Universitat de Barcelona

Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

SUPERCONVERGENCE OF WEIGHTED BIRKHOFF AVERAGES FOR QUASIPERIODIC ORBITS

Autor: Sara Xiang Ru Guaitoli

Director: Dr. Alex Haro Provinciale

Realitzat a: Facultat de Matemàtiques i Informàtica

Barcelona, 10 de juny de 2024

Contents

Abstract

A quasiperiodic trajectory $(x_n) \subset X_0$, where X_0 is a d-dimensional differentiable manifold, is characterised by a diffeomorphism $h : \mathbb{T}^d \to X_0$ and an irrational vector $\rho \in \mathbb{R}^d$ such that $x_n = h(n\rho \pmod 1)$. We will see that ρ and the Fourier coefficients of h can be expressed as a limit of a Birkhoff average, i.e. an average over a certain function along the trajectory (x_n). The Birkhoff Ergodic Theorem provides us the convergence of these averages $B_N(f) := \sum_{n=0}^{N-1} f(x_n)/N$ to the space average as $N \to \infty$.

Sad to say this convergence is slow. The main goal of this work is to show that if we modify the Birkhoff average by weighting each term such that the early and late terms of the set $\{0, ..., N-1\}$ are weighted much less than the terms with $n \sim N/2$ in the middle, the weighted average converges far faster to the space average. Hence, this numerical technique will allow us to obtain efficient numerical computation of ρ and h for quasiperiodic systems. This work is based on the research of S. Das, Y. Saiki, E. Sander and J. Yorke.

Our work proceeds as follows: Chapter 1 presents the formal definition of quasiperiodicity, Chapter 2 explains the necessary concepts for stating the Birkhoff Ergodic Theorem, Chapter 3 provides a detailed description of the superconvergence numerical technique of the weighted Birkhoff averages, and the last chapters show how this method applies to compute ρ and h alongside numerical examples.

Keywords: Quasiperiodicity, Birkhoff Ergodic Theorem, Rotation Vector.

2020 Mathematics Subject Classification. 37E10, 37E45, 37M21, 37M25, 65P99, 70K43

Resum

Una trajectòria quasiperiòdica $(x_n) \subset X_0$, on X_0 és una varietat diferencial d-dimensional, es caracteritza per un difeomorfisme $h : \mathbb{T}^d \to X_0$ i un vector irracional $\rho \in \mathbb{R}^d$ tal que $x_n = h(n\rho \text{ (mod 1)})$. Al llarg del treball, veurem que ρ i els coeficients de Fourier de h es poden expressar com un límit d'una mitjana de Birkhoff, és a dir, una mitjana sobre una determinada funció al llarg de la trajectòria (x_n). El Teorema Ergòdic de Birkhoff ens proporciona la convergència d'aquestes mitjanes $B_N(f) := \sum_{n=0}^{N-1} f(x_n)/N$ a la mitjana espacial quan $N \to \infty$.

Malgrat tot, aquesta convergència és lenta. L'objectiu principal d'aquest treball és demostrar que si modifiquem la mitjana de Birkhoff ponderant cada terme de manera que els termes inicials i finals del conjunt $\{0, ..., N-1\}$ tinguin una importància inferior que els termes amb $n \sim N/2$ al mig, la mitjana ponderada convergeix molt més ràpid a la mitjana espacial. Per tant, aquest mètode numèric ens permetrà obtenir una bona approximació de ρ i h per a sistemes quasiperiòdics. Aquest treball es basa en la recerca de S. Das, Y. Saiki, E. Sander i J. Yorke.

El nostre treball procedeix de la següent manera: el Capítol 1 conté la definició formal de quasiperiodicitat, el Capítol 2 explica els conceptes necessaris per enunciar el Teorema Ergòdic de Birkhoff, el Capítol 3 proporciona una descripció detallada del mètode numèric de superconvergència de les mitjanes ponderades de Birkhoff i els darrers capítols mostren com s'aplica aquest mètode per calcular ρ i h juntament amb exemples numèrics.

Paraules claus: Quasiperiodicitat, Teorema Ergòdic de Birkhoff, Vector de Rotació.

Acknowledgements

I wanted to thank my thesis director Dr. Alex Haro, without him this work would not have been possible. Thank you for your passion for teaching, for all the hours you have dedicated to my work and for making me enjoy the whole process.

Chapter 1

Quasiperiodicity

It has been hypothesised in [1] that for typical randomly chosen physical systems, there are only three kinds of maximal recurrent sets that are likely to be present: periodic orbits, quasiperiodic orbits and chaotic sets. For instance, a system as elementary as the double pendulum with zero air friction, depicted in the following Figure 1.1, can present all three behaviours. Note that the outcomes may differ depending on the specific characteristics of the double pendulum in question and the initial state, characterised by the initial angles and angular velocities.

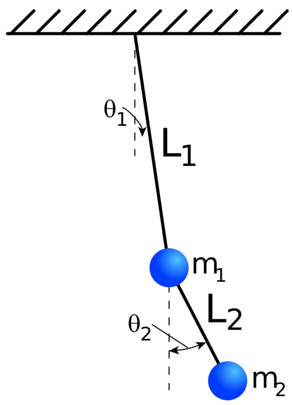


Figure 1.1: Double pendulum.

In this work, we will focus on quasiperiodicity. Quasiperiodic dynamics hold significant importance across many scientific disciplines. In mathematics, they provide deep insights into the structure of dynamical systems and their long-term evolution. In physics, quasiperiodic motion can be found everywhere: from celestial mechanics to solid-state physics. In material science, quasiperiodicity can be spotted in quasicrystals, which are the materials with quasiperiodic atomic arrangements, that have unique mechanical, thermal, and electric properties. Over-all, quasiperiodic dynamics are important for understanding the richness and complexity of natural phenomena and for extending mathematical abstraction over real-world applications. Currently, there is a significant amount of pioneering research across several scientific disciplines which study this dynamics.

In this first chapter, we start by explaining which are the conditions such that, given a trajectory (x_n), we can establish it has quasiperiodic motion. Most of the definitions seen in this chapter are sourced from [2].

