. Neither of these cases would seem to be desirable. But what intermediate value of w is best?

Fig. 7, shows how dig_T depends on w for two different T, for a regular trajectory of the two-wave model. Interestingly, these curves have local spikes indicating improved convergence for nearly isolated values of w. Since, as seen in the figure, the value of w for these spikes changes with T, and calculations (not shown) indicate that they also vary as the orbit is varied, consequently it is hard to argue that any specific choice for w would be optimal. Moreover, $w = 1$ seems to be a reasonable choice, since—if we ignore the spikes—neither large or small w offers a uniform improvement

5.2. A quasiperiodically forced system

As Theorem 3 showed, the weighted Birkhoff average is superconvergent for a quasiperiodic orbit with a Diophantine rotation vector. When the dynamics is a conjugate to rigid rotation, then there is an invariant measure on the torus. More generally, if there is no invariant measure, then the Birkhoff ergodic theorem does not apply. In this case it is not clear whether the accuracy of the weighted Birkhoff average can distinguish between regular and non-regular orbits.

In this subsection we study the quasiperiodically forced and damped pendulum model of [32,57]:

$$\begin{aligned} \dot\theta &= p, \\ \dot\psi_1 &= \gamma, \\ \dot\psi_2 &= 1, \\ \dot p &= -\nu p + a\cos(2\pi\theta) + b + c(\cos(2\pi\psi_1) + \cos(2\pi\psi_2)). \end{aligned} \tag{22}$$

Here $\gamma \in \mathbb{R} \setminus \mathbb{Q}$ is irrational and $a, b, c, \nu \in \mathbb{R}$ are parameters. Grebogi et al. [57] observed that this system can have a geometrically strange attractor with no positive Lyapunov exponents, a situation that they call a *strange, nonchaotic attractor*. Even though such a system may be thought of as nonchaotic because nearby orbits do not separate exponentially, the dynamics may still exhibit the weaker, topological form of sensitive dependence [33].

Formally the phase space for (22) is $\mathbb{T}^3 \times \mathbb{R}$, with coordinates $(\theta, \psi_1, \psi_2, p)$. A natural 3D Poincaré section is $\psi_2 = 0 \mod 1$. Following [32] we take

$$\nu = a = 6\pi, \quad c = 0.55\nu, \quad \gamma = \tfrac{1}{2}(-1 + \sqrt{5}), \tag{23}$$

leaving one free parameter, b.[2]

[2] In [32] the parameters are $\nu = a = 2\pi p$, $b = 2\pi Kp$, and $c = 2\pi Vp$, where p is a damping parameter. The case $p = 3$, $K = 1.33$, and $V = 0.55$ of [32, Fig. 5] corresponds to (23) with $b = 1.33\nu$.

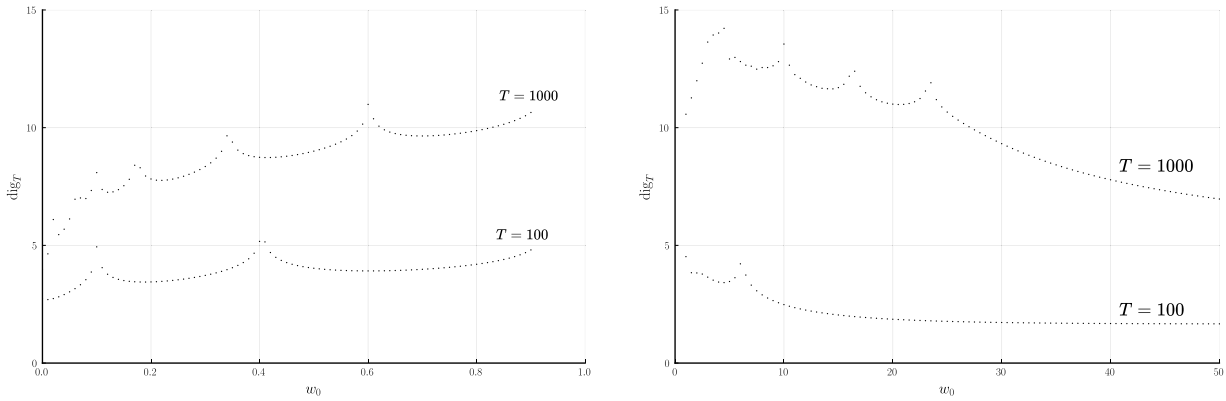


Fig. 7. The variation of $\mathrm{dig}_T(h)(z)$ with w, the width parameter of (21). The system is (18) with $\mu = 0.03$, $h = p$, and initial condition $(q_0, p_0) = (0, 0.1)$. (a) $w \in [0, 1]$ and (b) $w \in [0.5, 50]$.

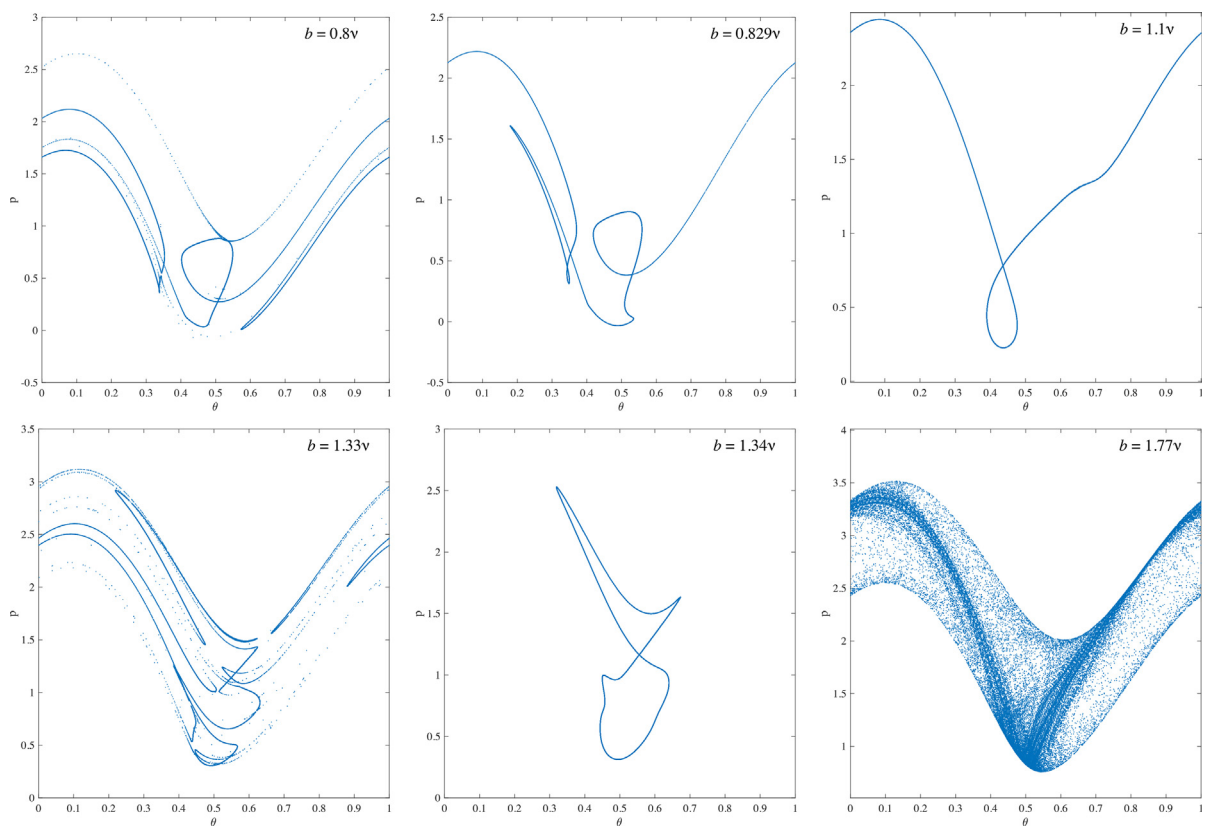


Fig. 8. Projections onto (θ, p) of 5×10^4 points on the 3D Poincaré section $\psi_2 = 0 \mod 1$ of the system (22) using (23). The six panels show $b = [0.8, 0.829, 1.1, 1.33, 1.34, 1.77]\nu$, respectively. The maximal Lyapunov exponent in the (q, p) subspace is negative, except for the last case, where attractor appears to be 3D on the Poincaré section.

