PAPER

Quantitative quasiperiodicity

View the article online for updates and enhancements.

Related content

- Measuring quasiperiodicity
  Suddhasattwa Das, Chris B. Dock, Yoshitaka Saiki et al.

- Scaling law in the standard map critical function
  Timoteo Carletti and Jacques Laskar

- Towards global models near homoclinic tangencies of dissipative diffeomorphisms
  Henk Broer, Carles Simó and Joan Carles Tatjer

Recent citations

- Super convergence of ergodic averages for quasiperiodic orbits
  Suddhasattwa Das and James A Yorke

Quantitative quasiperiodicity

Suddhasattwa Das[1,6], **Yoshitaka Saiki**[2,3,5]**, Evelyn Sander**[4]
and James A Yorke[5]

[1] Department of Mathematics, University of Maryland, College Park, MD,
United States of America
[2] Graduate School of Commerce and Management, Hitotsubashi University,
Tokyo, Japan
[3] JST, PRESTO, Saitama, Japan
[4] Department of Mathematical Sciences, George Mason University, VA,
United States of America
[5] University of Maryland, College Park, MD, United States of America

E-mail: dass@cims.nyu.edu

Abstract

The Birkhoff ergodic theorem concludes that time averages, i.e. Birkhoff averages, $B_N(f) := \Sigma_{n=0}^{N-1} f(x_n)/N$ of a function f along a length N ergodic trajectory (x_n) of a function T converge to the space average $\int f \mathrm{d}\mu$, where μ is the unique invariant probability measure. Convergence of the time average to the space average is slow. We use a modified average of $f(x_n)$ by giving very small weights to the 'end' terms when n is near 0 or $N-1$. When (x_n) is a trajectory on a quasiperiodic torus and f and T are C^∞, our weighted Birkhoff average (denoted $WB_N(f)$) converges 'super' fast to $\int f \mathrm{d}\mu$ with respect to the number of iterates N, i.e. with error decaying faster than N^{-m} for every integer m. Our goal is to show that our weighted Birkhoff average is a powerful computational tool, and this paper illustrates its use for several examples where the quasiperiodic set is one or two dimensional. In particular, we compute rotation numbers and conjugacies (i.e. changes of variables) and their Fourier series, often with 30-digit accuracy.

(Some figures may appear in colour only in the online journal)

---

[6] Author to whom any correspondence should be addressed.

1. Introduction

Quasiperiodicity is a key type of observed dynamical behavior in a diverse set of applications. We say a map T is (d-dimensionally) quasiperiodic (for $d \geqslant 1$) if (i) $T : \mathbb{T}^d \to \mathbb{T}^d$ and (ii) each trajectory is dense in $\mathbb{T}^d$ and (iii) there is a continuous choice of coordinates $\theta = (\theta_1, \cdots, \theta_d)$ and some $\rho = (\rho_1, \cdots, \rho_d) \in \mathbb{T}^d$ for which the T has the form

$$T(\theta) = \theta + \rho \bmod 1. \tag{1}$$

Condition (ii) can be replaced by saying in dimension $d = 1$ that ρ is irrational and in dimension $d > 1$ that all of the coordinates ρ_j of ρ are irrational and they are 'independent' over the reals; that is if $a = (a_1, \cdots, a_d)$ is a vector of integers and $\sum_{j=1}^{d} a_j \rho_j = 0$, then every $a_j = 0$. We then say such a ρ is irrational.

Let T be a C^∞ quasiperiodic map. The quasiperiodicity persists for most small perturbations by the Kolmogorov–Arnold–Moser theory. We believe that quasiperiodicity is one of only three types of invariant sets with a dense trajectory that can occur in typical smooth maps. The other two types are periodic sets and chaotic sets. See [1] for the statement of our formal conjecture of this triumvirate. For example, quasiperiodicity occurs in a system of weakly coupled oscillators, in which there is an invariant smooth attracting torus in phase space with behavior that can be described exclusively by the phase angles of rotation of the system. Indeed, it is the property of the motion being described using only a set of phase angles that always characterizes quasiperiodic behavior. In a now classical set of papers, Newhouse, Ruelle, and Takens demonstrated a route to chaos through a region with quasiperiodic behavior, causing a surge in the study of the motion [2]. There is active current interest in development of a systematic numerical and theoretical approach to bifurcation theory for quasiperiodic systems. Our goal in this paper is to present a numerical method for the fast calculation of the limit of Birkhoff averages in quasiperiodic systems, allowing us to compute various key quantities.

If f is integrable and the dynamical system is ergodic on the set in which the trajectory lives, then the Birkhoff ergodic theorem asserts that the Birkhoff average B_N defined as

$$\mathrm{B}_N(f) := \sum_{n=0}^{N-1} f(x_n)/N \tag{2}$$

of a function f along an ergodic trajectory (x_n) converges to the space average $\int f \mathrm{d}\mu$ as $N \to \infty$ for μ-almost every x_0, where μ is the unique invariant probability measure. In particular for quasiperiodic systems all trajectories with initial point in the ergodic set have the same limit of their Birkhoff averages. We develop a numerical technique for calculating the limit of such averages, where instead of weighting the terms $f(x_n)$ in the average equally, we weight the early and late terms of the set $\{0, \ldots, N-1\}$ much less than the terms with $n \sim N/2$ in the middle. That is, rather than using the equal weighting $(1/N)$ in the Birkhoff average, we use a weighting function $w(n/N)$.

1.1. Weighted Birkhoff averaging method

A function $w : \mathbb{R} \to [0, \infty)$ will be called a C^∞ weighting function if w is infinitely differentiable and $w > 0$ on $(0, 1)$ and $= 0$ elsewhere. The example of such a function that we will use in this paper is in equation (3) (figure 1) defined as

$$w(t) := \begin{cases} \exp\left(\frac{-1}{t(1-t)}\right), & \text{for } t \in (0, 1) \\ 0, & \text{for } t \notin (0, 1). \end{cases} \tag{3}$$
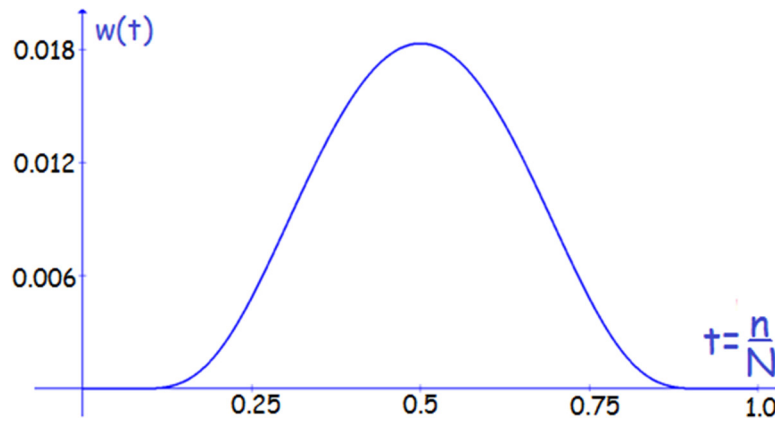
Figure 1. A C^∞ weighting function. The figure shows the graph of a the function $w(t)$ defined in equation (3). This function plays the role of a temporal weighting for the weighted Birkhoff average. That is, rather than using the equal weighting $(1/N)$ in the usual Birkhoff average, we use the non-uniform weight proportional to $w(n/N)$, where n is the iterate number and N is the orbit length. See equations (3) and (4) for a description of this weighted Birkhoff average.

See equation (16) for a family of weighting functions $w^{[p]}(t)$ that converge even faster when many digits of precision are required. This is an example of what is often referred to as 'window' functions in spectral analysis or a 'bump' function in the theory of partitions of unity.

For $d = 1$, $\mathbb{T}^d$ is a simple closed curve. For a continuous function f and a C^∞ quasiperiodic map T on $\mathbb{T}^d$, let $x_n \in \mathbb{T}^d$ be such that $x_n = T(x_{n-1})$ for all $n \geqslant 1$. We define a weighted Birkhoff (WB$_N$) average of f as

$$\mathrm{WB}_N(f)(x) := \sum_{n=0}^{N-1} \hat{w}_{n,N} f(x_n), \text{ where } \hat{w}_{n,N} = \frac{w(n/N)}{\sum_{j=0}^{N-1} w(j/N)}. \tag{4}$$

Note that $\mathrm{WB}_N(f)$ is indeed an average of the values $f(x_n)$ since $\sum_{n=0}^{N-1} \hat{w}_{n,N} = 1$.

1.2. Main convergence result

The main convergence result we are using is theorem 3.1. It is proved in [3] and an outline of the proof is given here in section 4. We now state a special case of the theorem that avoids unnecessary terminology and states only the C^∞ case.

Assume f and T are C^∞ and w is a C^∞ weighting function, then for almost every rotation number ρ and for every positive integer m there is a constant $C_m > 0$ such that

$$\left| \mathrm{WB}_N(f)(\theta) - \int f(\theta)\mathrm{d}\theta \right| \leqslant C_m N^{-m}.$$

We refer to the above as super-fast (super polynomial) convergence or exponential convergence. The above constant C_m depends on (i) $w(t)$ and its first m derivatives; (ii) the function $f(t)$; and (iii) the rotation number(s) of the quasiperiodic trajectory or more precisely, the small divisors arising out of the rotation vector. Our method of averaging does not give improved convergence results for chaotic systems.

In [4–7], Laskar employs a Hanning data weighting function and the analogue of our $\sin^2$ weighting function for his computations, which lead to the convergence of order $1/N^2$ or $1/N^4$, and $1/N^3$ respectively, (p 136 in [4]) where N is the length of a orbit. Specifically, the weighting function $\cos^2(\pi x)$ for $x \in (-1, 1)$ and averages over iterates from $-N$ to N. He mentions the C^∞ filter we use in Remark 2 of the appendix of [4], p. 146, though he does not use it. There seems to be no advantage to using his lower order methods than C^∞ filter. In particular the programming of both is quite simple. We will compare the two methods in figures 7(b), 9, 12(b) and 14.

Other authors have considered related numerical methods (see section 3.9), in particular [8–10], which we will compare to our approach when we introduce our averaging method in section 3. See also [11–20]. We announced some of the results presented here in [21].

1.3. The Babylonian problem of quasiperiodic rotation numbers

What constitutes a 'big-data' problem depends on the speed of computation available. With this understanding, the first big-data problem was 2500+ years ago when the Babylonians computed the three periods of the Moon from data on the position of the Moon collected almost daily for many years. The Moon's position through the fixed stars can be viewed as a quasiperiodic trajectory with $d = 3$ and their problem was to compute the rotation numbers from such a trajectory, which they did with high accuracy. See [22].

1.4. Applications

We demonstrate our weighted Birkhoff averaging method and its convergence rate by computing rotation numbers, conjugacies (i.e. changes of variables), and their Fourier series in dimensions one and two. We will refer to a one-dimensional quasiperiodic curve as a *curve*.

We start by describing our results for a key example of quasiperiodicity: the (circular, planar) restricted three-body problem (R3BP). This is an idealized model of the motion of a planet, a large moon, and an asteroid governed by Newtonian mechanics, in a model studied by Poincaré [23, 24]. In particular, we consider a planar three-body problem consisting of two massive bodies ('planet' and 'moon') moving in circles about their center of mass and a third body ('asteroid') whose mass is infinitesimal, having no effect on the dynamics of the other two.

We assume that the moon has mass μ and the planet mass is $1 - \mu$ where $\mu = 0.1$, and writing equations in rotating coordinates around the center of mass. Thus the planet remains fixed at $(q_1, p_1) = (-0.1, 0)$, and the moon is fixed at $(q_2, p_2) = (0.9, 0)$. In these coordinates, the satellite's location and velocity are given by the *generalized position vector* (q_1, q_2) and *generalized velocity vector* (p_1, p_2). The equations of motion are as follows.

$$
\begin{aligned}
\frac{dq_1}{dt} &= p_1 + q_2, \\
\frac{dq_2}{dt} &= p_2 - q_1, \\
\frac{dp_1}{dt} &= p_2 - \mu \frac{q_1 - 1 + \mu}{d_{\text{moon}}^3} - (1 - \mu) \frac{q_1 + \mu}{d_{\text{planet}}^3}, \\
\frac{dp_2}{dt} &= -p_1 - \mu \frac{q_2}{d_{\text{moon}}^3} - (1 - \mu) \frac{q_2}{d_{\text{planet}}^3},
\end{aligned}
\tag{5}
$$

where

$$
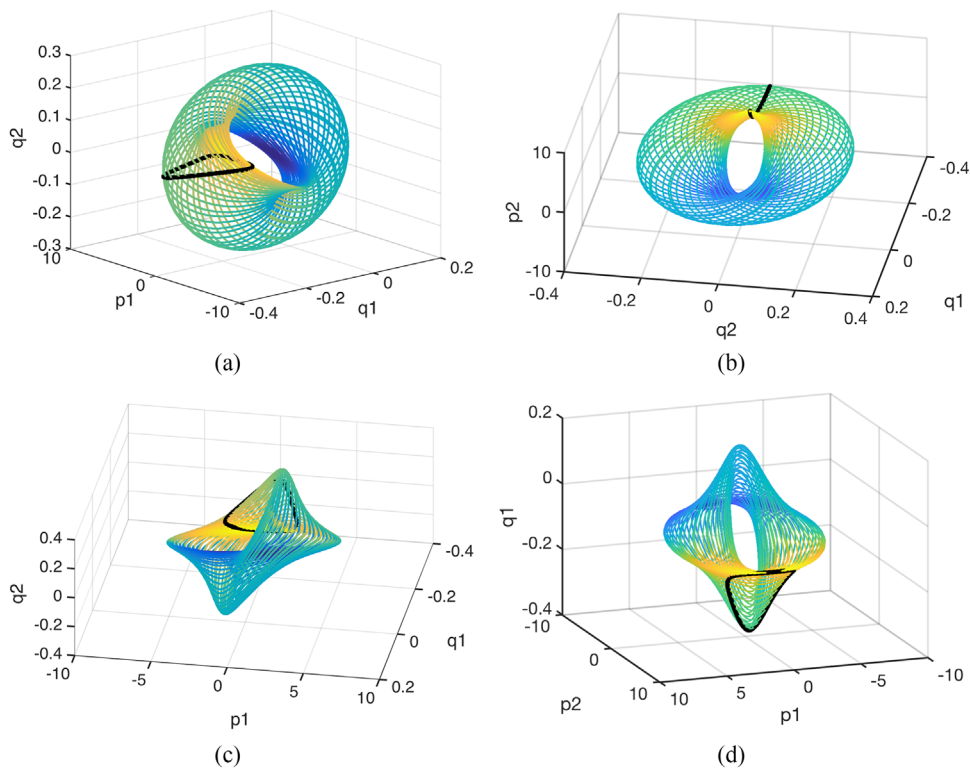d_{\text{moon}}^2 = (q_1 - 1 + \mu)^2 + q_2^2
$$