1.1 Irrational Rotations

Throughout this section, we will introduce irrational rotations, which represent the most basic form of quasiperiodic maps and provide an intuitive understanding of quasiperiodicity. Mastering the properties of these maps serves as a foundation for exploring more intricate quasiperodic maps. We will denote the *d-dimensional torus* as $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d \simeq (\mathbb{R}/\mathbb{Z})^d$.

Definition 1.1. Let $\rho = (\rho_1, \ldots, \rho_d) \in \mathbb{T}^d$. A (pure) *rotation* on the d-dimensional torus $\mathbb{T}^d$ is a map defined as:

$$\begin{array}{rccl} T_\rho : & \mathbb{T}^d & \longrightarrow & \mathbb{T}^d \\ & \theta & \longmapsto & \theta + \rho \text{ mod } 1 \text{ in each coordinate} \end{array}$$

We call ρ the *rotation vector* and each ρ_j a *rotation number*, for $j = 1, \ldots, d$. Notice that we are rotating each coordinate θ_j by its correspondent angle ρ_j. In some papers, this map is referred to as a *translation* by ρ.

In some cases, it is useful to consider $\rho \in \mathbb{R}^d$, and then take modulo 1 when computing the image.

Definition 1.2. We say $\rho = (\rho_1, \ldots, \rho_d) \in \mathbb{R}^d$ is *irrational* (or *non-resonant*) if $\forall k = (k_1, \ldots, k_d) \in \mathbb{Z}^d \setminus \{0\}$, $k \cdot \rho := k_1\rho_1 + \ldots + k_d\rho_d \notin \mathbb{Z}$. If the condition is not satisfied, we will say that ρ is *resonant*.

Example 1.3. The vector $(\sqrt{2}, \sqrt{2}) \in \mathbb{R}^2$ is irrational (or non-resonant), while the vector $(1, \sqrt{2}) \in \mathbb{R}^2$ is resonant.

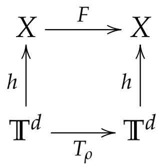
Definition 1.4. A (pure) *irrational rotation* is a rotation with an irrational rotation vector.

Theorem 1.5. *Let $T_\rho : \mathbb{T}^d \to \mathbb{T}^d$ be a rotation. T_ρ is an irrational rotation iff each trajectory is dense in $\mathbb{T}^d$.*

Proof. See proof in [3], Proposition 1.4.1. □

1.2 The Conjugacy Problem

Definition 1.6. Let X be a C^r-manifold of dimension d and T_ρ be an irrational rotation on the d-dimensional torus. A map $F : X \to X$ is *quasiperiodic with irrational rotation vector* ρ if there exists a C^r-diffeomorphism $h : \mathbb{T}^d \to X$ such that $F(h(\theta)) = h(T_\rho(\theta))$ for all $\theta \in \mathbb{T}^d$. Equivalently, if the following diagram commutes:



Note that this implies that F is also a C^r-diffeomorphism. We refer to h as the *C^r-conjugacy of F to T_ρ*.

Definition 1.7. Let X be a C^r-manifold of dimension $n \geq d$ and $X_0 \subseteq X$ be a d-dimensional invariant manifold with respect to F, i.e. $F(X_0) \subseteq X_0$. We will say that $F : X \to X$ is *quasiperiodic with irrational rotation vector ρ on $X_0 \subseteq X$* if there exists a conjugacy map $h : \mathbb{T}^d \to X$ such that $F(h(\theta)) = h(T_\rho(\theta))$ for all $\theta \in \mathbb{T}^d$ and, in addition, $h(\mathbb{T}^d) = X_0 \subseteq X$ and $h : \mathbb{T}^d \to X_0$ is a C^r-diffeomorphism.

The following figure tries to illustrate this concept, which is the scenario in most of the cases. As h acts as a parametrisation of X_0, we will sometimes refer to X_0 as a d-dimensional invariant torus embedded in X.

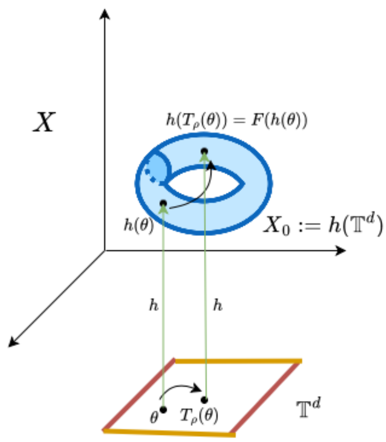


Figure 1.2: Main idea of the parametrisation h for the invariant manifold X_0, which presents quasiperiodic behaviour. Notice that in the angle coordinates $\theta \in \mathbb{T}^d$, the dynamics are a rotation by ρ.

The main challenge in proving the existence of a quasiperiodic behaviour lies in parameterising $X_0 \subseteq X$ such that the dynamics of the dynamical system are an irrational rotation. Formally, to establish that a d-dimensional trajectory $(x_n) \in X_0$ is quasiperiodic, it is sufficient to find a conjugacy map $h : \mathbb{T}^d \to X$ and an irrational vector $\rho \in \mathbb{T}^d$ such that $x_n = h(n\rho \pmod 1)$. It is important to realise that an irrational rotation is automatically a quasiperiodic map. The problem of finding h and ρ, knowing only the trajectory (x_n), is referred to as the *Conjugacy Problem*.

Theorem 1.8. *Let X be a C^r-manifold, $F : X \to X$ be a C^r-map, and $X_0 \subseteq X$ be an invariant C^r-manifold of dimension d. Assume the dynamics of $F|_{X_0}$ are known to be C^r-conjugated to a rotation T_ρ. F is quasiperiodic on X_0, i.e. ρ is irrational, iff each trajectory under F is dense in X_0.*

Proof. Let $x_0 \in X_0$ and consider its trajectory under F denoted as $(x_n := F^n(x_0))_{n\in\mathbb{Z}_{\geq 0}}$. The trajectory is quasiperiodic iff there exists a C^r-diffeomorphism $h : \mathbb{T}^d \to X_0$ and an irrational vector ρ such that $x_n = h(n\rho \mod 1)$. By hypothesis, we know there exists a C^r-conjugacy h of $F|_{X_0}$ to T_ρ. If we assume ρ is irrational, the sequence $(n\rho \mod 1)_{n\in\mathbb{Z}_{\geq 0}}$ follows an irrational rotation, so by Theorem 1.5 it is dense in $\mathbb{T}^d$. Since h is a conjugacy, this is equivalent to the fact that $(x_n)_{n\in\mathbb{Z}_{\geq 0}}$ is dense in X_0. □