Six examples of attractors for (22) are shown in Fig. 8. These figures are projections onto (θ, p) of the Poincaré section $\psi_2 = 0 \mod 1$. On the 3D section the attractor is sometimes a two-torus, sometimes geometrically strange with dimension between two and three, and sometimes fully 3D. For example, when $b = 1.33\nu$ the system has a geometrically strange attractor with a box-counting dimension larger than 1, but no positive Lyapunov exponents: in the (θ, p) subspace the maximal exponent is $\lambda = -0.45$ [32]. The maximal Lyapunov exponent reaches 0 at about $b = 1.66\nu$ where the attractor in the Poincaré section appears to be 3D, see the final panel of Fig. 8. The attractor collapses back to a two-torus as b nears 1.8ν (not shown).

5.2.1. Distinguishing strange attractors

The different geometric structures shown in Fig. 8 make this system a prime candidate for investigating whether weighted Birkhoff averaging can be used to distinguish between regular and strange nonchaotic attractors. Fig. 9 shows dig_T as a function of T for values of b that correspond two-torus, strange, and 3D attractors, respectively.

When $b = 1.1\nu$, the WBA appears to be super-convergent. The maximum digit accuracy reaches 13 by $T = 1200$ and then remains nearly constant; this is consistent with the accuracy of the numerical integration, which was set to 10^{-13} for both absolute and relative error. This suggests that the dynamics of this orbit are conjugate to a Diophantine rigid rotation. When

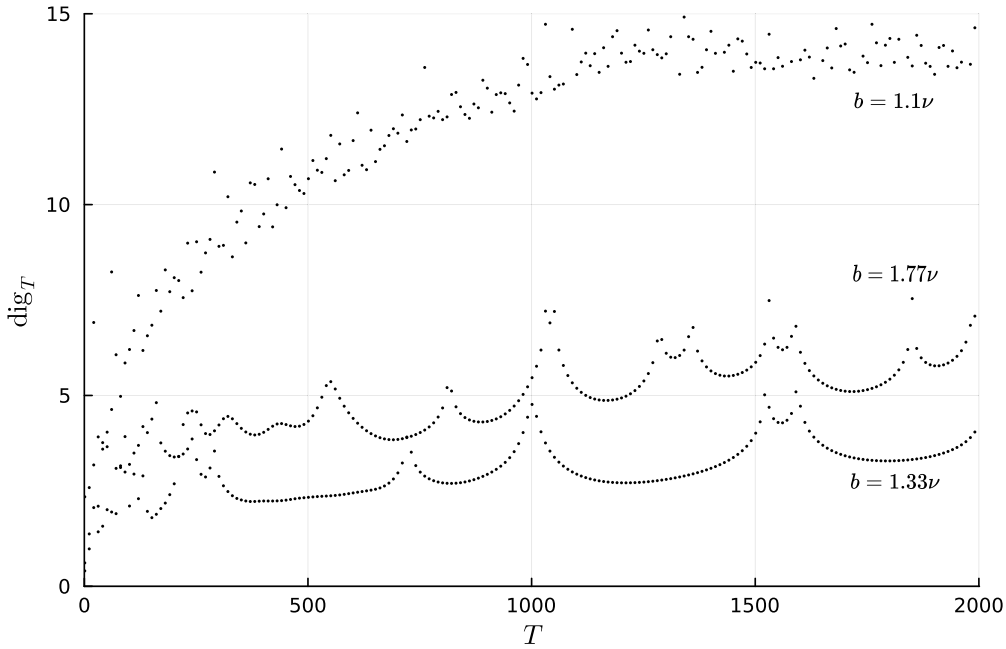


Fig. 9. The digit accuracy for the system (22) with parameters (23) and $b = 1.1\nu$, 1.33ν, and 1.77ν as a function of T using $h = p$. The initial condition is $(\theta, \psi_1, \psi_2, p) = (0, 0, 0, 2)$.

$b = 1.33\nu$, where the attractor is strange but nonchaotic, the convergence of the WBA in Fig. 9 is observed to be poor: it only reaches 3 by $T = 1200$. The convergence is also poor when $b = 1.77\nu$, where the attractor is 3D. Even though dig_T is larger than the previous case, it only reaches 5 when $T = 1200$, and in both cases the WBA appears to converge—at best—at a polynomial rate in T. Even if this attractor is simply a three-torus, the relatively poor convergence suggests that its dynamics are more complex than rigid rotation. Thus the weighted Birkhoff average effectively distinguishes between a two-torus attractor, and more complex or higher dimensional attractors.

5.2.2. Finding Two-Tori

We now look at how the accuracy of the WBA varies with b in order to distinguish between two-torus and strange or 3D attractors. Fig. 10 shows dig_T as a function of $b \in [0.6, 1.8]\nu$. The figure shows a clear stratification into three levels; the highest corresponds to $13 < \mathrm{dig}_T < 18$. This high accuracy occurs, for example, for the case $b = 1.1\nu$ and 1.34ν shown in Fig. 8 that are clearly two-tori; these are the blue points in Fig. 10. The highest accuracy, $\mathrm{dig}_T \sim 17$, occurs near $b = 0.6\nu$. The attractors in this case (not shown) are even simpler: they resemble librating orbits of the pendulum in (θ, p) that are simply extended in the ψ_1 direction.

The lowest level in Fig. 10 are those b values with $\mathrm{dig}_T \sim 4$. The two red points correspond to the values $b = 0.8\nu$, and 1.33ν, the strange attractors shown in Fig. 8.

The mid-level range, $\mathrm{dig}_T \sim 8$, for Fig. 10 corresponds to geometrically more complex attractors that are nevertheless, not strange. For example, the green points in the figure, represent the values $b = 0.829\nu$ and 1.77ν shown in Fig. 8. The first appears to be the projection of a two-torus, however, it is geometrically more complex than those tori that have higher values of dig_T. The second green point corresponds to the 3D attractor in Fig. 8. The rapid increase in dig_T as b increases beyond 1.77ν in Fig. 9 signals the collapse of this structure; by $b = 1.8\nu$, it has become a two-torus similar to that at 1.1ν though without the loop seen in Fig. 8.

The rotation number of θ:

$$\rho = \lim_{n\to\infty} \frac{\theta(n) - \theta(0)}{n},$$

is shown in Fig. 11. The figure is similar to the Devil's staircase shown in [32, Fig. 8b], however the weighted Birkhoff average for the function $h = p$ provides a much more accurate computation. The flat sections in ρ correspond to the two-torus attractors with $\mathrm{dig}_T \sim 14$, the highest level in Fig. 10. This figure shows almost no scatter when compared with the corresponding plot for two-wave model, Fig. 5.