Figure 2. Torus flow for the R3BP. This trajectory is a solution of equation (5), shown as curve B_1 in figure 3. All four views are of the same two-dimensional quasiperiodic torus lying in $\mathbb{R}^4$. Each picture consists of the same trajectory spiraling densely on this torus. We require four different views of this torus because the embedding into three dimensions gives a highly non-intuitive images. In all four panels, the color of the trajectory is the value of the variable p_2. The black curve is the set of values of the Poincaré return map with $q_2 = 0$ for this flow torus.

$$d_{\text{planet}}^2 = (q_1 + \mu)^2 + q_2^2.$$

The following function H is a Hamiltonian for this system

$$H = \frac{p_1^2 + p_2^2}{2} + (p_1 q_2 - p_2 q_1) + \left(-\frac{1 - \mu}{d_{\text{planet}}} - \frac{\mu}{d_{\text{moon}}} \right), \tag{6}$$

where $p_1 = \dot{q}_1 - q_2$ and $p_2 = \dot{q}_2 + q_1$ (see [25] p 59 equations (63)–(66)).

The three terms in equation (6) are respectively the kinetic energy, angular moment, and the potential. For fixed H, Poincaré reduced this problem to the study of the Poincaré return map for a fixed value of H, only considering a discrete trajectory of the values of (q_1, p_1) on the section $q_2 = 0$ and $\mathrm{d}q_2/\mathrm{d}t > 0$. Thus we consider a map in two dimensions rather than a flow in four dimensions. Figure 2 shows one possible motion of the asteroid for the full flow. The orbit is spiraling on a torus. The black curve shows the corresponding trajectory on the Poincaré return map. Figure 3 shows the Poincaré return map for the asteroid for a variety of starting points. A variety of orbits are shown, most of which are quasiperiodic invariant curves.
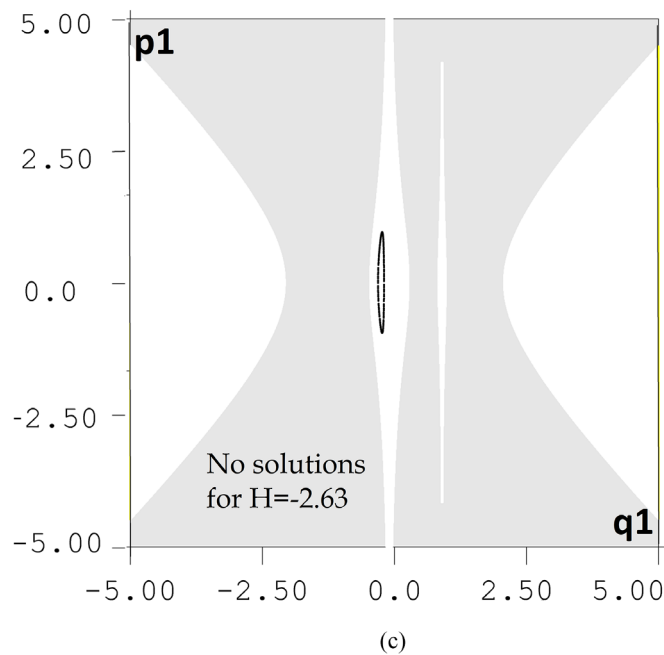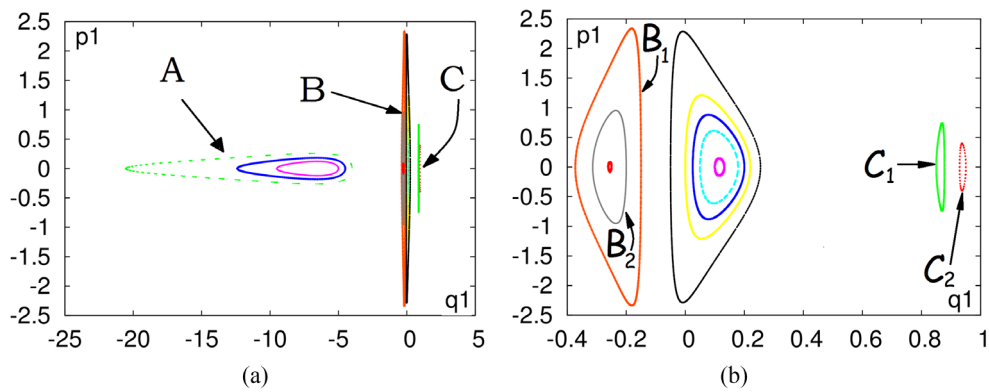
Figure 3. Poincaré-return map for R3BP. Panels (a) and (b) show a projection to the $q_1 - p_1$ plane of solutions to the R3BP in equation (5). The value of the Hamiltonian H for all the curves shown is the same and $H \approx -2.63$. They show various quasiperiodic trajectories on the Poincaré section corresponding to $q_2 = 0$. Each time the flow hits $q_2 = 0$ and $\mathrm{d}q_2/\mathrm{d}t > 0$, we plot (q_1, p_1). Note that the planet is fixed at the point $(-0.1, 0)$ and the moon at $(0.9, 0)$. Thus some trajectories orbit both the planet–moon system and some orbit only the planet or only the moon. Each trajectory shown is a quasiperiodic curve. In Panel (c), white indicates the region of the (q_1, p_1) plane where the Poincaré return map is defined for $H \approx -2.63$, and gray indicates the region of the plane where the Poincaré return map is not defined. Panel (c) also shows the trajectory which corresponds to the curve B_1.

An exception is trajectory A in figure 3(a), which is an invariant recurrent set consisting of 42 curves. Each curve is an invariant quasiperiodic curve under the 42nd iterate of the map.

Our paper proceeds as follows: in section 2, we give the formal definition of quasiperiodicity, rotation number, and the conjugacy map to the rigid rotation. In section 3, we describe

our numerical technique in detail. We illustrate our weighted Birkhoff averaging method for a series of four examples, including an example of a two-dimensionally quasiperiodic map. In all cases, we get fast convergence and are in most cases able to give results with about 30-digit accuracy. Section 4 describes what happens when a rotation number is unusually well approximated by a fraction with small denominator. In such cases we elliptically say the rotation number is 'nearly rational'. Finally, section 5 contains our concluding remarks.

2. Quasiperiodicity

In the introduction, we described quasiperiodic motion as motion that could be fully understood through a set of angles of rotation. We now formalize that idea in the following definition.

2.1. Quasiperiodicity

For a dimension $d \geqslant 1$, let $\rho = (\rho_1, \rho_2, \ldots, \rho_d)$ be a vector whose coordinates are irrational and are independent over the integers (see equation (1)). The following map $T_\rho : \mathbb{T}^d \to \mathbb{T}^d$ is called a rigid rotation:

$$T_\rho(\theta) = \theta + \rho \bmod 1, \text{ where 'mod' is applied to each coordinate.} \qquad (7)$$

A rigid rotation is the simplest, albeit least interesting example of a map with quasiperiodicity. Since T_ρ gives the same values on opposite sides of the unit cube, we identify the sides and refer to the domain of the rigid rotation as a curve in one dimension and a d-torus in dimension $d > 1$. In this paper, we will sometimes refer to the curve as a 1-torus. We define a map T on an ambient space $\mathbb{R}^D$ to be quasiperiodic if either T or some iterate T^k is topologically conjugate to a rigid rotation. We will assume $k = 1$ in the rest of this description. That is, a map T is quasiperiodic if there is a rigid rotation map T_ρ and an invertible conjugacy map (i.e. change of coordinates) $h : \mathbb{T}^d \to \mathbb{R}^D$ such that

$$T(h(\theta)) = h(T_\rho(\theta)). \qquad (8)$$

A flow has quasiperiodic behavior if its associated Poincaré return map has quasiperiodic behavior.

For an invertible map T to be quasiperiodic on a curve C, it is necessary and sufficient that T has a dense trajectory, as shown in [26]. In general, a one-time differentiable invertible map on a curve without periodic points may not be quasiperiodic. However, if we assume that the map T and the curve C are twice continuously differentiable, then Denjoy [27] showed that these conditions are both necessary and sufficient. Furthermore, clearly any rigid irrational rotation map is a real analytic map, but even if we assume that a quasiperiodic function is analytic, Arnold showed that the conjugacy map h may only be continuous for some atypical rotation number. However, Herman (see [28]) proved that for homeomorphisms on a circle, most conjugacy maps h are analytic. Yamaguchi and Tanikawa [29] and Hunt, Khanin, Sinai and Yorke [30] show that the critical KAM curve may not be C^2.

2.2. Diophantine rotations

An irrational vector $\rho \in \mathbb{R}^d$ is said to be Diophantine if for some $\beta > 0$ it is Diophantine of class β (see [28], definition 3.1), which means there exists $C_\beta > 0$ such that for every $k \in \mathbb{Z}^d$, $k \neq 0$ and every $n \in \mathbb{Z}$,

$$|k \cdot \rho - n| \geqslant \frac{C_\beta}{\|k\|^{d+\beta}}. \tag{9}$$

We conjecture that if the map is analytic, then almost every quasiperiodic torus having a rotation number that is Diophantine (i.e. far from rational) is real analytic.

2.2.1. Assigning angular coordinates. Let (x_n) be the forward orbit under T, and (θ_n) the forward orbit under T_ρ. That is, $x_{n+1} = T(x_n)$ and $\theta_{n+1} = \theta_n + \rho \pmod 1$.

In all of the $d = 1$ examples discussed here, the quasiperiodic curves are simple, closed, convex curves, and hence, angular coordinates can be obtained using polar coordinates (ϕ, r) with respect to a fixed and pre-defined center, where r is uniquely determined by ϕ. More generally we may have an image of a quasiperiodic curve or torus that is not an embedding as a closed, convex curve. Such is not needed, since we have a method based on the Takens delay coordinate maps that solves the problem. This general method of obtaining rotation numbers from the image of a quasiperiodic curve is announced in [21] with a complete description in [22].

Because of our nice embeddings the orbits $(\theta_n = n\rho)$, (x_n) and (ϕ_n) are all conjugate and there is a continuous map $V : S^1 \to S^1$ such that

$$x_n = h(\theta_n), \quad \phi_n = V(\theta_n), \quad \theta_n = n\rho \bmod 1 \text{ for all } n = 0, 1, 2, 3, \ldots \tag{10}$$

Since V is invertible, the following map is periodic.

$$g(\theta) := V(\theta) - \theta \tag{11}$$

For example, in figure 4(a), the angular coordinate ϕ for the quasiperiodic curve B_1 of the restricted three-body problem is measured from a point $(q_1, p_1) = (-0.2, 0)$.

3. Weighted Birkhoff averaging WB$_N$ and its applications

As mentioned in section 1, our approach is to modify the regular Birkhoff average in equation (2) using the weighting function w in equation (3) (see figure 1) so that for quasiperiodic dynamical systems, the weighted Birkhoff average in equation (4) convergences much faster to the same limit $\int f(\theta) \mathrm{d}\theta$. We will formalize this claim using the main result from the companion paper [3].

Definition. A function $w : \mathbb{R} \to [0, \infty)$ is said to be a C^∞ bump function if w is C^∞ and the support of w is $[0, 1]$ and $\int_{\mathbb{R}} w(x) \mathrm{d}x \neq 0$ and w and all of its derivatives vanish at 0 and 1.

Theorem 3.1 (Theorems 1.1, 3.1 in [3]). *For $r \in [1, \infty]$, let X be a C^r manifold and $T : X \to X$ be a C^r map which is d-dimensionally quasiperiodic on an invariant set $X_0 \subseteq X$, with invariant probability measure μ and a rotation vector of Diophantine class $\beta(> 0)$. Let $f : X \to \mathbb{R}^k$ be a C^r map. Let $m > 1$ be an integer such that $r \geqslant d + m(d + \beta)$, and let $w : \mathbb{R} \to \mathbb{R}$ be a C^m bump function. Then there is a constant C_m depending upon w, f, m, M, and β but independent of $x_0 \in X_0$ such that*

$$\left| (\mathrm{WB}_N f)(x_0) - \int_{X_0} f \, \mathrm{d}\mu \right| \leqslant C_m N^{-m}. \tag{12}$$

In particular, if $r = \infty$ and w is a C^∞ bump function, then equation (12) holds for every $m \in \mathbb{N}$.