Example 1.9. (Linear flow on a Torus) Consider a vector field $\mathcal{X}$ on $\mathbb{T}^2$ defined by:

$$\begin{array}{cccc} \mathcal{X}: & \mathbb{T}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x,y) & \longmapsto & (1,\alpha) \end{array}$$

where α is constant. Consider the Poincaré section $\sum := \{(0,y)|y \in \mathbb{T}\} \subset \mathbb{T}^2$. Recall we can represent the 2-dimensional torus $\mathbb{T}^2 = \{(x,y) \pmod 1 \,|(x,y) \in \mathbb{R}^2\}$ as a square with opposite sides identified, thus it is equivalent to consider $\sum = \{(1,y)|y \in \mathbb{T}\}$. Let $\pi_y : \mathbb{T}^2 \to \mathbb{T}$ be the projection onto the y-coordinate $\pi_y(x,y) = y$, and let $\varphi_1 : \sum \to \sum$ be the flow at time 1. Then, define $T : \mathbb{T} \to \mathbb{T}$ as $T(y) := \pi_y \circ \varphi_1(0,y) = y + \alpha \pmod 1$, which is a rotation by α.

If $\alpha = p/q \in \mathbb{Q}$ with $p,q \in \mathbb{Z}$ coprime, the orbit $\{T^n(y_0)\}_{n\geq 0}$ is periodic for any $y_0 \in \mathbb{T}$. Moreover, it returns to itself after q units of time. If instead $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, T is quasiperiodic. Notice that for different $n, m \in \mathbb{N}$, $T^n(y_0) = T^m(y_0)$ iff $n\alpha = m\alpha \pmod 1$ iff $(n-m)\alpha \in \mathbb{Z}$, which is impossible as α is irrational; therefore, each orbit is dense in $\mathbb{T}$. See Figure 1.3 and Figure 1.4.

1.3 Orientation-preserving Circle Homeomorphisms

In the previous section, we mentioned that to demonstrate the quasiperiodic nature of a trajectory (x_n), we must find an irrational vector ρ and a conjugacy

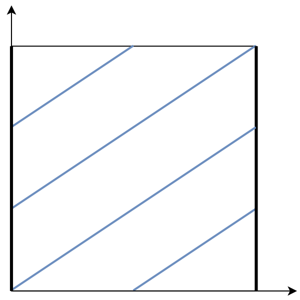


Figure 1.3: Exact orbit of $y_0 = 0$ under the flow T when $\alpha = 2/3 \in \mathbb{Q}$. It is 3-periodic.

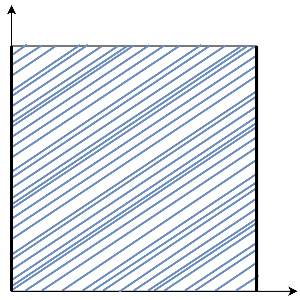


Figure 1.4: Approximation of the orbit of $y_0 = 0$ under the flow T when $\alpha = 1/\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$. It is dense.

map h satisfying $x_n = h(n\rho \pmod 1)$. For a more in-depth understanding of the Conjugacy problem, we will examine the case where $d = 1$ and see how these objects are computed.

The definitions and results we will present are well-known from Poincaré and most of them can be found in Katok and Hasselblatt's book [3]. Thoughout the section, we will refer to $\pi : \mathbb{R} \to \mathbb{R}/\mathbb{Z} =: \mathbb{T}$ as the projection $\pi(x) = x \pmod 1$ for $x \in \mathbb{R}$.

Definition 1.10. Let $T : \mathbb{T} \to \mathbb{T}$ be an homeomorphism. We say that an homemorphism $\hat{T} : \mathbb{R} \to \mathbb{R}$ is a *lift* of T if satisfies $T \circ \pi(x) = \pi \circ \hat{T}(x)$ for all $x \in \mathbb{R}$. Equivalently,

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{\hat{T}} & \mathbb{R} \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\ \mathbb{T} & \xrightarrow[T]{} & \mathbb{T} \end{array}$$

It is unique up to an additive integer constant.

Definition 1.11. An homeomorphism $T : \mathbb{T} \to \mathbb{T}$ is *orientation-preserving* if there exists a lift that is monotonically increasing.

Recall that any homeomorphism of $\mathbb{R}$ is monotone, hence this gives a notion of "preserving/swapping the orientation".

Lemma 1.12. *Let $T : \mathbb{T} \to \mathbb{T}$ be an orientation-preserving homeomorphism. Then:*

(1) There exists a lift $\hat{T} : \mathbb{R} \to \mathbb{R}$ of T.

(2) $\hat{T}$ is strictly increasing.

(3) $\hat{T}(x+k) = T(x) + k, \forall x \in \mathbb{R}, k \in \mathbb{Z}$.

(4) $G(x) := \hat{T}(x) - x$ *is* 1*-periodic in* $x \in \mathbb{R}$.

Proof. Property (4) can be proven as follows: $G(x+1) := \hat{T}(x+1) - (x+1) = \hat{T}(x) + 1 - x - 1 = \hat{T}(x) - x =: G(x)$, where we have used Property (3). □

Example 1.13. Consider the *Arnold (circle) family* defined as

$$\begin{array}{rccl} A_{\alpha,\epsilon}: & \mathbb{T} & \longrightarrow & \mathbb{T} \\ & \theta & \longmapsto & \theta + \alpha - \frac{\epsilon}{2\pi} sin(2\pi\theta) \text{ (mod 1)} \end{array}$$

for some constants $\alpha, \epsilon \in \mathbb{R}$. Its lifts are the form of $\hat{A}_{\alpha,\epsilon,K}(x) = x + \alpha - \frac{\epsilon}{2\pi} sin(2\pi x) + K$ for $x \in \mathbb{R}$, where $K \in \mathbb{Z}$.