5.3. Magnetic field line flow

As a final example we consider a family of model magnetic fields studied in [36]. Here the domain is the solid torus $D^2 \times \mathbb{S}^1$, where $(\psi, \theta) \in [0, 1] \times \mathbb{S}^1$ are polar coordinates on the disk D^2 and ζ is the toroidal angle on $\mathbb{S}^1$. The fields are generated from the vector potential $A = \psi\nabla\theta - \chi\nabla\zeta$ with

$$\chi(\psi, \theta, \zeta) = \tfrac{1}{2}\psi^2 - \sum_{m,n\in\mathbb{Z}} \varepsilon_{m,n}\psi(\psi - 1)\cos\left(2\pi(m\theta - n\zeta)\right),$$

This gives the magnetic field

$$B = \nabla \times A = \nabla\psi \times \nabla\theta - \nabla\chi(\psi, \theta, \zeta) \times \nabla\zeta. \tag{24}$$

The field line of B for this case can also be thought of as the flow of χ as a nonautonomous Hamiltonian using (θ, ψ) as canonical variables and ζ as "time". These are the solutions to

$$\begin{aligned} \dot\psi &= -2\pi \sum_{m,n} m\varepsilon_{m,n}\psi(\psi - 1)\sin(2\pi(m\theta - n\zeta)), \\ \dot\theta &= \psi - \sum_{m,n} \varepsilon_{m,n}(2\psi - 1)\cos(2\pi(m\theta - n\zeta)), \\ \dot\zeta &= 1. \end{aligned} \tag{25}$$

Note that this system has invariant two-tori at $\psi = 0$ and $\psi = 1$. Moreover, when all the amplitudes $\varepsilon_{m,n} = 0$ the system is completely integrable, since ψ is then invariant. More generally, each m, n Fourier mode creates a resonant magnetic island near $\psi = \frac{n}{m}$ with amplitude $\varepsilon_{m,n}$.

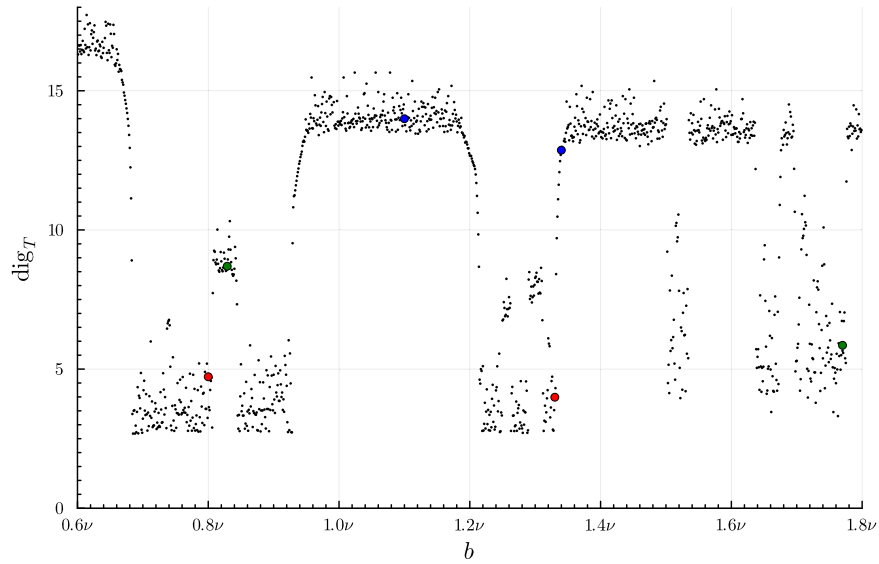


Fig. 10. The maximum digit accuracy for the system (22) with parameters (23) for a grid of 1200 values of $b \in [0.6, 1.8]\nu$. Here $T = 1500$ and $h = p$. The red points correspond to $b = 0.8\nu$ and 1.33ν, blue to $b = 1.1\nu$, 1.34ν, and green to 0.829ν and 1.77ν. The attractors for these b values were shown in Fig. 8.

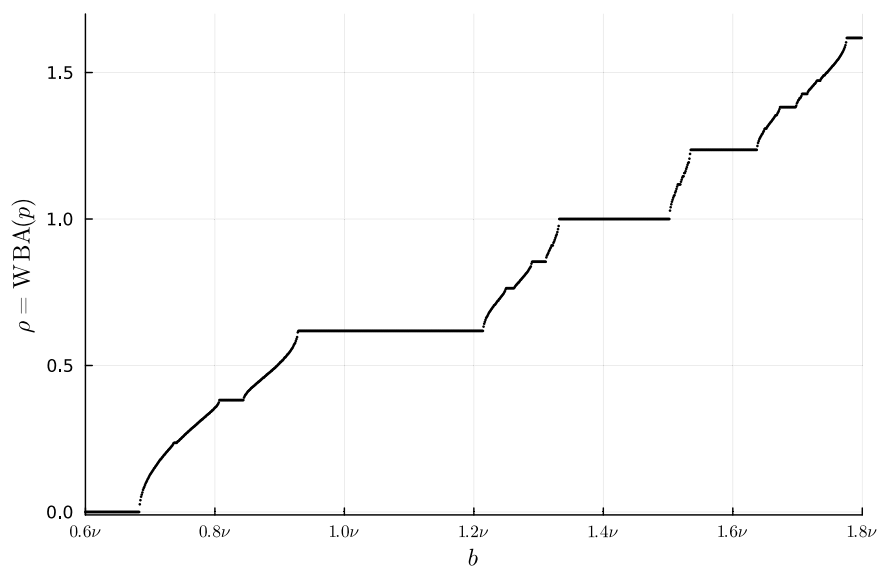


Fig. 11. The rotation number of θ for (22) for the same parameters as Fig. 10 computed using the WBA for $h = p$.

In [36], a set of resonances with fixed m and a range of n values are studied. To emphasize the creation of higher-order islands by resonant beating, we instead use a set of Fourier modes that correspond to resonances *up to* a given level on the Farey tree [58]. In particular, we take (m, n) corresponding to the resonances up to level three on the Farey tree with root $(\frac{0}{1}, \frac{1}{1})$, namely,

$$(m, n) \in \{(4, 1), (3, 1), (5, 2), (2, 1), (5, 3), (3, 2), (4, 3)\}. \tag{26}$$

Note that the Farey tree naturally generates a set of coprime (m, n) pairs.

We will choose a one-parameter family of the amplitudes so that there is a critical value at which the Chirikov overlap criterion [58] is simultaneously satisfied for each neighboring pair. The approximate island half-width in ψ for a single Fourier mode in (25) can be obtained by neglecting the $\mathcal{O}(\varepsilon)$ term in the $\dot{\theta}$ equation. Thus if $\varepsilon_{m,n} \ll 1$ the system is effectively a pendulum in the variables $(m\theta - n\zeta, m\psi - n)$. This gives a resonance at $\psi = \frac{n}{m}$ with the half-width

$$\Delta_{m,n} = 2\sqrt{\varepsilon_{m,n}\tfrac{n}{m}\left(1 - \tfrac{n}{m}\right)}.$$

Two neighboring resonant islands on the Farey tree then overlap when

$$\Delta_{m_1,n_1} + \Delta_{m_2,n_2} = \left|\frac{n_1}{m_1} - \frac{n_2}{m_2}\right| = \frac{1}{m_1 m_2}.$$

Here the last equality above follows because the two modes are Farey neighbors. If we scale the amplitudes as

$$(\varepsilon_{4,1},\ \varepsilon_{3,1},\ \varepsilon_{5,2},\ \varepsilon_{2,1},\ \varepsilon_{5,3},\ \varepsilon_{3,2},\ \varepsilon_{4,3}) = \frac{\varepsilon}{21600}\,(72, 27, 25, 96, 25, 27, 72) \tag{27}$$

then the resonances simultaneously overlap at $\varepsilon = 1$. For this value the system should be chaotic, in the sense that the rotational tori between each pair of islands are destroyed. Of course, as is well known, the overlap criterion overestimates the *critical* value for the destruction of the KAM tori [6].