Theorem 3.1 is proved in section 4 in a way that lets us determine what happens when an irrational rotation number is near a rational number.
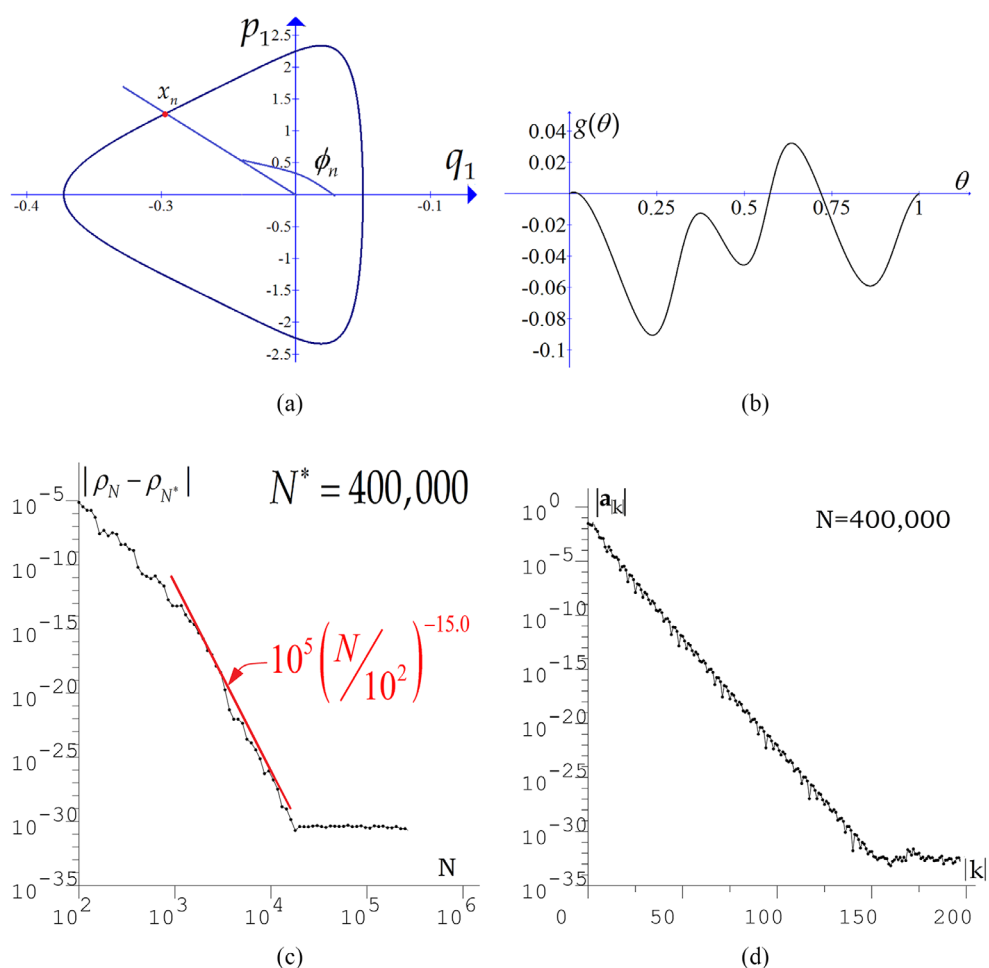
Figure 4. Quasiperiodicity for the R3BP. For the quasiperiodic curve B_1 from figure 3, panel (a) shows the curve B_1, along with the projection mapping from x_n, the *n*th iterate in the (q_1, p_1) coordinates, to angular coordinate ϕ_n. It is parameterized by coordinates $\phi \in S^1 \equiv [0, 1)$. Panel (b) depicts the periodic part $g(\theta)$ of the one-dimensional conjugacy map (equation (11)) between the quasiperiodic behavior and rigid rotation by ρ. The curve represents the true Fourier series up to 30 decimal digits. Panel (c) shows the convergence rate of the error in the rotation number ρ_N as a function of the number of iterates N. The 'error' is the difference $|\rho_N - \rho_{N^*}|$, where $N^* = 400\,000$ is large enough so that ρ_N appears to have converged. The straight line is an upper bound for the curve and its exponent indicates the rate of convergence. Panel (d) shows the norm of the Fourier coefficients of the conjugacy as a function of index. This exponential decay indicates that the conjugacy function is analytic, up to numerical precision. The step size used for the eighth order Runge–Kutta scheme is 2×10^{-5}.

3.1. Diophantine rotation numbers

The assumption on the rotation numbers being Diophantine means that the rotation numbers cannot be closely approximated by rational numbers with small denominators. For numbers which are not Diophantine, the trajectories have ergodic properties 'close' to periodic orbits and therefore, do not converge slower than the rate N^{-k} for some k.

3.2. Robustness of assumptions

It is well known that for every $\beta > 0$ the set of Diophantine vectors of class β have full Lebesgue measure in $\mathbb{R}^d$ (see for example, [28], 4.1). Thus, the assumption of the rotation number being Diophantine is robust in a measure theoretic sense, i.e. in physical experiments, the rotation number will be Diophantine with probability 1.

3.3. Computing a rotation number or rotation vector

We now show how to apply this averaging method in computation. We observe that N must generally be larger for $\mathbb{T}^2$ than for $\mathbb{T}^1$ to get a 30-digit accuracy.

According to the definition of 'quasiperiodicity', a quasiperiodic map is conjugate to a rigid rotation of the form equation (7) with rotation vector ρ. However, this vector ρ is not unique. When the dimension $d = 1$, there are two choices of ρ depending on whether you move clockwise or counterclockwise around the circle, and for $d > 1$, it is shown in [22] that there is a set of choices for ρ which are related to each other by unimodular transformations of $\mathbb{R}^d$ and are dense in $\mathbb{T}^d$. Our goal will be to find any one of these equivalent rotation vectors. The rotation vector of a quasiperiodic trajectory $(y_n)_{n=0}^{N-1}$ can be viewed as the average rotation traversed by the sequence of d-vectors $(T(\theta_n) - \theta_n)$, which by equation (10) is equivalent to the average of the angular increments $(\phi_{n+1} - \phi_n)$. To make the notion of the angular distance between ϕ_n and ϕ_{n+1} consistent, we must make this choice continuously. By conjugating with (hV^{-1}) from equation (10) if necessary, we may assume that the map T is in angular coordinates on $\mathbb{T}^d$.

We then associate with T a continuous map $\tilde{T}$ on the full Euclidean space $\mathbb{R}^d$ such that

$$\tilde{T}(z) \pmod 1 = T(z \pmod 1).$$

The map $\tilde{T}$ is called a lift of T, and $z \in \mathbb{R}^d$ is a lift of $\phi \in \mathbb{T}^d$.

Since T is invertible, the map $\tilde{T}(z) - z$ has period one in every coordinate direction. For example, the rigid rotation $T(\phi) = \phi + \rho \pmod 1$, for θ in $\mathbb{T}^d$ has a corresponding lift map $\tilde{T}(z) = z + \rho$. Of course if ρ was $\sqrt{2}$, we would have the same map T as for $\rho = \sqrt{2} - 1$ so we define $\tilde{T}$ using rotation numbers are in $(0, 1)$. Using the lift, we now give a formula for the rotation vector for the trajectory (y_n) starting at y_0:

$$\rho(T) := \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \left(\tilde{T}(z_n) - z_n \right). \tag{13}$$

This average converges slowly as $N \to \infty$, with order of at most $1/N$. However, since equation (13) can be written as a Birkhoff average by writing $f(z_n) = \tilde{T}(z_n) - z_n$, we can apply our method to this function. That is, let $(z_n)_{n=0}^{N-1}$ be an orbit for $\tilde{T}$. Our approximation of ρ is given by the weighted Birkhoff average of f,

$$\mathrm{WB}_N(z_{n+1} - z_n) := \sum_{n=0}^{N-1} \hat{w}_{n,N}(z_{n+1} - z_n) \to \rho \text{ as } N \to \infty. \tag{14}$$

3.4. Convergence rate of the weighted Birkhoff average WB_N

In order to illustrate the speed of convergence of our weighted Birkhoff average WB_N as $N \to \infty$, we introduce four different possible choices for the weighting function w, depicted
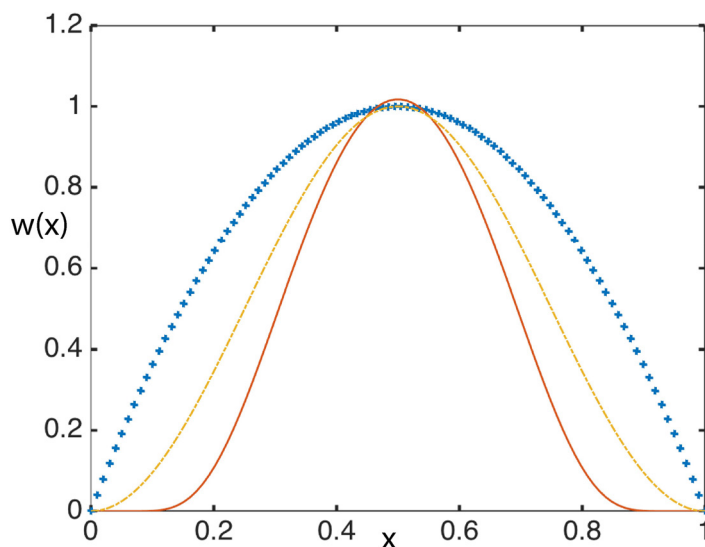
Figure 5. Variations of different weighting functions. This is a plot of three non-constant weighting functions from equation (15) and below w_{quad} (top), $w_{\sin^2}$ (second), $w_{\exp}$ (lowest). Since only the shape matters, they have been rescaled so that each has a peak of approximately 1.0.

in figure 5, and compare the convergence results for computing the rotation number for each of these choices of w.

$$w_{\text{equal}}(t) := 1 \text{ (Birkhoff's choice)}$$
$$w_{\text{quad}}(t) := t(1 - t)$$
$$w_{\sin^2}(t) := \sin^2(\pi t)$$
$$w^{[1]}(t) := w_{\exp}(t) := \exp\left(\frac{-1}{t(1 - t)}\right). \tag{15}$$

The weighting functions described above are defined to be 0 outside $(0, 1)$ and equal to the specified function inside $(0, 1)$. Recall that the last function in the list, $w_{\exp}$, is the function used in our calculations. It is the only one in the list that is C^∞. The others fail to be C^∞ at $t = 0$ and $t = 1$. A family of C^∞ weighting functions can be defined for $p \geqslant 1$ as

$$w^{[p]}(t) := \exp\left(\frac{-1}{t^p(1 - t)^p}\right) \text{ for } t \in (0, 1) \tag{16}$$

and $= 0$ elsewhere. This paper mainly uses $w = w^{[1]}$. The function $w^{[2]}$ results in an averaging method which converges noticeably faster than $w^{[1]}$ when using 30-digit precision, but not in 15-digit precision. We will write the weighted Birkhoff averages as $\text{WB}_N^{[1]}$ (or just WB_N) and $\text{WB}_N^{[2]}$ when using $w^{[1]}$ and $w^{[2]}$ respectively. See figure 7 where the two are compared.

When we compute with the first choice of w, we recover the truncated sum in the definition of the Birkhoff average. To estimate the error, we expect the difference $f(x_{N+1}) - f(x_N)$ to be of order one, implying that for w_{equal}, the error in the average is generally proportional to N^{-1} in our figures. The choice of a particular starting point also creates a similar uncertainty of order $1/N$. Every function w is always positive between 0 and 1. For all but the first choice, the function vanishes as t approaches 0 and 1. In addition, going down the list, increasing
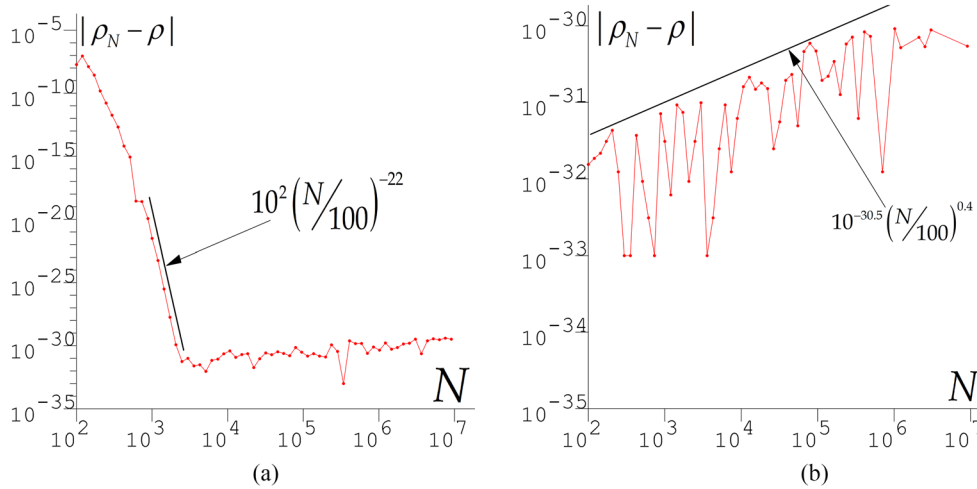
Figure 6. Testing how well the WB_N method can determine the rotation number. Panel (a) shows the convergence in the calculation of a known rotation number $\rho = \sqrt{2} - 1$, for the trajectory (ϕ_n) from equation ((17)), with $\alpha = 0.1$, $\beta = 0.2$, and $\rho = \sqrt{2} - 1$. The error quickly drops to the limit of numerical precision and then increases slowly as N increases. This increase in the error is apparently due to accumulated round-off error. Panel (b) shows the increasing round-off error in the rotation number for the trivial case $(\alpha = \beta = 0)$. Here, $\phi_n = n\rho \pmod 1$ and the error grows after attaining a minimum as N increases.

number of derivatives of w vanish for $t \to 0$ and $t \to 1$, with all derivatives of $w_{\exp}$ vanishing at 0 and 1. We thus expect the effect of the starting and endpoints to decay at the same rate as this number of vanishing derivatives. Indeed, we find that w_{quad} corresponds approximately to order $1/N^2$ convergence, $w_{\sin^2}$ to $1/N^3$ convergence, and $w_{\exp}$ to convergence faster than any polynomial in $1/N$, i.e. for every integer m, there is a constant $C > 0$ such that for N sufficiently large, $|WB_N f - \int f d\mu| \leqslant CN^{-m}$. Figures 12(b) and 9 show this effect. We have not tried other C^∞ weighting functions.