Proposition 1.14. (Rotation Number) *Let* $T : \mathbb{T} \to \mathbb{T}$ *be an orientation-preserving homeomorphism and consider one of its lifts* $\hat{T} : \mathbb{R} \to \mathbb{R}$*. Let* $\theta_0 \in \mathbb{T}$ *and define*

$$\rho(\hat{T}) := \lim_{N\to\infty} \frac{1}{N}(\hat{T}^N(\theta_0) - \theta_0)$$

Then, $\rho(\hat{T}) \in \mathbb{R}$*, it is independent of* θ_0*, and it is well-defined up to an integer. In fact, if* $\hat{T}_1, \hat{T}_2$ *are two different lifts of* T*, then* $\rho(\hat{T}_1) - \rho(\hat{T}_2) = \hat{T}_1 - \hat{T}_2 \in \mathbb{Z}$.

Proof. See proof in [3], Proposition 11.1.1. □

Definition 1.15. We define the *rotation number* of T as $\rho(T) := \pi(\rho(\hat{T})) \in \mathbb{T}$. When there is no-space for confusion, we will write just ρ. Note that as we have taken modulo 1, $\rho(T)$ is well-defined.

In 1-dimension, any homemorphism of the circle that preserves orientation has an associated rotation number, although it does not necessarily have to be quasiperiodic.

Proposition 1.16. *Let* $T : \mathbb{T} \to \mathbb{T}$ *be an orientation-preserving homeomorphism. Then,* $\rho(T) \in \mathbb{Q}$ *iff* T *has a periodic point.*

Proof. See proof in [3], Proposition 11.1.4. □

Proposition 1.17. (The Rotation Number is invariant under homeomorphisms) *Let* $T, h : \mathbb{T} \to \mathbb{T}$ *be orientation-preserving homeomorphisms. Then,* $\rho(h^{-1} \circ T \circ h) = \rho(T)$.

Proof. See proof in [3], Proposition 11.1.3. □

This proposition extends the definition of rotation number from rotations to more intrinsic maps, as long as they are conjugated to a rotation. In this work, we use this result to justify that a general 1-dimensional quasiperiodic map $F : X \to X$ conjugated to an irrational rotation T_ρ, has an associated rotation number $\rho(F) := \rho(T_\rho) = \rho$.

Hence, we have obtained a formula to compute the rotation number for circle homeomorphisms $T : \mathbb{T} \to \mathbb{T}$: we compute the limit defined in Proposition 1.14 and then take it modulo 1.

Notice that the rotation number can also be computed by averaging as follows:

$$\begin{aligned}\rho(\hat{T}) &:= \lim_{N\to\infty} \frac{1}{N}(\hat{T}^N(\theta_0) - \theta_0) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} (\hat{T}^{n+1}(\theta_0) - \hat{T}^n(\theta_0)) \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} (\hat{T}(\hat{T}^n(\theta_0)) - \hat{T}^n(\theta_0)) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} G(\hat{T}^n(\theta_0))\end{aligned}$$

where $G : \mathbb{R} \to \mathbb{R}$ is the 1-periodic function defined in Lemma 1.12. Thus, given a trajectory (θ_n) of length N, the rotation number can be computed as $\rho = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} G(\hat{T}^n(\theta_0))$. In other words, ρ can be computed as a Birkhoff average, which is an average of a function over an orbit. Although this may seem like just another way to compute ρ, understanding this approach is crucial when the limit is computed numerically. There is a result in the field of Ergodic Theory, known as the Birkhoff Ergodic Theorem, concerning the convergence of these averages to a certain integral. In the following chapter, we will see what useful information we can derive from it.

Moreover, the 1-dimensional Conjugacy Problem also involves finding a diffeomorphism $h : \mathbb{T} \to X$, where X is a 1-dimensional differentiable manifold. It is important to note that h must be 1-periodic in $\theta \in \mathbb{T}$ (since $\theta = \theta + 1$ in $\mathbb{T}$, then $h(\theta) = h(\theta + 1)$). Let us consider its Fourier series representation:

$$h(\theta) = \sum_{k\in\mathbb{Z}} c_k e^{2\pi i k\theta}, \text{ where } c_k = \int_{\mathbb{T}} h(\theta) e^{-2\pi i k\theta} d\theta$$

As previously mentioned, the Birkhoff Ergodic Theorem guarantees convergence of the Birkhoff averages to a certain integral; thus we will approximate the Fourier coefficients c_k of the conjugacy h through Birkhoff averages. We will explain in detail how to do this in Chapter 4. For now, the idea is that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} h(T^n(\theta_0)) e^{-2\pi i k T^n(\theta_0)} \longrightarrow c_k$$

Chapter 2

Ergodic Theory in Broad Brushstrokes

In the preceding chapter, we mentioned that to demonstrate the quasiperiodic nature of a trajectory (x_n), we must find an irrational vector ρ and a conjugacy map h satisfying $x_n = h(n\rho \pmod 1)$. We have also suggested that, as a result of the Birkhoff Ergodic Theorem, it is possible to find a numerical approximation of these crucial objects by computing the limit of Birkhoff averages.

Within this chapter, we will state the Birkhoff Ergodic Theorem. This requires introducing some foundational mathematical concepts of Ergodic Theory. This field studies the statistical properties of deterministic dynamical systems by analysing the behaviour of time averages of various functions along system trajectories. Ergodic Theory is extensively discussed in [4]. While we won't go into detail, we will provide an overview of its key concepts. It is important to stress that some definitions and theorems may be modified to allow the map F to have different domain and codomain.

2.1 Measure-preserving Maps

Let us start by defining the basics notions of Measure Theory that we will need.

Definition 2.1. Let X be a set and $P(X)$ be its power set. A subset $\mathcal{B} \subseteq P(X)$ is a *σ-algebra over X* if satisfies: $X \in \mathcal{B}$, $\mathcal{B}$ is closed under complementation (if $B \in \mathcal{B}$, then $X \setminus B \in \mathcal{B}$), and $\mathcal{B}$ is closed under countable unions (if $B_1, \ldots, B_n \in \mathcal{B}$, then $B_1 \cup \ldots \cup B_n \in \mathcal{B}$).

Definition 2.2. Let X be a set and $\mathcal{B}$ be a σ-algebra over X. The tuple $(X, \mathcal{B})$ is known as a *measurable space*.