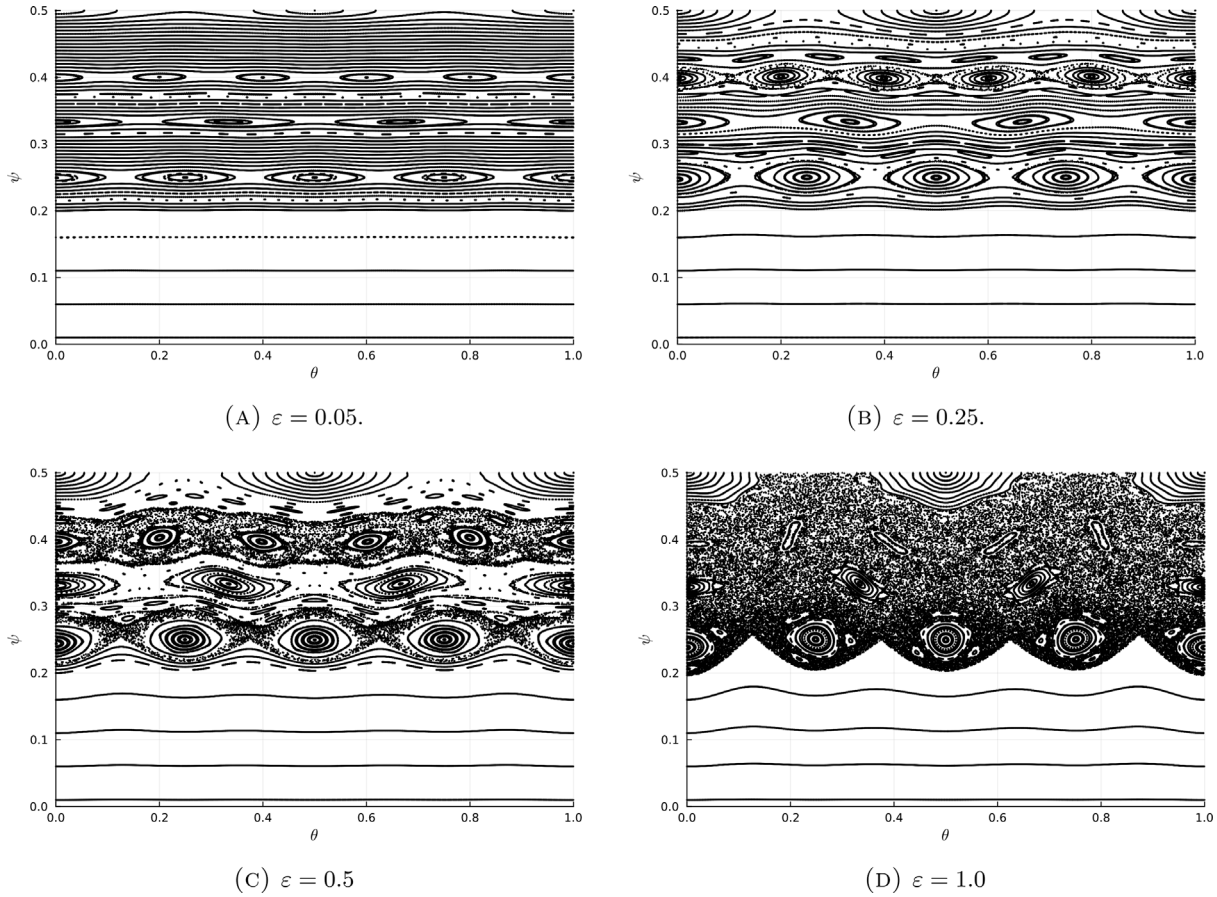


Fig. 12. Poincaré sections for (25) with parameters (27) at $\zeta = 0 \mod 1$ for four values of ε.

5.3.1. Detecting chaos

A Poincaré section at $\zeta = 0$ for the system (25) is shown in Fig. 12 for four values of ε. This figure shows only the range $\psi \in [0, \frac{1}{2}]$, but for the symmetric amplitudes (27), $\varepsilon_{m,n} = \varepsilon_{m,m-n}$, the phase portrait has the reflection symmetry $\psi \to 1 - \psi$. Thus the dynamics in the interval $\psi \in [\frac{1}{2}, 1]$ can be inferred.

When $\varepsilon = 0.05$, Fig. 12(A) shows that almost all orbits lie on tori, though there will invariably be small chaotic regions not observed at this scale near the separatrices of the island chains. When $\varepsilon = 0.25$, a small amount of chaos is visible near the separatrices of the period four and five islands. When $\varepsilon = 0.5$ these chaotic regions grow; however, there are still rotational tori that act as barriers to transport between each of the primary island chains in the set (26). For $\varepsilon = 1.0$, Fig. 12(D) shows that all of the rotational tori for ψ in the interval $(0.2, 0.8)$ have been destroyed, though some low-period islands persist in the sea of chaos.

The true critical value of ε can be estimated numerically by looking for an orbit that "crosses" all the resonances. Starting at $(\theta_0, \psi_0, \zeta_0) = (0.375, 0.27, 0)$, close to the hyperbolic-point of the $(4, 1)$ island, we found that the smallest ε for which $\psi(t) > 0.45$ for some time $t \in [0, 10^4]$ is $\varepsilon_{cr} = 0.665$. This is certainly consistent with the phase portraits in Fig. 12.

Note that the regions of regular tori around $\psi = 0$ and around $\psi = 1$ persist as ε grows. This is because the tori $\psi = 0$ and $\psi = 1$ are invariant, and the Farey island set (26) does not include any terms below $\frac{1}{4}$ and above $\frac{3}{4}$.

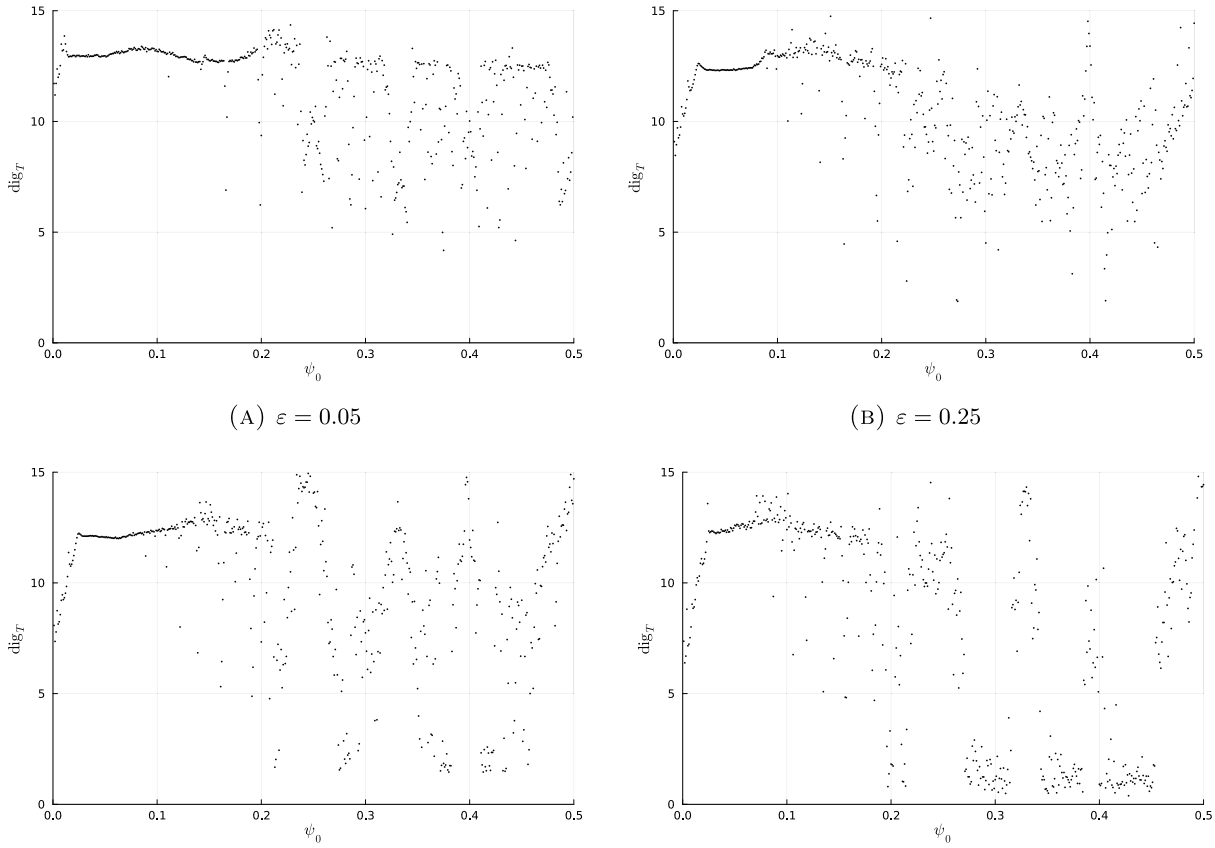
To study the onset of chaos using the weighted Birkhoff average, we computed the maximum digit accuracy, dig_T, for initial conditions $(\psi_0, 0, 0)$ with $\psi_0 \in [0, 0.5]$. We choose $h = \psi$ so that $\langle \psi \rangle$ is a first approximation of the rotation number of a regular torus. The results for the same four values of ε in Fig. 12 are shown in Fig. 13.