3.5. Estimating error when the true rotation number is known

In order to test the error in the calculation of rotation number, we present two examples below where we know the exact rotation number. This allows us to determine the actual error in the calculation for the WB_N method as N increases. In both cases the error decreases to less than 10^{-31} and then it grows as N increases, apparently due to accumulated round-off error.

Example 1. Let (θ_n) be an orbit under the rigid rotation described in equation (7) for a rotation by $\rho = \sqrt{2} - 1$. Assume that what we observe is ϕ, a perturbed version of θ, namely,

$$\phi_n = \theta_n + \alpha \cos(2\pi\theta_n) + \beta \sin(2\pi\theta_n), \text{ where } \theta_n = n\rho \pmod 1. \tag{17}$$

We use the weighted Birkhoff average as in equation ((14)) (changing y to ϕ) to obtain an estimate of the rotation number ρ from this orbit. Figure 6 shows the results for $\alpha = 0.1$ and $\beta = 0.2$ in (a) and for the case $\alpha = 0.0$ and $\beta = 0.0$ in (b).

Example 2. Figure 7 shows a geometric version of the problem from the previous example, and again the error in the rotation number is small.

(a)   (b)
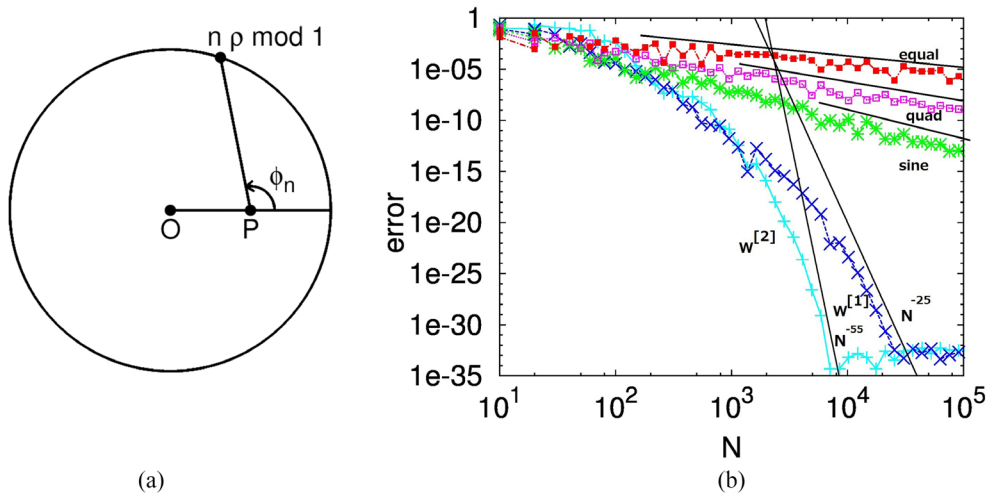
Figure 7. The error in the computed value of rotation number when the rotation number is known. Panel (a) shows the geometric configuration of the problem, with a constant rotation vector $\rho = \sqrt{2} - 1$ about the origin. The trajectory is $(n\rho \pmod 1)$, but the observer measures the angle ϕ as seen from its perspective at P, which is midway between the center of the circle, O, and the circle itself. Panel (b) shows the convergence in the rotation number calculation averaging $(\phi_{n+1} - \phi_n)$ five ways, using the Birkhoff average (top curve denoted 'equal') and next 'quad' and then 'sine' and then the weighted Birkhoff averages $\mathrm{WB}_N = \mathrm{WB}_N^{[1]}$ and $\mathrm{WB}_N^{[2]}$ (lowest curve), for the trajectory (ϕ_n) described in (a). The error from the known value $\sqrt{2} - 1$ is calculated for several values of number N. The weighted Birkhoff average $\mathrm{WB}_N^{[1]}$ reaches 32-digit accuracy by $N = 30\,000$ while $\mathrm{WB}_N^{[2]}$ does so by $N = 8000$, at which point the $sin^2(\pi t)$ curve has an accuracy to 10^{-9}, and its curve is proportional to $N^{-2.5}$.

3.6. Fourier coefficients and change of coordinates reconstruction

For a quasiperiodic curve as shown in figure 4(a), there are two approaches to representing the curve. Firstly, we can write the coordinates (X, Y) as a function of $\theta \in S^1$, or secondly, we can reduce the dimension and represent the points on the curve by an angle $\phi \in S^1$, that is, $\phi(X(\theta), Y(\theta))$, which is also $V(\theta) = \theta + g(\theta)$. We have shown g in figure 4(b) and the exponential decay of the norm of the Fourier coefficients in figure 4(d). We report the Fourier series for the periodic part $g(\theta)$.

Given a continuous periodic map $f : S^1 \to \mathbb{R}$, the Fourier series representation of f is the following:

$$\text{For every } \theta \in S^1, \, f(\theta) = \sum_{k=-\infty}^{\infty} a_k e^{i2\pi k\theta}, \tag{18}$$

where the complex coefficient a_k is given by the formula

$$a_k = \int_{\theta \in S^1} f(\theta) e^{-i2\pi k\theta} \, d\theta. \tag{19}$$

If we only have access to an ergodic orbit (x_n) on a curve, then we cannot use the fast Fourier transform as we only have the function values $f(x_n)$ along a quasiperiodic trajectory, and a rotation number ρ. Using interpolation to get the grid needed to apply a fast Fourier

(a)   (b)
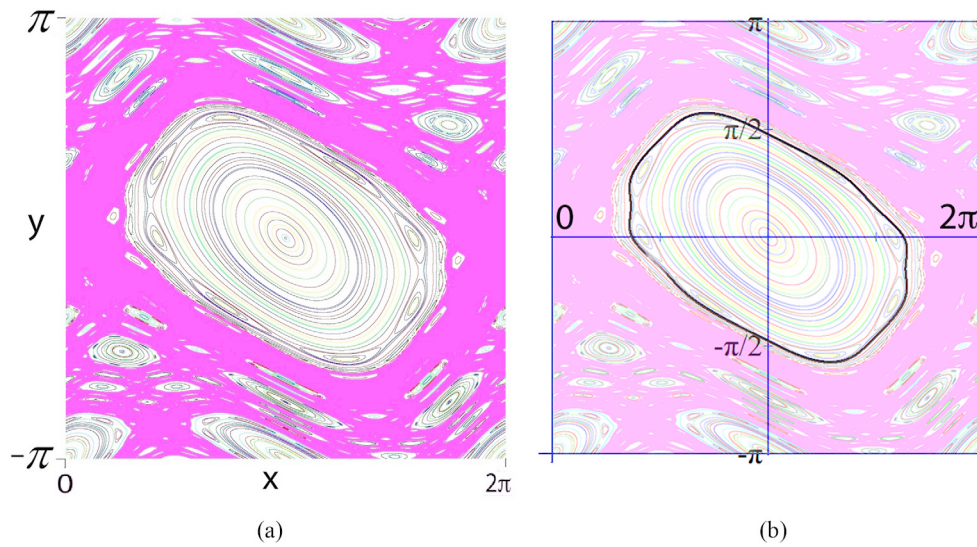
Figure 8. The standard map. Panel (a) shows a variety of orbits from different initial conditions in the standard map S_1 defined in equation (22). We can see both chaos (shaded area) and quasiperiodic orbits under this map. A single curve with quasiperiodic behavior is plotted in panel (b). The orbit has initial conditions $(x, y) \approx (-0.607, 2.01)$. That is, if we restrict the map to this invariant curve, then it appears to be topologically conjugate to a rigid irrational rotation.

transform introduces significant interpolation errors. So instead, we obtain these coefficients using a weighted Birkhoff average on a trajectory (x_n) by applying the functional WB_N. For $k = 0$, we find a_0 by applying WB_N to the function 1. Note that for all k, $a_{-k} = \bar{a}_k$. For $k > 0$, we find a_k as follows:

$$a_k = \mathrm{WB}_N(f(\theta)\mathrm{e}^{-\mathrm{i}2\pi k\theta}) = \sum_{n=0}^{N-1} f(x_n)\mathrm{e}^{-\mathrm{i}2\pi k n\rho}\hat{w}_{n,N}. \tag{20}$$

This is depicted for the R3BP in figure 4, for the standard map in figure 10, for the forced van der Pol equation in figure 11. In all three one-dimensional cases, we depict $|a_k|$ as a function of k for $k \geqslant 0$ only, as for all k, $|a_{-k}| = |a_k|$. Our main observation is that the Fourier coefficients decay exponentially; that is, for some positive numbers α and β, in dimension one, the Fourier coefficients satisfy

$$|a_k| \leqslant \alpha \mathrm{e}^{-\beta|k|} \text{ for all } k \in \mathbb{Z}. \tag{21}$$

This is characteristic of analytic functions. Therefore, the conjugacy functions of all our examples are effectively, analytic, 'effectively' meaning within the precision of our quadruple precision numerics. In two dimensions, the computation of Fourier coefficients is similar, but instead of only having one exponential functions, for each (j, k), we have two linearly independent sets of exponentials.

$$\mathrm{e}^{\mathrm{i}(jx+ky)} \text{ and } \mathrm{e}^{\mathrm{i}(jx-ky)}.$$

We define $a_{j,k}$ and $a_{j,-k}$ to be the complex-valued coefficients corresponding to these two functions.

3.7. Examples

3.7.1. The standard map. The standard map is an area preserving map on the two-dimensional torus, often studied as a typical example of analytic twist maps (see [29]). It is defined as follows

$$S_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ y + \alpha \sin(x + y) \end{pmatrix} \quad (\text{mod } 2\pi). \tag{22}$$

In this paper, we only consider the case $\alpha = 1.0$. Figure 8(a) shows the trajectories starting at a variety of different initial conditions plotted in different colors. The shaded set is a large invariant chaotic set with chaotic behavior, but many other invariant sets consist of one or more topological circles, on which the system has quasiperiodic behavior. For example, initial condition $(\pi, 1.65)$ yields chaos while $(\pi, 1.5)$ yields a quasiperiodic trajectory. As is clearly the case here, one-dimensional quasiperiodic sets often occur in families for non-linear processes, structured like the rings of an onion. There are typically narrow bands of chaos between quasiperiodic onion rings. Usually these rings are differentiable images of the *d*-torus. We have computed the rotation number to be $0.120\,552\,721\,973\,755\,133\,002\,981\,643\,698\,39$ for one such standard map orbit shown in the figure 8(b) using quadruple precision. Figure 9 shows a convergence rate of $O(N^{-8.75})$ in computing the rotation number using WB_N. Figure 10(a) shows the periodic parts *g* of conjugacies of the quasiperiodic orbit, and figure 10(b) shows the absolute values of Fourier coefficients representing an exponential decay.

3.7.2. The forced Van der Pol oscillator. Figure 11(a) shows attracting orbits for the time-$2\pi/0.83$ map of the following periodically forced Van der Pol oscillator with nonlinear damping [31]

$$\frac{d^2 x}{dt^2} - 0.2\left(1 - x^2\right)\frac{dx}{dt} + 20x^3 = F \sin(0.83t), \tag{23}$$

for several values of *F*. While the innermost orbit shown is chaotic, the outer orbits are topological circles with quasiperiodic behavior[7]. Each curve was assigned the angular coordinates of equation (13) by assigning the points on the curve the angle with respect to the origin $(0, 0)$. Figure 11(b) shows the periodic parts *g* of conjugacies of quasiperiodic orbits, and figure 11(c) shows the absolute values of Fourier coefficients representing exponential decays.