Definition 2.3. Let $(X, \mathcal{B})$ be a measurable space. A map $F : X \to X$ is *measurable* if for every $B \in \mathcal{B}$, $F^{-1}(B) := \{x \in X | F(x) \in B\} \in \mathcal{B}$.

Definition 2.4. A *measure* on a measurable space $(X, \mathcal{B})$ is a positive function $\mu : \mathcal{B} \to \mathbb{R}_{\geq 0} \cup \{+\infty\}$ with $\mu(\varnothing) = 0$ such that satisfies the *countable additivity*: for any countable collection $\{B_n\}_{n \in \mathbb{N}}$ of pairwise disjoint sets in $\mathcal{B}$, $\mu(\cup_{n \in \mathbb{N}} B_n) = \sum_{n \in \mathbb{N}} \mu(B_n)$. Moreover, we say that μ is a *probability measure* if $\mu(X) = 1$.

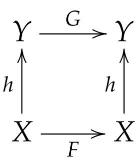
Definition 2.5. Let $(X, \mathcal{B})$ be a measurable space and μ be a measure on $(X, \mathcal{B})$. The triple $(X, \mathcal{B}, \mu)$ is called a *measure space*. Moreover, if μ is a probability measure, then $(X, \mathcal{B}, \mu)$ is known as a *probability measure space*

Now we are ready to define what is a measure-preserving map, which we will presume to have from Chapter 3 onward.

Definition 2.6. Let $(X, \mathcal{B}, \mu)$ be a measure space. A measurable map $F : X \to X$ is a *measure-preserving* (or *μ-invariant*) map if for all $B \in \mathcal{B}$, $\mu(B) = \mu(F^{-1}(B))$. We will refer to μ as the *invariant measure* for F.

Note that we define this concept in terms of the inverse to avoid discussing the injectivity of F.

Proposition 2.7. (Measure-preserving is invariant under homeomorphisms) *Let $(X, \mathcal{B}_X, \mu_X)$ and $(Y, \mathcal{B}_Y, \mu_Y)$ be measure spaces. Let $F : X \to X$ be measure-preserving with respect to μ_x and $h : X \to Y$ be a measurable homeomorphism. Assume $\mu_Y(B_Y) := \mu_X(h^{-1}(B_Y))$ for $B_Y \in \mathcal{B}_Y$. Then, $G := h \circ F \circ h^{-1}$ is measure-preserving with respect to μ_Y.*



Proof. As μ_X is F-invariant, $\mu_X(B_X) = \mu_X(F^{-1}(B_X))$ for all $B_X \in \mathcal{B}_X$. Consider $B_Y \in \mathcal{B}_Y$. Since h is measurable, it preserves the σ-algebras, i.e. $h^{-1}(B_Y) \in \mathcal{B}_X$. Then, $\mu_Y(B_Y) := \mu_X(h^{-1}(B_Y)) = \mu_X(F^{-1}(h^{-1}(B_Y)) = \mu_Y(h(F^{-1}(h^{-1}(B_Y))) = \mu_Y(G^{-1}(B_Y)))$. Hence, μ_Y is G-invariant. □

We will now state a theorem that proves, under certain additional conditions, the existence of such measures for dynamical systems in compact metric spaces.

Definition 2.8. Let X be a set. A *metric* on X is a function $d : X \times X \to \mathbb{R}$ such that for $x, y, z \in X$: $d(x, x) = 0$; if $x \neq y$, then $d(x, y) > 0$; $d(x, y) = d(y, x)$; and $d(x, z) \leq d(x, y) + d(y, z)$. The tuple (X, d) is known as a *metric space*.

Theorem 2.9. (Krylov-Bogolyubov) *Let X be a compact metric space. For any continuous map $F : X \to X$ there exists invariant (probability) Borel measures.*

Proof. See proof in [4], Theorem 3.5.1. □

There is a very particular context, where proving that a map F is measure-preserving with respect to the Lebesgue measure in $\mathbb{R}^d$ is very simple.

Lemma 2.10. *Let $(\mathbb{R}^d, \mathcal{B}, \lambda)$ be the measure space where $\mathcal{B}$ is the σ-algebra formed by the Borel measurable sets and λ stands for the Lebesgue measure. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a continuous function such that $\forall B \in \mathcal{B}$, $\int_B f dx = 0$. Then, $f(x) = 0$ for all $x \in \mathbb{R}^d$.*

Proof. Assume there exists $x_0 \in \mathbb{R}^d$ such that $f(x_0) \neq 0$. Without loss of generality, assume $f(x_0) > 0$. Then, as f is continuous, there exists $\epsilon > 0$ such that $f(x) > 0$ $\forall x \in B(x_0, \epsilon)$. Hence, $\int_{B(x_0,\epsilon)} f dx > 0$, which contradicts the hypothesis. □

Lemma 2.11. *Let $(\mathbb{R}^d, \mathcal{B}, \lambda)$ be the measure space where $\mathcal{B}$ is the σ-algebra formed by the Borel measurable sets and λ stands for the Lebesgue measure. Let $F : \mathbb{R}^d \to \mathbb{R}^d$ be a C^1-diffeomorphism. F is measure-preserving with respect to λ iff $|det(DF(x))| = 1$ for all $x \in \mathbb{R}^d$.*

Proof. Since F is a diffeomorphism, it is injective and we can rewrite the measure-preserving condition as $\lambda(B) = \lambda(F(B))$, for all $B \in \mathcal{B}$. This is equivalent to

$$\int_B dx = \int_{F(B)} du \iff \int_B dx = \int_B |det(DF(x))| dx$$

where we used the *change of variables* formula for multivariable integration.