As can be seen in the figure, the distinction between dig_T for chaotic and regular orbits becomes clearer as ε increases. The criterion (20) suggests that chaotic orbits have $\mathrm{dig}_T \lesssim 5$. Note that some of the regular orbits have dig_T up to 15, and that the nearly horizontal, rotational tori in the range $0 < \psi_0 < 0.2$ have $\mathrm{dig}_T \simeq 13$. When $\varepsilon = 0.05$, there are only 4 initial conditions on the 501 point grid that would be designated as chaotic; though since $4 \le \mathrm{dig}_T \le 5$ for each of these, they are near the threshold. As ε increases, chaotic regions surrounding the islands appear and grow. By $\varepsilon = 0.25$, there are small intervals of low-digit accuracy around the $\frac{1}{4}$ and $\frac{2}{5}$ islands; however, only 14 initial conditions have $\mathrm{dig}_T \le 5$. As ε increases, the chaotic regions around the low-period islands grow, and for $\varepsilon = 0.5$, initial conditions with low dig_T are seen near $\psi_0 = 0.2$, 0.3, 0.4 and 0.5. Finally, when $\varepsilon = 1.0$ the chaotic regions around the low period islands merge, though the line of initial conditions with $\theta_0 = \zeta_0 = 0$ goes through the elliptic center of each of the forced resonances in the set (26). This gives rise to the peaks in dig_T near the low-order islands that are visible in Fig. 12(D). Of course, the orbits for $\psi_0 < 0.18$, where regular tori seen in Fig. 12(D) persist, have a maximum digit accuracy that remains high.

To conclude, it is clear by comparing Fig. 12 and Fig. 13, that the weighted Birkhoff average accurately detects the onset of chaos for this family of Farey magnetic fields.

5.3.2. A measure of non-integrability

In [36], the authors developed a measure of the "effective volume of parallel diffusion" as a proxy for measuring the non-integrable region. To do this, they first solve the steady-state



(A) $\varepsilon = 0.05$ (B) $\varepsilon = 0.25$ (C) $\varepsilon = 0.5$ (D) $\varepsilon = 1.0$

Fig. 13. The maximum digit accuracy, $\mathrm{dig}_T(h)$, as a function of initial condition ψ_0 on a grid with steps of 0.001, for four ε values. Here $T = 1000$ and $h(\psi, \theta, \zeta) = \psi$.

temperature using the anisotropic diffusion equation,

$$\nabla \cdot (\kappa_\parallel \nabla_\parallel T + \kappa_\perp \nabla_\perp T) = 0. \tag{28}$$

Here T is the temperature, $\kappa_\parallel, \kappa_\perp$ are the parallel, perpendicular diffusion coefficients, respectively, and $\nabla_\parallel = \hat{b}\hat{b} \cdot \nabla$ and $\nabla_\perp$ are the gradients parallel and perpendicular to the unit vector $\hat{b} = B/\|B\|$, respectively. Eq. (28) is solved with the Dirichlet boundary conditions that fix T on the boundary tori, $\psi = 0$ and 1. The measure they use is the fraction of the volume Ω in which the local parallel heat transport is larger than the perpendicular transport:

$$\mathcal{V}_{PD} = \frac{1}{\mathrm{Vol}(\Omega)} \int_\Omega \Theta\left(\kappa_\parallel |\nabla_\parallel T|^2 - \kappa_\perp |\nabla_\perp T|^2\right) dx^3. \tag{29}$$

Here Θ is the Heaviside step function.

In order that (29) be an effective measure of integrability, Paul et al. [36] argue that when $\kappa_\perp \ll \kappa_\parallel$, T is approximately constant along field lines of B. It follows, for a region foliated by invariant tori, $|\nabla_\parallel T|$ will be relatively small, and thus the measure $\mathcal{V}_{PD}$ will be essentially zero. Conversely, they argue that regions of phase-space with chaotic field lines will have relatively large parallel diffusion. This second claim is shown by first proving that surfaces of constant temperature must have the same topology as the boundary surfaces. Consequently, these isotherms will not be able to completely align to the structure of the field in chaotic regions, increasing the value of $|\nabla_\parallel T|$. However, as the authors remark, the effective parallel diffusion will also be high within islands, even if they are not chaotic. Thus regions where the invariant surfaces do not have the same topology as the boundary will also contribute to (29). In this regard, the measure of parallel diffusion is analogous to converse KAM theory [17,31].

Regardless of what (29) precisely measures, such a measure may in fact be more useful for the original purpose of [36] in optimizing the structure of magnetic fields for plasma confinement. However, the weighted Birkhoff averaging may provide a reasonable alternative measure of chaos, if that is desired.

A simple measure of integrability is the relative fraction of initial conditions deemed chaotic by weighted Birkhoff averaging. For the model (25), we first used the same initial conditions as Fig. 13: $(\psi_0, \theta_0, \zeta_0) = (\psi_0, 0, 0)$. For each $\varepsilon \in [0, 1.0]$ in steps of 0.01, we computed dig_T for $T = 1000$ and $h = \psi$. An orbit was deemed chaotic if $\mathrm{dig}_T < 5$ and regular otherwise. The relative fraction of chaotic initial conditions for each ε is shown using the symbol '+' in Fig. 14. A similar computation was performed along the line $\theta_0 = 0.15$. The result points are shown in Fig. 14 using the symbol •.

For both sets of initial conditions, the fraction of chaotic initial conditions is observed to vanish when $\varepsilon = 0$, and increase—though not always monotonically—with ε. This is consistent with Fig. 12. For small ε (e.g., $\varepsilon < 0.2$) we see the fraction of chaotic initial conditions is very small. This is consistent with the expected exponentially small splitting of separatrices. For larger ε, the rate of increase of the chaotic fraction slows. This is most prominent for $\theta_0 = 0.15$ near $\varepsilon = 0.6$, just below $\varepsilon_{cr} = 0.665$ where the last rotational tori in the interval $\psi_0 \in [0.18, 0.5]$ are destroyed.