3.7.3. A two-dimensional torus map. So far, the quasiperiodic sets studied here are closed curves. We now describe an example [32–34] of a two-dimensional quasiperiodic torus map on $\mathbb{T}^2$. This is a two-dimensional version of Arnold's family of one-dimensional maps (see [35]). The map is given by (T_1, T_2) where

$$T_1(x, y) = \left[x + \omega_1 + \frac{\epsilon}{2\pi}P_1(x, y)\right] \quad (\text{mod } 1),$$
$$T_2(x, y) = \left[y + \omega_2 + \frac{\epsilon}{2\pi}P_2(x, y)\right] \quad (\text{mod } 1), \tag{24}$$

and $P_i(x, y), i = 1, 2$ are periodic functions with period one in both variables, defined by:

---

[7] As with the standard map, we have specified all non-essential parameters rather than stating the most general form of the Van der Pol equation. Our computed rotation numbers for the three orbits $F = 5.0, 15.0$, and 25.0 are $0.292\,061\,263\,291\,995\,892\,855\,775\,787\,189\,59$, $0.375\,534\,411\,131\,440\,108\,849\,089\,280\,833\,18$ and $0.562\,353\,700\,926\,850\,566\,344\,192\,213\,361\,54$ respectively.
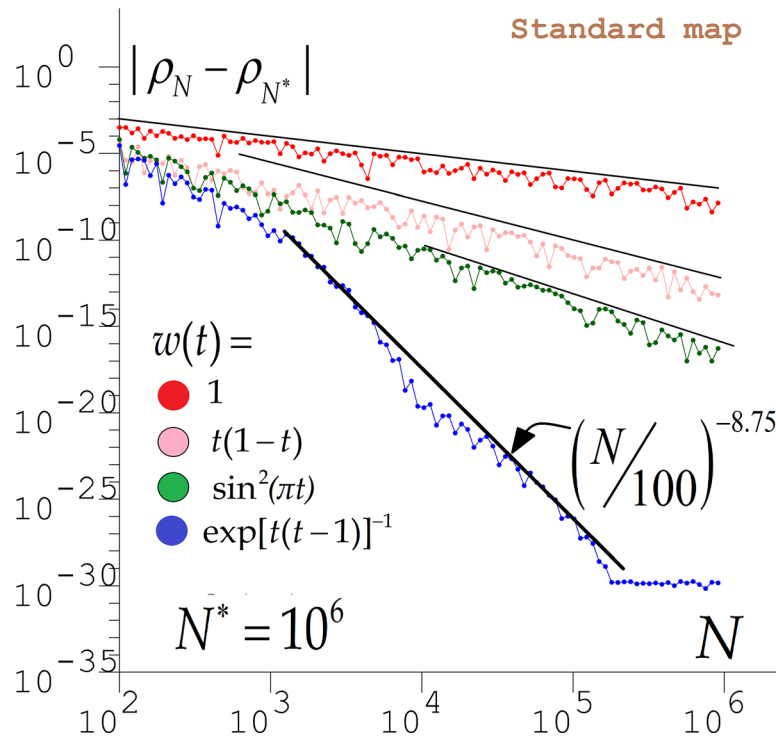
Figure 9. Rate of convergence in the rotation number for different weighting functions. This shows results for the standard map using the same trajectory for each plot. For a given w and a given number of iterates N, the rotation number $\hat{\rho}$ approximation is calculated for the curve all using the same trajectory. The plotted error of the calculation is the difference $|\rho - \hat{\rho}|$ as a function of N. The exponential weighting function w_{exp} allows WB_N to reach a limit by approximately $N = 150\,000$ at which point it is about 10^{16} times more accurate than the next best curve, the $\sin^2(\pi t)$ weighting. After that the error fluctuates by approximately 10^{-30}. The black line indicates the average slope of the bottom-most curve for large N and hence the approximate rate of convergence for the C^∞ weighted Birkhoff average. The $\sin^2(\pi t)$ weighting curve convergence rate is proportional to $N^{-3.0}$.

$$P_i(x, y) = \sum_{j=1}^{4} a_{i,j} \sin(2\pi \alpha_{i,j}), \text{ with } \alpha_{i,j} = r_j x + s_j y + b_{i,j}.$$

The values of all coefficients are given in table 1. This choice of this function is based on [32, 33]. The papers use the same form of equation, though the constants are close to but not precisely the same as the ones used previously. This fits with the point of view advocated by these papers: that the constants should be randomly chosen. Since we are using higher precision, we have chosen constants that are irrational to the level of our precision. The forward orbit is dense on the torus, and the map is a nonlinear map which exhibits two-dimensional quasiperiodic behavior.

Figure 12(a) depicts iterates of the orbit, indicating that it is dense in the torus. We use our weighted Birkhoff average to compute the two Lyapunov exponents, which have super convergence to zero. Figure 12(b) shows one of them. In terms of method, this is just a matter of
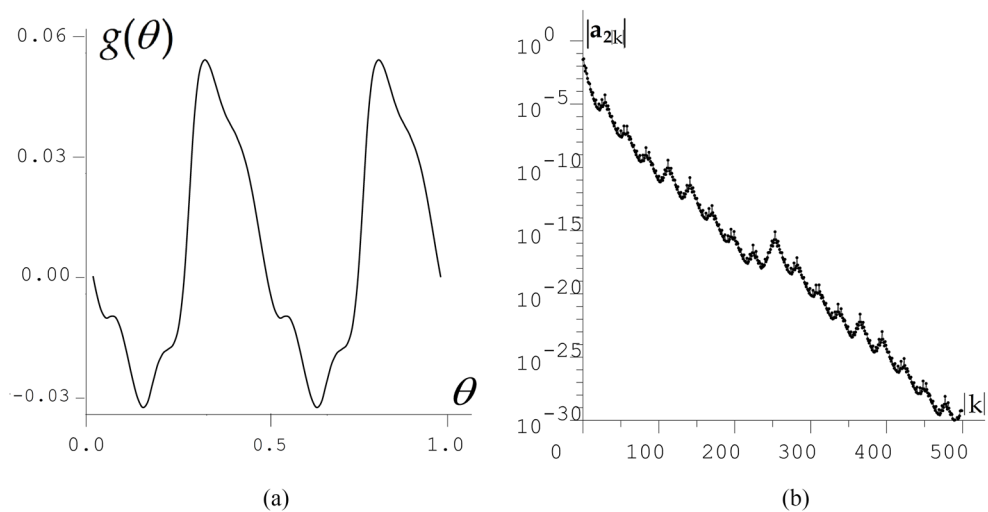
(a)   (b)

Figure 10. The standard map conjugacy. This figure shows the analysis of the quasiperiodic trajectory in figure 8. Panel (a) depicts the periodic part $g(\theta)$ of the conjugacy between the quasiperiodic behavior and rigid rotation by ρ, measured with respect to the center point $(x, y) = (\pi, 0)$. See equation (11) for a description of $g(\theta)$. Panel (b) shows the decay of the Fourier coefficients. Since the conjugacy is an odd function, the odd-numbered Fourier coefficients are zero and therefore have been omitted from the picture. The decay of the Fourier terms can be bounded from above be an exponential decay, which suggests that the conjugacy is analytic. An orbit of length $N = 10^7$ is used for these computations. A smaller orbit of length $N = 10^6$ does not lead to any significant changes.
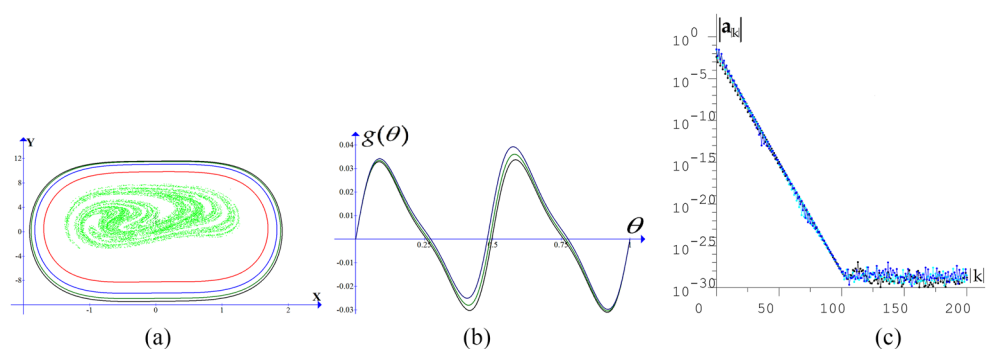

(a)   (b)   (c)

Figure 11. Forced Van der Pol oscillator. Panel (a) shows attracting orbits for a number of different forcing values F for the stroboscopic map of the van der Pol flow given in equation (23). The plot depicts points $(X, Y) = (x(t_k), x'(t_k))$, where $t_k = 2k\pi/0.83, k = 0, 1, 2, \ldots$. The chaotic orbit lying inside the cycles corresponds to $F = 45.0$. There are stable quasiperiodic orbits shown as curves, which from outermost to innermost correspond to $F = 5.0, 15.0, 25.0$ and 35.0 respectively. Panel (b) is the periodic part $g(\theta)$ of the conjugacy (equation (10)) to a rigid rotation, measured with respect to the center point $(0, 0)$ for $F = 5.0, 15.0,$ and 25.0. Panel (c) shows the absolute values of the Fourier coefficients of g. Their linear decay as $|k|$ increases in this log-linear plot means the coefficients decay exponentially fast, which is a signature of the analyticity g and hence of the conjugacy to a rigid rotation. The decay is exponentially fast down to the resolution of the numerics.

Table 1. Coefficients for the torus map. All values are used in quadruple precision, but in this table the repeated zeros on the end of the number are suppressed.

Coefficient	Value
ϵ	0.423 4823
ω_1	0.711 511 344 577 763 622 646 812 066 970 062 38
ω_2	0.877 350 098 112 614 561 009 170 866 728 499 71
$a_{1,j}$	$(-0.268, -0.9106, 0.3, -0.04)$
$a_{2,j}$	$(0.08, -0.56, 0.947, -0.4003)$
$b_{1,j}$	$(0.985, 0.504, 0.947, 0.2334)$
$b_{2,j}$	$(0.99, 0.33, 0.29, 0.155)$
r_j	$(1, 0, 1, 0)$
s_j	$(0, 1, 1, -1)$
Computed ρ_1	0.718 053 759 982 066 107 095 244 936 117
Computed ρ_2	0.885 304 666 596 099 792 113 366 824 157

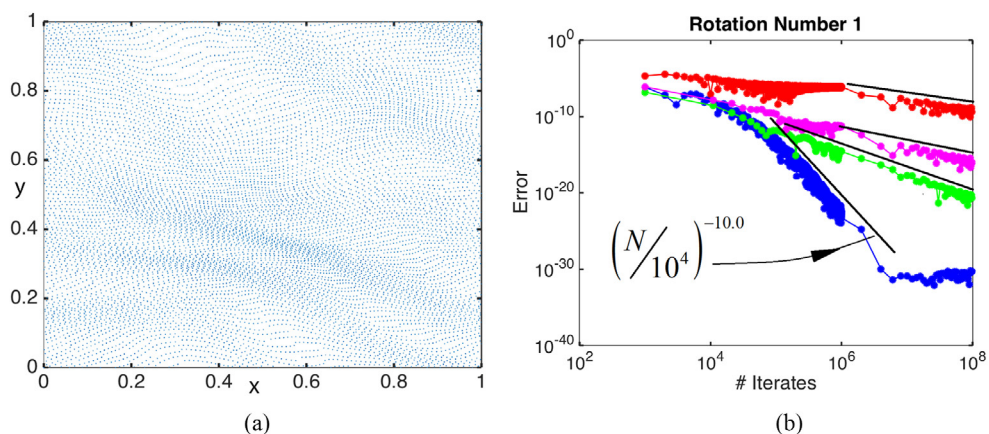(a)   (b)

Figure 12. Two-dimensional torus map. Panel (a) shows an orbit of length 10^4 for the two-dimensional quasiperiodic torus map. The orbit appears to be dense, which is consistent with quasiperiodicity. Panel (b) shows the convergence rate for the first rotation number for the four different weighting functions given in equation (15), respectively from highest curve to lowest, plotted in red, magenta, green, and blue (online). The black line is an upper bound for the blue curve and its exponent indicates the rate of convergence when using the C^∞ weight. It has a slope of -10 in the log-linear graph while the corresponding slope for the $\sin^2(\pi t)$ weighting function would be -3. That is, its rotation rate's error is proportional to N^{-3}.

changing the function f used in WB_N in equation (4). Likewise, finding rotation numbers in two dimensions uses the same technique as in the one-dimensional case (see figure 12(c)). In all of our calculations, the computation is significantly longer than in dimension one in order to get the same accuracy, perhaps because in dimension two sufficient coverage by a trajectory may vary like the square of the side length of the domain. Figure 12(b) shows a convergence rate of $O(N^{-10.0})$.

The reconstructed conjugacy function for the two-dimensional torus is depicted in figure 13(a). The decay of Fourier coefficients shown in figure 13(b) displays $\sqrt{j^2 + k^2}$ on the horizontal axis, and $|a_{j,k}|$ on the vertical axis, where both of these coefficients are complex,
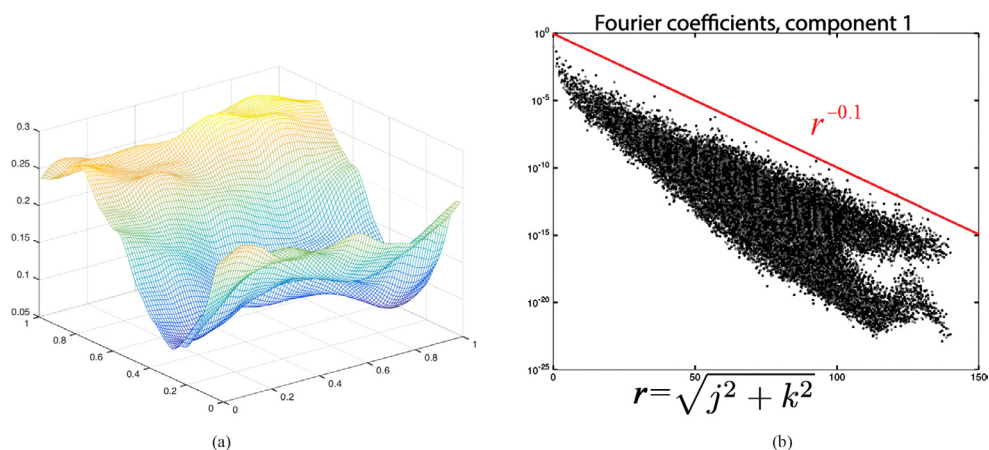
Figure 13. Conjugacy for the two-dimensional torus map. Panel (a) depicts the reconstruction of the periodic part g (see equation (11)) of the first component of the conjugacy function for the torus map. The surface is colored by height. The surface is created using the Fourier coefficients shown in (b). The second conjugacy function is similar but not depicted here. Panel (b) shows the norm of the Fourier coefficients $|a_{j,k}|$ and $|a_{j,-k}|$, for $0 \leqslant j, k \leqslant 100$ for the first component of the conjugacy function from panel (a).

meaning that $|\cdot|$ represents the modulus. Again here, the coefficients decay exponentially, though the decay of coefficients is considerably slower in two dimensions. The data set looks quite a lot more crowded in this case, since there are 10^4 coefficients and many different values of (j, k) have almost same values of $\sqrt{j^2 + k^2}$. In addition, the coefficients $a_{j,k}$ generally converge at different exponential rates. This is why there is a strange looking set consisting of an upper and a lower cloud of data in figure 13(b). While more information on the difference between these coefficients is gained by interactively viewing the data in three dimensions, we have not been able to find a satisfactory static flat projection of this data. We feel that in a still image, the data cloud shown conveys the maximum information.