If we assume that $|det(DF(x))| = 1$ for all $x \in \mathbb{R}^d$, then the equality of integrals holds, so F is measure-preserving. To show the other implication, we must demonstrate that $\forall B \in \mathcal{B}$, $\int_B (1 - |det(DF(x))|) dx = 0$. As $f(x) := 1 - |det(DF(x))|$ is continuous, the previous Lemma holds and $f(x) = 0$ for all $x \in \mathbb{R}^d$, which implies that $|det(DF(x))| = 1$ for all $x \in \mathbb{R}^d$. □

Example 2.12. Consider the *Standard map* defined as:

$$\begin{array}{rccc} F_\epsilon : & \mathbb{T} \times \mathbb{R} & \longrightarrow & \mathbb{T} \times \mathbb{R} \\ & (x, y) & \longmapsto & (x + y - \frac{\epsilon}{2\pi} sin(2\pi x), y - \frac{\epsilon}{2\pi} sin(2\pi x)) \end{array}$$

See Figure 5.1 for a plot of its orbits. The differential matrix of F is

$$\begin{pmatrix} 1 - \epsilon cos(2\pi x) & 1 \\ -\epsilon cos(2\pi x) & 1 \end{pmatrix} \tag{2.1}$$

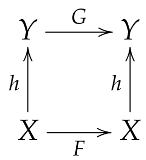
As $det(DF) = 1$, F is an area-preserving map with respect to the Lebesgue measure on $\mathbb{T} \times \mathbb{R}$, induced by the one on $\mathbb{R}^2$.

2.2 Ergodicity

Definition 2.13. Let $(X, \mathcal{B}, \mu)$ be a measure space. A measure-preserving dynamical system $F : X \to X$ is *ergodic* if for every F-invariant set $B \in \mathcal{B}$, i.e. $F^{-1}(B) = B$, it is the case that either $\mu(B) = 0$ or $\mu(X \setminus B) = 0$. We refer to μ as an *ergodic measure* for F.

In other words, the definition implies the absence of non-trivial invariant sets under system's dynamics.

Proposition 2.14. (Ergodicity is invariant under homeomorphisms) *Let $(X, \mathcal{B}_X, \mu_X)$ and $(Y, \mathcal{B}_Y, \mu_Y)$ be measure spaces. Let $F : X \to X$ be ergodic with respect to μ_x and $h : X \to Y$ be a measurable homeomorphism. Assume $\mu_Y(B_Y) := \mu_X(h^{-1}(B_Y))$ for $B_Y \in \mathcal{B}_Y$. Then, $G := h \circ F \circ h^{-1}$ is ergodic with respect to μ_Y.*



Proof. By Proposition 2.7, we know μ_Y is G-invariant. Notice that since h is measurable, it preserves the σ-algebras, i.e. $h^{-1}(B_Y) \in \mathcal{B}_X$. Let us start by characterising G-invariant sets:

$$G^{-1}(B_Y) = B_Y \iff (h \circ F \circ h^{-1})^{-1}(B_Y) = B_Y$$

$$\iff (h \circ F^{-1} \circ h^{-1})(B_Y) = B_Y \iff F^{-1}(h^{-1}(B_Y)) = h^{-1}(B_Y)$$

Hence, B_Y is G-invariant iff $h^{-1}(A)$ is F-invariant. We will now show that B_Y has either zero-measure or full-measure with respect to μ_Y. As F is ergodic, we know that any set $B_X \in \mathcal{B}_X$ such that $F^{-1}(B_X) = B_X$, satisfies either $\mu_X(B_X))) = 0$ or $\mu_X(X \setminus B_X) = 0$. Thus, if B_Y is G-invariant, then $h^{-1}(B_Y) \in \mathcal{B}_X$ is F-invariant and has either zero-measure or full-measure with respect to μ_X. Since h is a bijection, this implies that B_Y satisfies either $\mu_Y(B_Y) = 0$ or $\mu_Y(Y \setminus B_Y) = 0$. □

We will now present a lemma that provides a necessary and sufficient condition for determining if a dynamical system is ergodic. Its utility will become apparent in the following chapter.

Lemma 2.15. (Ergodicity via invariant square integrable functions) *Let $(X, \mathcal{B}, \mu)$ be a probability measure space and $F : X \to X$ be a measure-preserving map. F is ergodic iff for all $f \in L^2(X, \mu)$ such that $f \circ T = f$ μ-a.e., f is μ-a.e. constant.*

Proof. See proof in [4], Lemma 3.6.3. □

2.3 Birkhoff Ergodic Theorem

Once the previous sections are understood, we can state one of the main results on which our work is based: the Birkhoff Ergodic Theorem.

Definition 2.16. Let $(X, \mathcal{B}, \mu)$ be a measure space and $F : X \to X$ be a measure-preserving map. Let $n \in \mathbb{N}$ and $f : X \to \mathbb{R}^n$ a function. Given $x_0 \in X$ and $N \in \mathbb{N}$, consider the N-segment of the forward orbit through x_0. We define the *Birkhoff average* of the function f at the point x_0 as the time average of the form:

$$B_{N,F}(f)(x_0) := \frac{1}{N} \sum_{n=0}^{N-1} f(F^n(x_0))$$

Theorem 2.17. (Birkhoff Ergodic Theorem for ergodic transformations) *Consider the same notation as in the previous Definition 2.16. Let F be an ergodic map with respect to μ. For all $f \in L^1(X, \mu)$ and for μ-a.e. $x_0 \in X$, the Birkhoff time average converges to the space average, i.e*

$$\lim_{N\to\infty} B_{N,F}(f)(x_0) = \int_X f d\mu$$

Moreover, for any non-constant f, there exists a constant $K > 0$ such that for infinitely many N:

$$\Big|\frac{1}{N} \sum_{n=0}^{N-1} f(F^n(x_0)) - \int_X f d\mu\Big| \geq \frac{K}{N}$$

Proof. See proof of the first part in [4], Theorem 3.8.2. See proof of the second part in [5]. □

It is important to note that for numerical computations, a rate of convergence greater than $O(1/N)$ is quite slow. In the following chapter, we will see how the rate of convergence can be improved assuming F is quasiperiodic.

Observe that the fact that the convergence holds for μ-a.e. $x \in X$, implies that the average behaviour of the system can be deduced from the trajectory of a "typical" point. In the particular case that $f = \mathcal{X}_B$, where $\mathcal{X}_B$ stands for the *characteristic function of B*, the Birkhoff Ergodic Theorem implies that for all $B \subseteq X$ the proportion of time the system spends in B, as time goes to infinity, is the same regardless of the starting point, provided the initial condition is in a set of full-measure.

Chapter 3

Superconvergence of Ergodic Averages for Quasiperiodic Orbits

By the Birkhoff Ergodic Theorem 2.17, we know that the time average converges to the space average. But for general ergodic dynamical systems, the rate of convergence of these sums can be arbitrarily slow, as mentioned in the theorem. For many purposes the speed of convergence is irrelevant but it is important for numerical computations. We will show that we can improve the rate of convergence if the dynamical system is quasiperiodic.