An issue with the use of the fraction of chaotic orbits along a line of initial conditions in the 3D phase space is evident in the difference between the two cases in Fig. 12. The line $\theta_0 = 0$ is special because it intersects the rotational periodic orbits at their elliptic points; this is evidently due to a time-reversal symmetry of the system (25) under the involution

$$(\psi, \theta, \zeta) \to (\psi, -\theta, -\zeta).$$

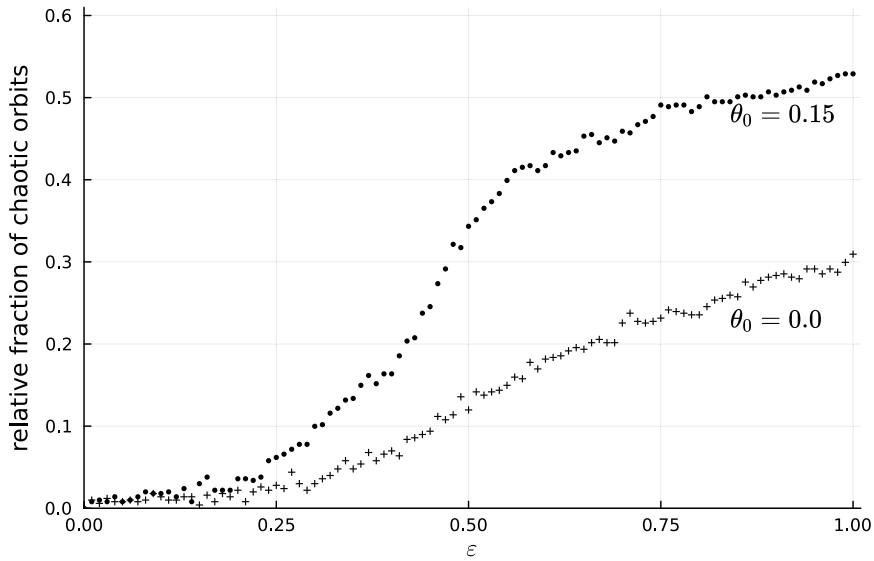


Fig. 14. The fraction of initial conditions that are chaotic for (25) with $\varepsilon \in [0, 1.0]$, and initial conditions $(\psi_0, \theta_0, 0)$ with $\psi_0 \in [0, 0.5]$ in steps of 0.001.

A similar time-reversal symmetry of the Chirikov standard map results in the so-called "dominant symmetry line" that contains all the minimax rotational periodic orbits [58]—these are the orbits that are elliptic for small perturbations. The result is that the sample of initial conditions along the line $\theta_0 = \zeta_0 = 0$ includes more regular orbits, the elliptic islands around each of these periodic orbits. In contrast, the line $\theta_0 = 0.15$ intersects fewer of the elliptic regions around the islands. Hence, the fraction of chaotic orbits for $\theta_0 = 0$ is less than that for $\theta_0 = 0.15$.

Of course, a more general system than (25), e.g., one with added phases in each Fourier term, would not have this symmetry and would thus be less susceptible to this problem. In any case, this issue could be ameliorated by sampling initial conditions on a 2D grid in the Poincaré section; of course, this would add the computational expense. One could keep track of which grid cells are visited by each orbit, so as to reduce the number that need to be considered.

6. Discussion

We have investigated the utility of the WBA for distinguishing between regular and chaotic orbits for the two-wave Hamiltonian system, a quasiperiodically forced, dissipative system that has a strange attractor with no positive Lyapunov exponents, and a model for magnetic field line flow. It was shown that the WBA is super-convergent when the dynamical system and phase space function are smooth and the dynamics is either conjugate to a rigid rotation with Diophantine rotation vector or more generally satisfies (30). The contrasting, relatively slow convergence of chaotic trajectories provides an efficient discrimination criterion. However, there remain some open questions and interesting further directions.

A first theoretical question is that of the convergence of the weighted Birkhoff average for general ergodic flows. In each application it was observed that the WBA for chaotic orbits converged relatively slowly in comparison to the regular orbits. This formed the basis for the WBA as a method to detect chaos. However, this slow convergence does not yet have a theoretical foundation. It may be possible to show that (30) is not only sufficient for super-polynomial convergence, but also necessary. If this is true, then it may provide a path forward to theoretically confirming the slow convergence for chaotic orbits observed in this paper and for the case of maps in [29,30].

One of the benefits of the WBA is that it can provide an accurate computation of the average of a phase space function. Indeed, when the average converges, one gets—for free—a good approximation to $\langle h \rangle$. Consequently, given a physically important h, such as rotation number, its value is computed as a free by-product of the method. Conversely, if the main goal is to compute an orbit average of some smooth function, then super-convergence of the WBA on regular orbits, also gives—for free—a criterion distinguishing between regular and chaotic behavior.

This poses the question: which h is optimal for chaos detection? This appears to be a difficult question. It is clear that some choices are poor, and this is supported by the convergence theorems for the Birkhoff average in [45,46]. Moreover, an everywhere constant function is also obviously a poor choice for a different reason: its average over any orbit is the same for any T. In some sense, an ideal function for distinguishing chaos would be constant on regular orbits, but vary on chaotic ones. In this case the average along the latter should still exhibit the characteristically slow convergence that we observed above. Of course, if one were able to construct such a function, then one would already know the orbital structure of the flow, obviating the necessity of a computation.

This argument suggests that an approximate integral of the system might be a good choice for h. Such a choice would ensure that h has little variation on regular orbits, still leaving room to see the distinction between convergence for regular and chaotic orbits. We obtained several such approximate integral functions for the two-wave system in [17]. However, our preliminary studies using these approximate integrals as h in the two-wave system did not appear to produce a stronger contrast in dig_T between chaotic and regular orbits. In the future, we hope to investigate the choice of h in the weighted Birkhoff average as a means of detecting chaos.

A further line of future study is that of an effective measure of (non)-integrability. This was one of the primary aims of the work in [36]. Such a measure would help in optimizing field configurations by minimize chaos. There are several improvements of the crude measure of Fig. 14 that one could implement and use to understand chaos in magnetic confinement devices.

Finally, it was evident in Section 5.2 that the WBA can distinguish between regular and strange "non-chaotic attractors". Thus, the convergence rate distinction for the WBA does not rely on exponential divergence of orbits. Future investigation is needed to understand precisely which types of dynamics this method can accurately discern.

CRediT authorship contribution statement

Nathan Duignan: Conceptualization, Software, Formal analysis, Writing, Visualization. **James D. Meiss:** Conceptualization, Validation, Formal analysis, Writing, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that support the findings of this study are available within the article.

Acknowledgments

The authors acknowledge support of the Simons Foundation through grant #601972 "Hidden Symmetries and Fusion Energy". Useful conversations with E. Sander are gratefully acknowledged.

Appendix. Kachurovskii's convergence result

Following the work of [28,46], super-polynomial convergence can be guaranteed under weaker hypothesis on the flow than those in Theorem 3, provided h has a particular structure.

Theorem 4. *Let $(M, \mathcal{B}, \mu)$ be a probability space and $\varphi_t : X \to X$be a smooth, ergodic flow with invariant probability measure μ. Suppose that $h \in L^1(X, \mu)$ and there exists bounded function $H \in C^m(X, \mathbb{R})$ such that*

$$h(x) = \langle h \rangle - \frac{d^m}{dt^m} H(\varphi_t(x))\bigg|_{t=0}, \tag{30}$$

for all $x \in S$ where S is an invariant set of full measure. Moreover, suppose the bump function $g \in \mathcal{G}_\infty$. Then, there exists a constant c_m such that,

$$|WB_T(h)(x) - \langle h \rangle| \le \frac{c_m}{T^m} \quad \text{for all } x \in S.$$

Proof. As S is invariant, the group property of φ_t guarantees, for all $x \in S,\ \tau \in \mathbb{R}$,

$$h(\varphi_\tau(x)) = \langle h \rangle - \frac{d^m}{dt^m} H(\varphi_{t+\tau}(x))\bigg|_{t=0} = \langle h \rangle - \frac{d^m}{dt^m} H(\varphi_t(x))\bigg|_{t=\tau}.$$

Using this relation we can rewrite the difference between weighted Birkhoff average and the phase space average as

$$\begin{aligned} |WB_T(h)(x) - \langle h \rangle| &= \left| \frac{1}{T} \int_0^T g\left(\frac{t}{T}\right) h(\varphi_t(x))\, dt - \langle h \rangle \right| \\ &= \left| \int_0^1 g(s) h(\varphi_{Ts}(x))\, ds - \langle h \rangle \right| \\ &= \left| \int_0^1 g(s) \frac{d^m}{dt^m} H(\varphi_t)(x)\bigg|_{t=sT} ds \right|. \end{aligned}$$