3.7.4. The three-body problem revisited. We have computed trajectories for the Poincaré return map using an eighth order Runge–Kutta method with time step 2×10^{-5}, in quadruple precision and figure 4(c) is consistent with 30-digit accuracy of the rotation number (see item 1 below) for the quasiperiodic orbit of the three-body problem labeled B_1 in figure 3(b). The initial condition for this orbit is $(q_1, q_2, p_1, p_2) = (-0.15, 0, 0, p_2)$ where $p_2 \approx 5.41$ is chosen so that the Hamiltonian is about -2.63. While it is more straightforward to obtain a numerical trajectory when the quasiperiodic trajectory is stable, it is also possible when it is a saddle or a repeller. Our weighted Birkhoff average WB_N approach works equally well for both cases. The extent of convergence of WB_N is limited by the accuracy of the trajectory data. We now list results of our numerical methods for the restricted three-body problem.

1. The rotation number is $0.063\,961\,728\,757\,453\,097\,164\,077\,724\,400\,302$, computed to 30-digit accuracy, and Figure 4(c) shows the accuracy plateauing at about 30-digit accuracy.
2. We compute 200 terms of the Fourier series, the last 125 of which have magnitude near 0 (i.e. less than 10^{-30}). There is a conjugacy map h between the first return map and a rigid rotation on the circle. Evaluating the Fourier series allows us to reconstruct the conjugacy map (see figure 4(b)).

3. We find that the coefficients for the conjugacy map decrease exponentially fast (see figure 4(d)). That is a signature of a real analytic function as in figure 4(b).
4. The high rate of convergence of $O(N^{-15.0})$ for the rotation number in figure 4(c) suggests that we have an effective computational method that yields an accuracy that is close to the limit of numeric precision, provided N is sufficiently large.

3.7.5. Speed of convergence of Fourier series for a conjugacy.   In a separate report [36], we examine conjugacies of the quasiperiodic curves in the Siegel disk. The map is a simple one-dimensional complex dynamical system $z_{n+1} = f(z_n)$, where

$$f(z) = z^2 + e^{2\pi i \rho} z \tag{25}$$

and $\rho = (\sqrt{5} - 1)/2$ [9, 37].

It is found (see figure 3 in [36]) that the invariant curves get more irregular near the boundary of the disk. While typical smooth quasiperioidc curves require about 70 Fourier coefficients for 30-digit precision, near the boundary of the Siegel disk the curves are much more irregular and can require 24 000 coefficients (or more) for the same precision. The curves thus become increasingly fractal looking near the boundary and the Fourier series converges much more slowly. We remark that we would expect a similar slower convergence when exploring a case like the well known last KAM circle of the Standard Map.

3.7.6. Sources of error.   We end this section by noting a few sources of error in the computation of Fourier coefficients. If the number of iterates N is too small, then we will not have sufficient coverage to get a good approximation of the coefficients, and the problem becomes more acute as the coefficient number $|k|$ grows. If the approximation of the rotation number is not accurate, then we cannot expect the approximations of our Fourier coefficients to be good either, and given an error in the rotation number, there will be a $k_{\max}$ such that the Fourier coefficients a_k with $|k| > k_{\max}$ cannot be approximated with any reasonable accuracy. A more subtle form of a error comes from the fact that if the rotation number we are trying to estimate is close to being commensurate with the rotation number, then we will get unexpectedly insufficient coverage of the space when performing iteration. See section 4.2.

3.8. Lyapunov exponents computed as a weighted average

Lyapunov exponents are an important characterization of the dynamics resulting from the map T. They measure the rate at which nearby trajectories diverge or converge and can be used to distinguish between chaos and quasiperiodicity, for example, the existence of a positive Lyapunov exponent implies chaos, while for quasiperiodic systems, all d Lyapunov exponents for equation (1) are zero. In this section, we will show a way to obtain 'super-convergence' to the Lyapunov exponents of a quasiperiodic dynamics on a 2D torus. Since we are considering the dynamics restricted to the torus, we do not calculate the Lyapunov exponents for the normal sub-bundle.

Lyapunov exponents are usually calculated numerically as an average of logarithms with all terms weighted equally, and the computations are therefore limited by extremely slow convergence rates. For quasiperiodic systems we would expect the error in the exponents to be $O(1/N)$. For 2-dimensional quasiperiodic systems, the Lyapunov exponent can be expressed as a weighted Birkhoff average (using WB_N) and we observe that we obtain the Lyapunov exponents much faster. Figures 14(a) and (b) show convergence rates of $O(N^{-20.0})$ and $O(N^{-16.7})$ for the computation of the Lyapunov exponents using WB_N. (We make no claim
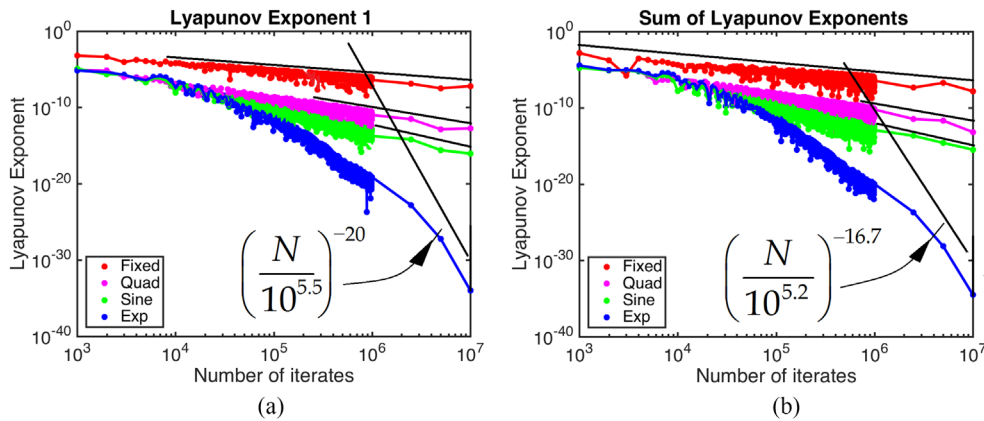
Figure 14. Lyapunov exponents computed with WB_N. The two Lyapunov exponents of the torus map from figure 12 are computed using WB_N and we find them to be 0 up to our numerical accuracy. (a) The computation for the first Lyapunov exponent is shown in blue for up to 10^7 iterates. The other three curves depict the values attained using the three other weighting functions, using the same order (and color scheme online) as in figure 12. Panel (b) shows the sum of the two Lyapunov exponents computed via direct computation of the determinant as described in equation (27). In the two panels, the $\sin^2(\pi t)$ graphs are approximately proportional to N^{-3}.

that we can prove convergence in (a) is super fast.) Using the weighted Birkhoff average does not change the limit.

Recall that given a d-manifold M and a map $T : M \to M$ with an invariant probability measure μ, Oseledets' multiplicative ergodic theorem (see [38]) states that there exists numbers $\lambda_1 \leqslant \ldots \leqslant \lambda_d$ such that for μ-almost every point x and every vector v in the tangent space at x, the limit

$$\lim_{N \to \infty} \frac{\ln \|DT^N(x)v\|}{N} \tag{26}$$

exists and equals one among $\lambda_1, \ldots, \lambda_d$. These numbers are the Lyapunov exponents of the map. For the rest of the section, we will assume that we have a quasiperiodic orbit (x_n) filling out a two-dimensional torus $\mathbb{T}^2$ and therefore $d = 2$ and $M = \mathbb{T}^2$.

One of the properties of the largest Lyapunov exponent λ_2 is that for almost every vector v in the tangent space at x_0, the quantity $N^{-1} \ln \|DT^N(x_0)v\|$ converges to λ_2. Note that this limit can be expressed as an average along the trajectory (x_n) in the following manner:

$$\frac{\ln \|DT^n(x_0)v\|}{N} = \sum_{n=0}^{N-1} \frac{\ln \|v_n\|}{N}, \text{ where } v_n = DT(x_{n-1}) \frac{v_{n-1}}{\|v_{n-1}\|}, v_0 = v.$$

We obtain the same limit by taking our weighted Birkhoff average instead of a uniformly weighted average as above. We observe (without proof) that for quasiperiodic orbits, we get super-convergence to the same limit by taking a weighted Birkhoff average in the following manner:

$$\lambda_2 = \lim_{N \to \infty} \sum_{n=0}^{N-1} \hat{w}_{n,N} \ln \|v_n\|, \text{ where } v_n = DT(x_{n-1}) \frac{v_{n-1}}{\|v_{n-1}\|}, v_0 = v.$$

See chapter 3 from [39] for an explanation of this method.

Once we have calculated λ_2, λ_1 can be obtained from the the sum $\lambda_1 + \lambda_2$ of the Lyapunov exponents. $\lambda_1 + \lambda_2$ can be expressed as the Birkhoff average of the function $f_1(x) := \ln|\det DT(x)|$. Note that $\det DT(x_n) > 0$ for an orientation preserving map T. Therefore, we have

$$\lambda_1 + \lambda_2 = \lim_{N \to \infty} \frac{\ln|\det DT^N(x)|}{N} = \lim_{N \to \infty} B_N(f_1).$$

As mentioned before, the quantity $B_N(f_1)$ converges to the limit $\int_X \ln|\det DT(x)|\mathrm{d}\mu(x)$ and we can obtain super-convergence to the same limit using the quantity $\mathrm{WB}_N(f_1)$ as written below:

$$\lambda_1 + \lambda_2 = \lim_{N \to \infty} \mathrm{WB}_N(f_1) = \lim_{N \to \infty} \sum_{n=0}^{N-1} \hat{w}_{n,N} \ln|\det DT(x_n)|. \tag{27}$$

3.9. Related methods

See [4–8, 10] for references to earlier methods for computing Birkhoff averages along a quasiperiodic orbit.

For higher dimensional quasiperiodicity ($d > 1$) Laskar [4] has an interesting technique of finding the frequency σ that maximizes a function $\phi(\sigma)$. That corresponds to our rotation rate ρ. Each evaluation of $\phi(\sigma)$ requires application of the window filter (whereas our method uses only one application of WB). But his has the advantage of subtracting off this frequency and repeating this method to find the next frequency. Hence he has an automatic method for finding multiple rotation rates. Presumably our method could be combined with his to improve on both methods—when dealing with higher dimensional quasiperiodicity.

Luque and Villanueva [8, 10] develop fast methods for obtaining rotation numbers for smooth or analytic functions on a quasiperiodic torus, sometimes with quasiperiodic forcing with several rotation numbers. The paper [10] develops a technique to compute rotation numbers (but not other function integrals) with error satisfying $|error| \leqslant C_p N^{-p}$ for $N \geqslant 2^p$ where C_p is a constant. Their method of computation is defined recursively, with the $p+1^{st}$ method being defined in terms of the pth. As p increases the computational complexity increases for fixed N. If $\tau(p, N_p)$ is their computation time when using $N = N_p$, it appears that $\tau(p, N_p)/N_p \to \infty$ as $p \to \infty$ (and $N_p \to \infty$). In comparison, computation time for our Weighted Birkhoff average is simply proportional to N since it requires a sum of N numbers. Figure 11 from [10] shows the rate of convergence to the rotation number for a quasiperiodic orbit arising from the R3BP. There, they get 30-digit accuracy for the rotation number using approximately $2\,000\,000$ trajectory points while we get the same accuracy with $20\,000$. The rate of convergence of their method is $\approx N^{-7.8}$. Their methods were extended in [40] to compute the derivatives of the rotation numbers for one-parameter families of circle diffeomorphisms, and in [41] to compute the Fourier coefficients of the conjugacy function (i.e. the change of variables between the map and rigid rotation).

3.9.1. Newton methods in the literature.

An alternative approach to our approach for finding a conjugacy is considered in [42–45]. In the current paper, we are assuming that we are starting with only a set of iterates for a single finite length forward trajectory, rather than having access to the functional form of the defining equation. In contrast, the approach in the papers above assumes access to the full form of the original defining equations. The Fourier series for the conjugacy is obtained and validated by using automatic differentiation. In addition,

in this current paper, we assume that we start with a point in an invariant torus with quasiperiodic dynamics, whereas the methods referenced above include a fast Newton's method for finding invariant tori, Lyapunov multipliers, and invariant stable and unstable manifolds. See also Jorba [46].

Several variants of the Newton's method have been employed to determine quasiperiodic trajectories in different settings. In [47] a variant of Newton's method was applied to locate periodic or quasi-periodic relative satellite motion in a non-linear, non-conservative setting.