The idea is to modify the Birkhoff average by weighting each term such that the early and late terms of the set $\{0, ..., N-1\}$ are weighted much less than the terms with $n \sim N/2$ in the middle. Assuming in addition that (x_n) is a quasiperiodic trajectory and $f \in C^\infty$, we will demonstrate that the weighted time average converges far faster to the space average.

3.1 Rotations are Measure-preserving Maps

Since $\mathbb{T}^d$ is a compact metric space and rotations are continuous, by Theorem 2.9, we know there exist a measure such that the rotations are measure-preserving maps. We will now demonstrate that the Lebesgue probability measure satisfies this requirement.

Definition 3.1. For any open interval $I := (a, b) \subseteq [0, 1]$, we consider its length as $l(I) := b - a$. Let $d \in \mathbb{N}$ and $C \subseteq [0, 1]^d$ be a rectangular cuboid, i.e. $C := I_1 \times \ldots \times I_d$ is a product of open intervals. We consider its volume as $vol(C) := \lambda_1(I_1) \cdot \ldots \cdot \lambda_1(I_d)$. The *Lebesgue probability measure on* $[0, 1]^d$ is $\lambda_d : \mathcal{A}_d \to [0, 1]$, where $\mathcal{A}_d$ is a borelian σ-algebra of $[0,1]^d$, defined so that for each $A \in \mathcal{A}$,

$$\lambda_d(A) := \inf\{\sum_{n=1}^{\infty} vol(C_n) | (C_n)_{n\in\mathbb{N}} \text{ is a sequence of cuboides and } A \subseteq \bigcup_{n=1}^{\infty} C_n\}$$

Definition 3.2. The *Lebesgue probability measure on* $\mathbb{T}^d$ is $\lambda_d : \mathcal{B}_d \to [0,1]$, where $\mathcal{B}_d$ is a borelian σ-algebra of $\mathbb{T}^d$, defined in the same way of the Lebesgue probability measure on $[0,1]^d$ under the identification $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d \simeq ([0,1]/0 \sim 1)^d$.

Lemma 3.3. *The Lebesgue probability measure λ_d is invariant for rotations.*

Proof. We will proof the result only for $d = 1$. A 1-dimensional rotation is a map $T_\rho : \mathbb{T} \to \mathbb{T}$ such that $T_\rho(\theta) = \theta + \rho \pmod 1$. Consider $(a,b) \subseteq \mathbb{T}$, its Lebesgue probability measure $\lambda_1((a,b))$ is given by its *positive difference*, i.e. $\lambda_1((a,b)) = \min\{|b-a|, |1+a-b|\}$. Notice that $T_\rho^{-1} = T_{-\rho}$ and thus the inverse is also a rotation. As the image of an interval under a rotation has the same length, $\lambda_1((a,b)) = \lambda_1(T_\rho^{-1}((a,b)))$. Therefore, T_ρ is λ_1-invariant. □

Now, we will state that there is only one possible choice of measure for which an irrational rotation is a measure-preserving map. According to the precedent Lemma 3.3, this measure must be the Lebesgue probability measure.

Proposition 3.4. *An irrational rotation has a unique measure for which it is a measure-preserving map, which is the Lebesgue probability measure.*

Proof. See proof in [3], Proposition 4.2.1. □

3.2 How Ergodicity applies to Irrational Rotations

In this section, we will see whether the concept of ergodicity defined in the previous chapter is applicable to rotations. In particular, we will proceed to show that an irrational rotation is ergodic with respect to the Lebesgue probability measure, while rational rotations are not, as illustrated by a counterexample.

Theorem 3.5. (Uniqueness of Fourier Coefficients) *Let $f \in L^2(\mathbb{T}^d, \lambda_d)$. Then, its Fourier coefficients $c_k = 0$ for all $k \in \mathbb{Z}^d$ iff $f(\theta) = 0$ for λ_d-a.e. $\theta \in \mathbb{T}^d$. In particular, taking differences, if two functions have the same Fourier coefficients, then they are the same except on a zero-measure set.*

Proposition 3.6. *An irrational rotation is ergodic with respect to λ_d.*

Proof. We will employ Lemma 2.15. Consider $f \in L^2(\mathbb{T}^d, \lambda_d)$ and assume $f(T_\rho(\theta)) = f(\theta)$ for λ_d-a.e. point $\theta \in \mathbb{T}^d$, we want to show that f is constant λ_d-a.e. Consider the Fourier series of f and $f \circ T_\rho$:

$$f(\theta) = \sum_{k\in\mathbb{Z}^d} c_k e^{2\pi i k\cdot\theta}, \text{ where } c_k = \int_{\mathbb{T}^d} f(\theta) e^{-2\pi i k\cdot\theta} d\theta$$

$$f \circ T_\rho(\theta) = \sum_{k\in\mathbb{Z}^d} \hat{c}_k e^{2\pi i k\cdot\theta}, \text{ where } \hat{c}_k = \int_{\mathbb{T}^d} f \circ T_\rho(\theta) e^{-2\pi i k\cdot\theta} d\theta$$

Using $T_\rho(\theta) := \theta + \rho \pmod 1$ and doing a change of variables,

$$\hat{c}_k = \int_{\mathbb{T}^d} f(\theta+\rho) e^{-2\pi i k\cdot\theta} d\theta = \int_{\mathbb{T}^d} f(\theta) e^{2\pi i k\cdot(\rho-\theta)} d\theta = e^{2\pi i k\cdot\rho} c_k$$

Thus,

$$f \circ T_\rho(\theta) = \sum_{k\in\mathbb{Z}^d} \left[c_k e^{2\pi i k\cdot\rho}\right] e^{2\pi i k\cdot\theta}$$

Since we have assumed that $f(T_\rho(\theta)) = f(\theta)$ for λ_d-a.e., by Theorem 3.5 we can equate both expressions of the Fourier series, which implies:

$$c_k = c_k e^{2\pi i k\cdot\rho} \iff c_k(1 - e^{2\pi i k\cdot\rho}) = 0 \text{ for all } k \in \mathbb{Z}^d$$

As ρ is an irrational vector, $\nexists k \in \mathbb{Z}^d \setminus \{0\}$ such that $k \cdot \rho \in \mathbb{Z}$, so $(1 - e^{2\pi i k\cdot\rho}) \neq 0$ and $c_k = 0$ for all $k \in \mathbb{Z}^d \setminus \{0\}$. Thus, all Fourier coefficients of f are zero except possibly c_0. Then, $f = c_0$ for λ_d-a.e. as we wanted to show. □

Note that this can be extended to show that a general quasiperiodic map is ergodic, as follows.