Finally, integrating by parts m times and noting that the boundary terms vanish by (8) shows that, for all $x \in S$,

$$\begin{aligned} |WB_T(h)(x) - \langle h \rangle| &= \frac{1}{T^m} \left| \int_0^1 g^{(m)}(s) H(\varphi_{sT}(x))\, ds \right| \\ &\le \frac{1}{T^m} \|g^{(m)}\|_1 \|H\|_\infty. \end{aligned}$$

Since $\|g^{(m)}\|_1 < \infty$ and $\|H\|_\infty < \infty$ by assumption, we can take $c_m = \|g^{(m)}\|_1 \|H\|_\infty$ to give the result. □

Remark 2. The condition (30) is equivalent to a function being 'cohomologous' to its average [46], with some further regularity assumed. It is difficult to ascertain whether, for a given function h, there exists a function H. This is due to the fact that finding H in (30) requires knowledge of both the function h and the flow φ_t. The numerical observations that were given in Section 5 imply that for a chaotic orbit super-polynomial convergence is not observed. This indicates it is impossible to find the desired H for a chaotic orbit.

Remark 3. Theorem 3 can be obtained as a corollary of Theorem 4. As in the proof of Theorem 3, assume $h \in L^2(\mathbb{T}^d)$ and assume coordinates $\theta \in \mathbb{T}^d$ is taken so that $\varphi_t(\theta) = \theta + t\omega$. Denoting the Fourier series for h by $h = \sum_{k \in \mathbb{Z}^d} a_k e^{2\pi i k \cdot \theta}$, we can then set

$$H = -\sum_{k \in \mathbb{Z}^d} \frac{a_k}{(2\pi i k \cdot \omega)^m} e^{2\pi i k \cdot \theta}.$$

Then, for each m, h is of the form (30). Note that as m grows, H is bounded provided ω is Diophantine.

References

[1] V.I. Arnold, Mathematical Methods of Classical Mechanics, Springer-Verlag, New York, 1978, https://www.springer.com/us/book/9780387968902.

[2] J.M. Greene, A method for determining a stochastic transition, J. Math. Phys. 20 (1979) 1183–1201, http://dx.doi.org/10.1063/1.524170.

[3] R.S. MacKay, A renormalisation approach to invariant circles in area-preserving maps, Physica D 7 (1983) 283–300, http://dx.doi.org/10.1016/0167-2789(83)90131-8.

[4] A. Haro, R. de la Llave, A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: Numerical algorithms, Discrete Contin. Dyn. Syst. B 6 (6) (2006) 1261–1300, http://dx.doi.org/10.3934/dcdsb.2006.6.1261.

[5] J. Auslander, J.A. Yorke, Interval maps, factors of maps, and chaos, Tôhoku Math. J. 32 (2) (1980) 177–188, http://dx.doi.org/10.2748/tmj/1178229634.

[6] J.D. Meiss, Differential Dynamical Systems: Revised Edition, revised ed., in: Mathematical Modeling and Computation, vol. 22, SIAM, 2017, http://dx.doi.org/10.1137/1.9781611974645.

[7] C. Robinson, Dynamical Systems: Stability, Symbolic Dynamics, and Chaos, second ed., in: Studies in Advanced Mathematics, CRC Press, Boca Raton, Fla, 1999, http://dx.doi.org/10.1201/9781482227871.

[8] P. Cincotta, C. Giordano, Theory and applications of the mean exponential growth factor of nearby orbits (megno) method, in: C. Skokos, G.A. Gottwald, J. Laskar (Eds.), Chaos Detection and Predictability, in: Lecture Notes in Physics, vol. 915, Springer, Heidelberg, 2016, pp. 93–126, http://dx.doi.org/10.1007/978-3-662-48410-4_5.

[9] C. Froeschle, R. Gonczi, E. Lega, The fast Lyapunov indicator: A simple tool to detect weak chaos. application to the structure of the main asteroidal belt, Planet. Space Sci. 45 (7) (1997) 881–886, http://dx.doi.org/10.1016/S0032-0633(97)00058-5.

[10] E. Lega, M. Guzzo, C. Froeschle, Theory and applications of the fast Lyapunov indicator (fli) method, in: C. Skokos, G.A. Gottwald, J. Laskar (Eds.), Chaos Detection and Predictability, in: Lecture Notes in Physics, vol. 915, Springer, Heidelberg, 2016, pp. 35–54, http://dx.doi.org/10.1007/978-3-662-48410-4_2.

[11] C. Giordano, P. Cincotta, Chaotic diffusion of orbits in systems with divided phase space, Astron. Astrophys. 423 (2004) 745–753, http://dx.doi.org/10.1051/0004-6361:20040153.

[12] C.V. Abud, I.L. Caldas, On Slater's criterion for the breakup of invariant curves, Physica D 308 (2015) 34–39, http://dx.doi.org/10.1016/j.physd.2015.06.005.

[13] G.A. Gottwald, I. Melbourne, The 0-1 test for Chaos: A review, in: C. Skokos, G.A. Gottwald, J. Laskar (Eds.), Chaos Detection and Predictability, in: Lecture Notes in Physics, Springer, Berlin, 2016, pp. 221–247, http://dx.doi.org/10.1007/978-3-662-48410-4_7.

[14] C. Skokos, T. Manos, The smaller (SALI) and the generalized (GALI) alignment indices: Efficient methods of chaos detection, in: C. Skokos, G.A. Gottwald, J. Laskar (Eds.), Chaos Detection and Predictability, in: Lecture Notes in Physics, Springer, Berlin, 2016, pp. 129–181, http://dx.doi.org/10.1007/978-3-662-48410-4_5.

[15] B.R. Hunt, E. Ott, Defining chaos, Chaos 25 (9) (2015) 097618, http://dx.doi.org/10.1063/1.4922973.

[16] R.S. MacKay, I.C. Percival, Converse KAM: Theory and practice, Comm. Math. Phys. 98 (1985) 469–512, http://dx.doi.org/10.1007/BF01209326.