A partial differential equation-based approach was taken in [48], where the authors defined an invariance equation which involves partial derivatives. The invariant tori are then computed using finite element methods. See also section 2 in [48] for more references on the numerical computation of invariant tori. In [49], the authors used an application of the Newton's method to compute elliptical, low dimensional invariant tori in Hamiltonian systems.

Computing the conjugacy to the rigid rotation is key to the methods of de la Llave *et al* [42–45]. They use this to obtain a numerical 'proof' of the existence of tori, and, what they call ' *a posteriori* KAM theory'.

4. Why our method works and when *N* must be large

We will assume that we have a C^∞ quasiperiodic map $F : \mathbb{T}^d \to \mathbb{T}^d$ and a C^M map $f : \mathbb{T}^d \to \mathbb{C}$. To better understand convergence on an averaging method applied to an f, write

$$f(\theta) = \sum_k a_k f_k(\theta), \text{ where } f_k(\theta) := e^{2\pi i k \cdot \theta}. \tag{28}$$

In particular $a_0 = \int_{\mathbb{T}^d} f$. WB_N is of course a weighted average of $(f(n\rho))_{n=0}^{N-1}$ and by theorem 3.1 $\text{WB}_N(f) \to a_0$ as $N \to \infty$. Since an averaging process is linear, we can define $\psi_{N,k,\rho} = \text{WB}_N(f_k)$, a collection of numbers that is independent of the choice of f. (More generally one could define $\psi_{N,k,\rho}^w$ for any weighting function w and the advantages of different choices of w are reflected in the magnitudes $|\psi_{N,k,\rho}^w|$). They depend only on the averaging method (e.g. WB_N), k, N, and the rotation number ρ. This set encapsulates all errors that arise in the use of WB_N. In particular, $f(\theta) = \sum_k a_k f_k(\theta)$, we have

$$\text{WB}_N(f) = a_0 + \sum_{k \neq 0} a_k \psi_{N,k,\rho}. \tag{29}$$

In particular $\psi_{N,0,\rho} = 1$ and for each $k \neq 0$, $\lim_{N \to \infty} \psi_{N,k,\rho} = 0$. The rate of convergence to 0 depends on ρ and k and convergence can be slow when $e^{2\pi i k \cdot \rho} - 1 \approx 0$, as we show in the next section.

To investigate the effects of the rotation number ρ on the numerical errors in calculating a Fourier coefficient, we show $\psi_{N,k,\rho} = \text{WB}_N(f_k)$ for each k ($0 \leqslant k \leqslant 250$) for two cases of $\rho = \sqrt{2} - 1$ and $\pi - 3$ in figure 15. The figure suggests that $N = 10^5$ is sufficient to calculate Fourier coefficients for $\rho = \sqrt{2} - 1$, whereas for $\rho = \pi - 3$ even $N = 10^5$ is not sufficient.

4.1. Sketch of proof of theorem 3.1

Here we sketch a proof that enables us to determine what happens when a rotation number is near a rational number.

Note that for any constant a_0, $\text{WB}_N(a_0 + f) = a_0 + \text{WB}_N(f)$ so for simplicity we will assume f has mean $a_0 = 0$. Let I denote the identity operator and U denote the Koopman operator on $L^2(X, \mu)$, defined as
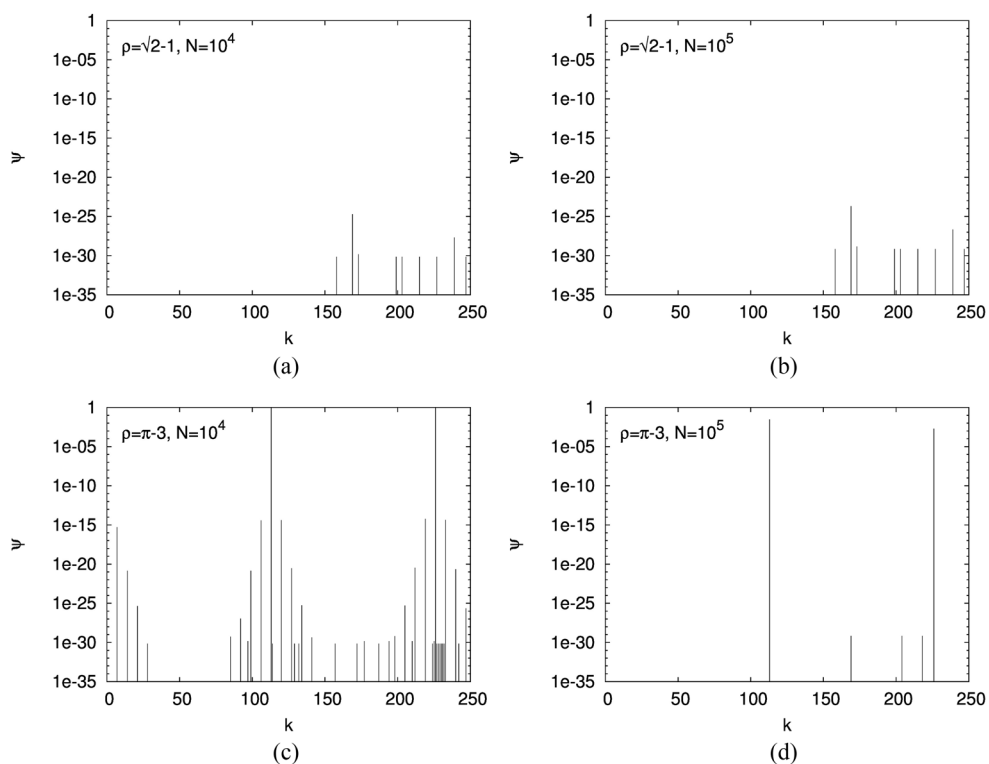
Figure 15. Estimations of the numerical error in the calculations of Fourier coefficients. Panels (a) and (b) show the value of $\psi_{k,N,\rho} = \mathrm{WB}_N(f_k)$ for $\rho = \sqrt{2} - 1$, when the number of iterations is $N = 10^4$ and 10^5, respectively. Panels (c) and (d) show for $\rho = \pi - 3$, when $N = 10^4$ and 10^5, respectively.

$$(Uf) : x \mapsto f(F(x)), \text{ for all } f \in L^2(X, \mu).$$

The idea of the proof is to provide two different estimates of the quantity

$$\mathrm{WB}_N(U - I)^m f = \frac{1}{N} \sum_n w\left(\frac{n}{N}\right)(U - I)^m f(n\rho). \tag{30}$$

First let $m = 1$ in equation (30). Then

$$\frac{1}{N} \sum_n w\left(\frac{n}{N}\right)(U - I)f(n\rho) = \frac{1}{N} \sum_n w\left(\frac{n}{N}\right) f((n+1)\rho) - \frac{1}{N} \sum_n w\left(\frac{n}{N}\right) f(n\rho)$$

$$= \frac{1}{N} \sum_n \left[w\left(\frac{n+1}{N}\right) - w\left(\frac{n}{N}\right)\right] f(n\rho).$$

Now taking absolute values on both sides give,

$$\left| \frac{1}{N} \sum_n w\left(\frac{n}{N}\right)(U - I)f(n\rho) \right| \leqslant \frac{1}{N} \|f\|_{C^0} \sum_n \left| \frac{1}{N} w^{(1)}\left(\frac{n}{N}\right) \right| \leqslant \frac{1}{N} \|f\|_{C^0} \|w^{(1)}\|_{C^0}.$$

Applying this procedure m times gives a constant $c_m > 0$ such that,

$$|\mathrm{WB}_N(U - I)^m f| \leqslant c_m N^{-m} \|f\|_{C^0}, \tag{31}$$

where c_m depends only on the first m derivatives of w.

The second way to evaluate equation (30) is by using the Fourier series for f. Note that f being C^M implies that there is a constant $C_{f,m} > 0$ such that $|a_k| \leqslant C_{f,m}\|k\|^{-M}$. If $k \neq 0$, then

$$\mathrm{WB}_N(U - I)^m f_k = (e^{2\pi i k \cdot \rho} - 1)^m \mathrm{WB}_N f_k.$$

Hence, by equation (31),

$$|\psi_{N,k}| = |\mathrm{WB}_N f_k| \leqslant c_m \left(\frac{1}{N \left| e^{2\pi i k \cdot \rho} - 1 \right|} \right)^m. \tag{32}$$

This implies

$$
\begin{aligned}
|\mathrm{WB}_N(f)| = \left| \sum_{k \neq 0} a_k \mathrm{WB}_N(f_k) \right| &\leqslant \sum_{k \neq 0} |a_k| \, |\mathrm{WB}_N(f_k)| \\
&\leqslant c_m \left(\frac{1}{N} \right)^m \sum_{k \neq 0} |a_k| \left| \frac{1}{e^{2\pi i k \cdot \rho} - 1} \right|^m.
\end{aligned}
\tag{33}
$$

The claim of the theorem now follows from equation (9) and the decay rate of $|a_k|$.

4.2. Difficulties when ρ is approximately rational.

While theorem 3.1 requires ρ to be Diophantine to get fast convergence of $\mathrm{WB}_N(f)$ to $\int f \, \mathrm{d}\mu$ (the integral in the Birkhoff ergodic theorem), our computations have finite precision so we have to ask what this condition means in a finite precision world. With this in mind, we can restate the Diophantine condition, in equation (9), saying ρ is Diophantine (of class β) if there is some $\beta > 0$ for which

$$\liminf_{|k| \to \infty} |k|^{d+\beta} \left| e^{2\pi i k \cdot \rho} - 1 \right| > 0.$$

In other words, the values of which k for which $|k|^d \left| e^{2\pi i k \cdot \rho} - 1 \right|$ should not go to 0 too fast. This idea suggests defining

$$\Delta(k, \rho) := |k|^d \left| e^{2\pi i k \cdot \rho} - 1 \right|.$$

For fast convergence in computing Fourier series coefficients a_k, the quantity $\Delta(k, \rho)$ should not be too small for the relevant k, those for which $|a_k|$ is likely to be larger than our error threshold, which in this paper is about 10^{-30}.

4.2.1. What values of $\Delta(k, \rho)$ are we likely to encounter?
For $\rho_1 := (\sqrt{5} + 1)/2$, the golden mean, we find

$$\min_{k=2,\cdots,10^6} \Delta(k, \rho_1) \approx 2.655.$$

It appears that $\liminf_{|k| \to \infty} \Delta(k, \rho_1) = 2.809\,925 \cdots$. Note the last term in equation (33) is similar to Δ except for the leading $|k|$. Of course the power $-m$ can greatly multiply the problem of $\left| e^{2\pi i k \cdot \rho} - 1 \right|$ being small. To offset $\left| e^{2\pi i k \cdot \rho} - 1 \right|$ being smaller by a factor of 100, we might expect that convergence would require N in equation (33) to be larger by a factor of 100.
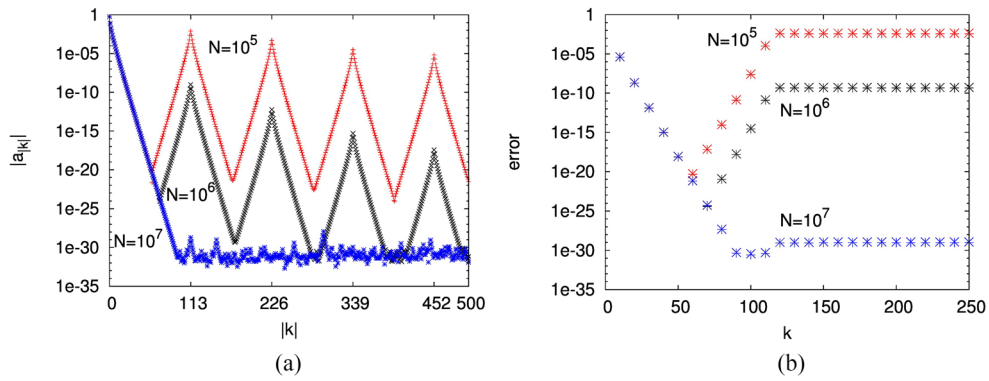
Figure 16. When ρ is near a rational. This figure is for the same case as figure 7(a), but with the rotation number $\rho = \pi - 3$. This rotation number is chosen because we require number of iterates N to be large to get accurate results. These plots show a slower convergence compared to the case in $\rho = \sqrt{2} - 1$. The number of iterates N used to calculate Fourier coefficients is changed from the top to the bottom ($N = 10^5$ (top), 10^6 (middle), 10^7 (bottom)). Panel (a) shows the norm of the Fourier coefficients a_k of the periodic part of the conjugacy map—as a function of index k. There are peaks at the multiples of 113 for $N = 10^5$ and 10^6. Panel (b) shows the agreement and the disagreement of the Fourier series and the original function in L^1 (+) and L^2 (×) norms in equation (34) for the corresponding N. Note that the L^1 and L^2 errors are so similar that the + and × overlap, yielding eight-legged spiders. The graph suggests that $N = 10^7$ is large enough to calculate Fourier coefficients, whereas for $N = 10^5$ and 10^6, it is best to include only 60 or 70 coefficients respectively, stopping when the error is minimum.

Of course both are raised to the same power m in these equations. Indeed in figure 16 we must increase N by a factor of 100 to get 30-digit convergence.