[17] N. Duignan, J.D. Meiss, Nonexistence of invariant tori transverse to foliations: An application of converse KAM theory, Chaos 31 (1) (2021) 013124, http://dx.doi.org/10.1063/5.0035175.
[18] S. Das, Y. Saiki, E. Sander, J.A. Yorke, Quasiperiodicity: Rotation numbers, in: C. Skiadas (Ed.), The Foundations of Chaos Revisited: From Poincaré to Recent Advancement, Understanding Complex Systems, Springer, 2016, http://dx.doi.org/10.1007/978-3-319-29701-9_7.
[19] S. Das, C.B. Dock, Y. Saiki, M. Salgado-Flores, E. Sander, J. Wu, J.A. Yorke, Measuring quasiperiodicity, Euro. Phys. Lett. 114 (2016) 40005, http://dx.doi.org/10.1209/0295-5075/114/40005.
[20] S. Das, Y. Saiki, E. Sander, J.A. Yorke, Quantitative quasiperiodicity, Nonlinearity 30 (11) (2017) 4111, http://dx.doi.org/10.1088/1361-6544/aa84c2.
[21] J. Laskar, C. Froeschlé, A. Celletti, The measure of chaos by the numerical analysis of the fundamental frequencies. Application to the standard mapping, Physica D 56 (1992) 253–269, http://dx.doi.org/10.1016/0167-2789(92)90028-L.
[22] R. Bartolini, A. Bazzani, M. Giovanniozzi, E. Todesco, Tune evaluation in simulations and experiments, Part. Accel. 52 (1996) 147–177, http://cdsweb.cern.ch/record/292773?ln=en.
[23] J. Sánchez, M. Net, C. Simó, Computation of invariant tori by Newton–Krylov methods in large-scale dissipative systems, Physica D 239 (3) (2010) 123–133, http://dx.doi.org/10.1016/j.physd.2009.10.012.
[24] N. Miguel, C. Simó, A. Vieiro, From the Hénon conservative map to the Chirikov standard map for large parameter values, Regul. Chaotic Dyn. 20 (2013) 469–489, http://dx.doi.org/10.1134/S1560354713050018.
[25] A. Luque, J. Villanueva, Numerical computation of rotation numbers of quasi-periodic planar curves, Physica D 238 (20) (2009) 2025–2044, http://dx.doi.org/10.1016/j.physd.2009.07.014.
[26] T.M. Seara, J. Villanueva, On the numerical computation of Diophantine rotation numbers of analytic circle maps, Physica D 217 (2) (2006) 107–120, http://dx.doi.org/10.1016/J.Physd.2006.03.013.
[27] S. Das, J.A. Yorke, Super convergence of ergodic averages for quasiperiodic orbits, Nonlinearity 31 (2) (2018) 491–501, http://dx.doi.org/10.1088/1361-6544/aa99a0.
[28] A.G. Kachurovskii, I.V. Podvigin, A.A. Svishchev, The maximum pointwise rate of convergence in Birkhoff's Ergodic theorem, J. Math. Sci. 255 (2) (2021) 119–123, http://dx.doi.org/10.1007/s10958-021-05354-x.
[29] E. Sander, J.D. Meiss, Birkhoff averages and rotational invariant circles for area-preserving maps, Physica D 411 (2020) 132569, http://dx.doi.org/10.1016/j.physd.2020.132569.
[30] J.D. Meiss, E. Sander, Birkhoff averages and the breakdown of invariant tori in volume-preserving maps, Physica D 428 (15) (2021) 133048, http://dx.doi.org/10.1016/j.physd.2021.133048.
[31] R.S. MacKay, A criterion for non-existence of invariant tori for Hamiltonian systems, Physica D 36 (1989) 64–82, http://dx.doi.org/10.1016/0167-2789(89)90248-0.
[32] F.J. Romeiras, E. Ott, Strange nonchaotic attractors of the damped pendulum with quasiperiodic forcing, Phys. Rev. A 35 (10) (1987) 4404–4412, http://dx.doi.org/10.1103/PhysRevA.35.4404.
[33] P. Glendinning, T.H. Jäger, G. Keller, How chaotic are strange non-chaotic attractors? Nonlinearity 19 (2006) 2005–2022, http://dx.doi.org/10.1088/0951-7715/19/9/001.
[34] R.D. Hazeltine, J.D. Meiss, Plasma Confinement, second ed., Dover Publications, Mineola, NY, 2003, https://store.doverpublications.com/0486151034.html.
[35] P. Helander, Theory of plasma confinement in non-axisymmetric magnetic fields, Rep. Progr. Phys. 77 (8) (2014) 087001, http://dx.doi.org/10.1088/0034-4885/77/8/087001.
[36] E.J. Paul, S.R. Hudson, P. Helander, Heat conduction in an irregular magnetic field. Part 2. Heat transport as a measure of the effective non-integrable volume, J. Plasma Phys. 88 (1) (2022) http://dx.doi.org/10.1017/S0022377821001306.
[37] P. Walters, An introduction to ergodic theory, in: Graduate Texts in Mathematics, Springer-Verlag, 1982.
[38] G.D. Birkhoff, Proof of the ergodic theorem, Proc. Natl. Acad. Sci. 17 (1931) 656–660, http://dx.doi.org/10.1073/pnas.17.2.656.
[39] G.D. Birkhoff, Proof of a recurrence theorem for strongly transitive systems, Proc. Natl. Acad. Sci. 17 (12) (1931) 650–655, http://dx.doi.org/10.1073/pnas.17.12.650.
[40] P. Billingsley, Ergodic Theory and Information, Wiley, New York, 1965.
[41] I.P. Cornfeld, S.V. Fomin, Ya.G. Sinai, Ergodic Theory, in: Grundlehren Der Mathematischen Wissenschaften, vol. 245, Springer-Verlag, New York, 1982.
[42] L. Breiman, Probabilty, in: Classics in Applied Mathematics, vol. 7, SIAM, Philadelphia, 1992.
[43] V. Bergelson, A. Leibman, C.G. Moreira, From discrete-to continuous-time ergodic theorems, Ergodic Theory Dynam. Systems 32 (2) (2012) 383–426, http://dx.doi.org/10.1017/S0143385711000848.
[44] C.C. Pugh, M. Shub, Ergodic elements of ergodic actions, Compos. Math. 23 (1) (1971) 115–122, http://www.numdam.org/item/CM_1971__23_1_115_0/.
[45] U. Krengel, On the speed of convergence in the ergodic theorem, Monatshefte Math. 86 (1978) 3–6, http://dx.doi.org/10.1007/BF01300052.
[46] A.G. Kachurovskii, The rate of convergence in ergodic theorems, Russian Math. Surveys 51 (4) (1996) 653, http://dx.doi.org/10.1070/RM1996v051n04ABEH002964.
[47] S. Das, Y. Saiki, E. Sander, J.A. Yorke, Solving the Babylonian problem of quasiperiodic rotation rates, Discrete Contin. Dyn. Syst. 12 (8) (2019) 2279–2305, http://dx.doi.org/10.3934/dcdss.2019145.
[48] L.L. Silverman, On the notion of summability for the limit of a function of a continuous variable, Trans. Amer. Math. Soc. 17 (3) (1916) 284–294, http://dx.doi.org/10.1090/S0002-9947-1916-1501042-8.
[49] Z. Levnajic, I. Mezic, Ergodic theory and visualization I: Mesochronic plots for visualization of ergodic partition and invariant sets, Chaos 20 (3) (2010) 033114, http://dx.doi.org/10.1063/1.3458896.
[50] P. Lochak, Canonical perturbation theory via simultaneous approximation, Russian Math. Surveys 47 (6) (1992) 59–140, http://dx.doi.org/10.1070/RM1992v047n06ABEH000965.
[51] T.W. Cusick, The two-dimensional Diophantine approximation constant, Monatshefte Math. 78 (1974) 297–304, http://dx.doi.org/10.1007/BF01294641.
[52] J.W.S. Cassels, An introduction to the geometry of numbers, in: Classics in Mathematics, Springer-Verlag, 1997, http://dx.doi.org/10.1007/978-3-642-62035-5.
[53] L. Grafakos, Classical Fourier Analysis, in: Graduate Texts in Mathematics, vol. 249, Springer, New York, 2015, http://dx.doi.org/10.1007/978-1-4939-1194-3.
[54] D.F. Escande, F. Doveil, Renormalization method for computing the threshold of the large scale stochastic instability in two degree of freedom Hamiltonian systems, J. Stat. Phys. 26 (1981) 257–284, http://dx.doi.org/10.1007/BF01013171.
[55] C. Rackauckas, Q. Nie, Differentialequations.jl–a performant and feature-rich ecosystem for solving differential equations in Julia, J. Open Res. Softw. 5 (1) (2017) http://dx.doi.org/10.5334/jors.151.
[56] J.H. Verner, Explicit Runge–Kutta methods with estimates of the local truncation error, SIAM J. Numer. Anal. 15 (4) (1978) 772–790, http://dx.doi.org/10.1137/0715051.
[57] C. Grebogi, E. Ott, S. Pelikan, J.A. Yorke, Strange attractors that are not chaotic, Physica D 13 (1984) 261–268, http://dx.doi.org/10.1016/0167-2789(84)90282-3.
[58] J.D. Meiss, Symplectic maps, variational principles, and transport, Rev. Modern Phys. 64 (3) (1992) 795–848, http://dx.doi.org/10.1103/RevModPhys.64.795.