An illustration of the problem can be seen for $\rho_2 = \pi - 3$ since $\pi \approx \frac{355}{113}$, and $|k| = 113$ yields the rather small value $\Delta(\pm 113, \rho_2) \approx 0.021$. Also $\Delta(\pm 226, \rho_2) \approx 0.085$. $\Delta(\pm 339, \rho_2) \approx 0.193$. These suggest slow convergence for those k values. When computing a Fourier series for f using WB_N with this ρ, when $N < 10^7$, we obtain poor values for many Fourier coefficients as we illustrate in figure 16(a).

4.2.2. Comparing a function to its Fourier series. To estimate how accurate a computed Fourier series of a function $f : \mathbb{T}^d \to \mathbb{R}^1$ is, define $a_k = \mathrm{WB}_N(ff_{-k})$ and $f_k(\theta) := \mathrm{e}^{2\pi i k \cdot \theta}$ (where $k \cdot \theta$ denotes an inner product in $\mathbb{R}^d$), and

$$f^{K,N}(\theta) := \sum_{|k| \leqslant K} a_k f_k(\theta).$$

To test how similar f and $f^{K,N}$ are in the L^2 and L^1 sense, we compute the errors

$$\delta_2^{K,N}(f) := \sqrt{\mathrm{WB}_N((f - f^{K,N})^2)} \text{ and } \delta_1^{K,N}(f) := \mathrm{WB}_N(|f - f^{K,N}|). \qquad (34)$$

They are calculated in figure 16(b).

4.2.3. A saw-tooth pattern of errors. In the above example with $N = 10^5$, figure 16 shows a saw-tooth pattern with peaks at multiples of 113 where the slopes are all the same. Many coefficients have a big error, and all errors here are the result of $\psi_{N,k}$ when k is a non-zero multiple of 113. To explain this saw-tooth effect, we note that we have found for this example

that $\psi_{N,k} \approx 0$ for $N = 10^5$, when k is not a multiple of 113 and in particular $k \neq 0$, and here we assume all those values are indeed 0 to simplify computation except for $k = 0$ and $|k| = k^*$ for some $k^* \gg 1$ such as $k^* = 113$.

Suppose we wish to compute the Fourier coefficients of f and determine how accurate the result is. The computed kth coefficient, denoted $\hat{a}_k$ is

$$\hat{a}_{|k|} = \hat{a}_{\pm k} = \mathrm{WB}_N(f(\theta)\mathrm{e}^{\pm 2\pi \mathrm{i} k \cdot \theta}) = \sum_m a_m \psi_{N,m \pm k} \tag{35}$$

which has significant contributions from $m \pm k = 0$ and $|m \pm k| = k^*$. Since $k^* \gg 1$, we can ignore $a_{\pm k}$, and we conclude that for small integers n,

$$\hat{a}_{|k^* \pm n|} \approx a_{|n|}\psi_{N,k^*}.$$

On the log-linear plot of the graph, the coefficients a_k are almost linear, so it follows from this equation that the erroneous $\hat{a}_{|k^* \pm n|}$ has the same slope (the absolute value of the derivative), and that each non-zero multiple of $k^* = 113$ has the same error pattern. The heights of the peaks are $a_0 \psi_{N,mk^*}$ at $k = mk^*$ where $m = 1, 2, \cdots$.

4.2.4. Remark on figure 10. The panel (b) of this figure shows a jump in the Fourier series terms at coefficient $|a_{506}|$ (i.e. $k = 253$ in that figure). This does not seem to be a numerical artifact. In fact Δ does have a local minimum at 506 but $\Delta(506, \rho) = 1.03$ is not particularly small for such minimum. Furthermore, changing the number of iterates (as mentioned in the caption) does not change the graph.

5. Concluding remarks

We have developed a straightforward but effective computational tool for quickly computing a large variety of quantities for quasiperiodic orbits. These quantities include rotation vectors, Fourier reconstruction of conjugacy maps, and in some cases Lyapunov exponents. The methods work well in one and higher dimensions. They are effective using both double and quadruple precision, though we have chosen to do most of our calculations in higher precision to show the full possibilities and quick convergence properties of our method.

The literature on quasiperiodicity is vast and windowing techniques analogous to ours are often used. But our goals in this paper are limited: to introduce the C^∞ weighted Birkhoff averages WB_N $(=\mathrm{WB}_N^{[1]})$ and $\mathrm{WB}_N^{[2]}$ as numerically useful tools and to present some of its applications.

We note that the computational time for computing our weighting functions $w(t)$ is almost the same as for the weighting function $w_{\sin^2}$ that uses $\sin^2(\pi t)$. Both are equally easy to program. But convergence is far faster with the weighted Birkhoff averages, as seen in figures 7(b), 9, 12(b) and 14.

Quasiperiodic orbits can occur in many different situations, for example, subject to periodic forcing (see Luque and Villanueva [10]); as high-dimensional tori that are not simply embedded (see Medvedev *et al* [50]); in the presence of noise; etc. The question of whether our methods extend to these situations is worthy of further consideration.

5.1. When must N be large to get convergence?

We have developed some diagnostics in section 4.2 to detect when N must be chosen especially large to get high accuracy—at least for the $d = 1$ dimensional cases. Computation of

Δ can be used to detect cases when N must be large to get accurate values for Fourier coefficients. For example we found that because $\Delta(\pm 113, \rho_2)$ is so small, N must be increased by a factor of 100 to get an accurate Fourier series for the conjugacy map. One might ask if there are other k for which $\Delta(\pm 113, \rho_2)$ is quite small. We find $\Delta(\pm 113, \rho_2) < \Delta(k, \rho_2)$ for all $k \neq 113$ and $k < 10^7$.

Acknowledgments

ORCID iDs

Suddhasattwa Das ⓘ https://orcid.org/0000-0003-2085-4743

References

[1] Sander E and Yorke J A 2015 The many facets of chaos *Int. J. Bifurcation Chaos* **25** 15300
[2] Newhouse S, Ruelle D and Takens F 1978/79 Occurrence of strange Axiom A attractors near quasiperiodic flows on $T^m m \geqslant 3$ *Commun. Math. Phys.* **64** 35–40
[3] Das S and Yorke J A 2015 Super convergence of ergodic averages for quasiperiodic orbits (arXiv:1506.06810 [math.DS])
[4] Laskar J 1999 Introduction to frequency map analysis *Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995)* vol 533 (Dordrecht: Kluwer) pp 134–50
[5] Laskar J 1993 Frequency analysis of a dynamical system *Celest. Mech. Dyn. Astron.* **56** 191–6
[6] Laskar J 1993 Frequency analysis for multi-dimensional systems. Global dynamics and diffusion *Physica* D **67** 257–83
[7] Laskar J 2003 Frequency map analysis and particle accelerators *Proc. 20th IEEE Particle Accelerator Conf. (Portland, Oregon, 2–16 May 2003) (*IEEE) pp 378–82
[8] Seara T M and Villanueva J 2006 On the numerical computation of Diophantine rotation numbers of analytic circle maps *Physica* D **217** 107–20
[9] Luque A and Villanueva A J 2009 Numerical computation of rotation numbers of quasi-periodic planar curves *Physica* D **238** 2025–44
[10] Luque A and Villanueva J 2014 Quasi-periodic frequency analysis using averaging-extrapolation methods *SIAM J. Appl. Dyn. Syst.* **13** 1–46
[11] Simó C 1994 Averaging under fast quasiperiodic forcing *Hamiltonian Mechanics (Toruń, 1993)* vol 331 (New York: Plenum) pp 13–34
[12] Simó C, Sousa-Silva P and Terra M 2013 *Practical Stability Domains near $L_{4,5}$ in the Restricted Three-Body Problem: Some Preliminary Facts* vol 54 (Berlin: Springer)
[13] Durand F and Schneider D 2002 Ergodic averages with deterministic weights *Ann. Inst. Fourier* **52** 561
[14] Baesens C, Guckenheimer J, Kim S and MacKay R S 1991 Three coupled oscillators: mode-locking, global bifurcations and toroidal chaos *Physica* D **49** 387–475
[15] Broer H W and Huitema G B 1995 Unfoldings of quasi-periodic tori in reversible systems *J. Dyn. Differ. Equ.* **7** 191–212
[16] Vitolo R, Broer H and Simó C 2011 Quasi-periodic bifurcations of invariant circles in low-dimensional dissipative dynamical systems *Regul. Chaotic Dyn.* **16** 154–84

[17] Hanßmann H and Simó C 2012 Dynamical stability of quasi-periodic response solutions in planar conservative systems *Indag. Math.* **23** 151–66

[18] Sevryuk M B 2012 Quasi-periodic perturbations within the reversible context 2 in KAM theory *Indag. Math.* **23** 137–50

[19] Broer H W, Huitema G B, Takens F and Braaksma B L J 1990 Unfoldings and bifurcations of quasi-periodic tori *Mem. Am. Math. Soc.* **83** viii–175

[20] Kuznetsov A P, Migunova N A, Sataev I R, Sedova Y V and Turukina L V 2015 From chaos to quasi-periodicity *Regul. Chaotic Dyn.* **20** 189–204

[21] Das S, Dock C, Saiki Y, Salgado-Flores M, Sander E and Yorke J A 2016 Measuring quasiperiodicity *Europhys. Lett.* **114** 40005

[22] Das S, Saiki Y, Sander E and Yorke J A 2017 Solving the Babylonian Problem of quasiperiodic rotation rates (arXiv:1706.02595)

[23] Poincaré H 1905 *Leçons de Mécanique Céleste : Professées à la Sorbonne* (Paris: Gauthier-Villars)

[24] Greene J B 1996 *Poincaré and the Three Body Problem* (Providence, RI: American Mathematical Society)

[25] Szebehely V 1967 *Theory of Orbits: the Restricted Problem of Three Bodies* (Cambridge, MA: Academic)

[26] Markley N G 1968 Transitive homeomorphisms of the circle *Math. Syst. Theory* **2** 247–9

[27] Van Kampen E R 1935 The topological transformations of a simple closed curve into itself *Am. J. Math.* **57** 142–52

[28] Herman M R 1979 Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations *Publ. Math. Inst. Hautes Études Sci.* **49** 5–233

[29] Yamaguchi Y and Tanikawa K 1999 A remark on the smoothness of critical KAM curves in the standard mapping *Prog. Theor. Phys.* **101** 1–24

[30] Hunt B, Khanin K A, Sinai Y G and Yorke J A 1996 Fractal properties of critical invariant curves *J. Stat. Phys.* **85** 261–76

[31] van der Pol B 1920 A theory of the amplitude of free and forced triode vibrations *Radio Rev.* **1** 701–10

[32] Grebogi C, Ott E and Yorke J A 1983 Are three-frequency quasiperiodic orbits to be expected in typical nonlinear dynamical systems? *Phys. Rev. Lett.* **51** 339

[33] Grebogi C, Ott E and Yorke J A 1985 Attractors on an N-torus: quasiperiodicity versus chaos *Physica* D **15** 354–73

[34] Kim S, MacKay R S and Guckenheimer J 1989 Resonance regions for families of torus maps *Nonlinearity* **2** 391–404

[35] Arnold V 1965 Small denominators. I. Mapping of the circumference onto itself *Am. Math. Soc. Transl. (2)* **46** 213–84

[36] Das S, Saiki Y, Sander E and Yorke J A 2017 Quasiperiodic orbits within a Siegel disk and a Siegel ball in complex dynamical systems (unpublished)

[37] de la Llave R and Petrov N P 2008 Boundaries of Siegel disks: numerical studies of their dynamics and regularity *Chaos* **18** 033135, 1–11

[38] Raghunathan M S 1979 A proof of Oseledec's multiplicative ergodic theorem *Isr. J. Math.* **32** 356–62

[39] Alligood K T, Sauer T D and Yorke J A 1997 *Chaos* (Berlin: Springer)

[40] Luque A and Villanueva J 2008 Computation of derivatives of the rotation number for parametric families of circle diffeomorphisms *Physica* D **237** 2599–615

[41] Seara T M and Villanueva J 2009 Numerical computation of the asymptotic size of the rotation domain for the Arnold family *Physica* D **238** 197–208

[42] Haro A and de la Llave R 2007 A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: explorations and mechanisms for the breakdown of hyperbolicity *SIAM J. Appl. Dyn. Syst.* **6** 142–207

[43] Cabré X, Fontich E and de la Llave R 2003 The parameterization method for invariant manifolds. I. Manifolds associated to non-resonant subspaces *Indiana Univ. Math. J.* **52** 283–328

[44] de la Llave R, González A, Jorba À and Villanueva J 2005 KAM theory without action-angle variables *Nonlinearity* **18** 855–95

[45] Huguet G, de la Llave R and Sire Y 2013 Fast iteration of cocycles over rotations and computation of hyperbolic bundles *Discrete Contin. Dyn. Syst.* **2** 323–33

[46] Jorba À 2001 Numerical computation of the normal behaviour of invariant curves of n-dimensional maps *Nonlinearity* **14** 943976

[47] Becerra V M, Biggs J D, Nasuto S J, Ruiz V F, Holderbaum W and Izzo D 2006 Using Newton's method to search for quasi-periodic relative satellite motion based on nonlinear Hamiltonian models *7th Int. Conf. on Dynamic and Control of Systems and Structures in Space* vol 7

[48] Schilder F, Osinga H M and Vogt W 2005 Continuation of quasi-periodic invariant tori *SIAM J. Appl. Dyn. Syst.* **4** 459–88

[49] Luque A and Villanueva A J 2011 A KAM theorem without action-angle variables for elliptic lower dimensional tori *Nonlinearity* **24** 1033

[50] Medvedev A G, Neishtadt A I and Treschev D T 2015 Lagrangian tori near resonances of near-integrable Hamiltonian systems *Nonlinearity* **28** 2105–30