Corollary 3.7. *Let $F : X \to X$ be a quasiperiodic map on X_0, where $X_0 \subseteq X$ is a d-dimensional manifold. Assume it is conjugated to an irrational rotation T_ρ through a measurable diffeomorphism $h : \mathbb{T}^d \to X_0$. Then, F is ergodic with respect to the measure defined by $\mu(B) := \lambda_d(h^{-1}(B))$, for $B \subseteq X$.*

Proof. By definition, $F = h(T_\rho)$ where h is a diffeomorphism and T_ρ is an irrational rotation, which by the previous Proposition 3.6 is ergodic. Therefore by Proposition 2.14, F is also ergodic. □

Example 3.8. (Rational rotations are not ergodic with respect to λ) Let $\rho = p/q$ with $p, q \in \mathbb{Z}$ coprime. Let λ_1 be the Lebesgue probability measure of $\mathbb{T}$. Consider $T_\rho : \mathbb{T} \to \mathbb{T}$ such that $T_\rho(\theta) := \theta + \rho$, for $\theta \in \mathbb{T}$. Define for $q \in \mathbb{N}$,

$$A_q := \bigcup_{i=0}^{q-1} [\frac{i}{q}, \frac{i}{q} + \frac{1}{2q}] \subseteq \mathbb{T}$$

The set A_q is clearly invariant under T_ρ, since the rational rotation sends each interval into another one. Since A is a union of q intervals of equal length $1/2q$, $\lambda_1(A_q) = 1/2 \in (0,1)$. We have constructed an invariant set whose measure is neither 0 nor 1, thus T_ρ is not ergodic with respect to λ_1.

We will now proof a version of the Birkhoff Ergodic Theorem for circle homeomorphisms. Note that in this particular case, the result is self-contained and does not require Measure Theory. Moreover, we will see that we obtain a bound of order $O(1/N)$ for the rate of convergence of the Birkhoff averages to the space average. The proof is taken from [6].

Theorem 3.9. (Birkhoff Ergodic Theorem for Circle Homeomorphisms) *Let $\rho \in \mathbb{R} \setminus \mathbb{Q}$. For every continuous function $f : \mathbb{T} \to \mathbb{C}$ and $\theta_0 \in \mathbb{T}$, the temporal average converges to the spatial average, i.e.*

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T_\rho^n(\theta_0)) \longrightarrow \int_{\mathbb{T}} f(\theta)d\theta \text{ as } N \to \infty$$

Proof. Consider $\theta_0 = 0$ so that the temporal average is $1/N\sum_{n=0}^{N-1} f(n\rho)$. Let us define $E_N(f) := 1/N\sum_{n=0}^{N-1} f(n\rho) - \int_{\mathbb{T}} f(\theta)d\theta$. Our goal is to show that $E_N(f) \to 0$.

(1) First, notice that if $f(\theta) = 1$, then $E_N(f) = 1 - 1 = 0$. Secondly, if $f(\theta) = e^{2\pi ik\theta}$ for $k \neq 0$, then

$$E_N(f) = \frac{1}{N}\sum_{n=0}^{N-1} e^{2\pi ikn\rho} - \int_{\mathbb{T}} e^{2\pi ik\theta}d\theta = \frac{1}{N}\sum_{n=0}^{N-1} e^{2\pi ikn\rho} = \frac{1}{N}\frac{e^{2\pi ikN\rho}-1}{e^{2\pi ik\rho}-1}$$

Using the triangular inequality and $|e^{i\alpha}| = 1$ for all $\alpha \in \mathbb{R}$,

$$|E_N(f)| \le \frac{1}{N}\frac{2}{|e^{2\pi ik\rho}-1|} \longrightarrow 0 \text{ as } N \to \infty$$

Therefore, if $f(\theta) = \sum_{k=-d}^{d} a_k e^{2\pi ik\theta}$, by linearity $E_N(f) \to 0$ as $N \to \infty$.

(2) If $f, g : \mathbb{T} \to \mathbb{C}$ and $\|f - g\|_\infty \le \epsilon$ for a constant $\epsilon > 0$ as small as we want,

$$|E_N(f) - E_N(g)| \le \frac{1}{N}\sum_{n=0}^{N-1} |f(n\rho) - g(n\rho)| + \int_{\mathbb{T}} |f(\theta) - g(\theta)|d\theta \le \epsilon + \epsilon = 2\epsilon$$

We know that for any arbitrary continuous function $f : \mathbb{T} \to \mathbb{C}$, and any constant $\epsilon > 0$, there exists a trigonometric polynomial $p : \mathbb{T} \to \mathbb{C}$ such that $\|f - p\|_\infty \le \epsilon/3$. Because of (1), there exists N_0 such that for all $N \ge N_0$, $|E_N(p)| \le \epsilon/3$. By (2), for all N, $|E_N(f) - E_N(p)| \le 2\epsilon/3$. Putting this together, we have that for all $N > N_0$,

$$|E_N(f)| \le |E_N(f) - E_N(p)| + |E_N(p)| \le 2\epsilon/3 + \epsilon/3 = \epsilon$$

which proofs the theorem. □

3.3 Weighted Birkhoff Averages

In this section, we aim to define what are the weighted Birkhoff averages and state the main theorem of this work about the superconvergence of these averages. This is based on the research of S. Das, Y. Saiki, E. Sander and J. Yorke, which is explained in [2] and [7].

Definition 3.10. A C^m function $w : \mathbb{R} \to [0, \infty)$ is a *bump function* if its support $supp(w) := Cl\{x \in \mathbb{R} | w(x) \neq 0\} = [0,1]$, $\int_{\mathbb{R}} w(x)dx \neq 0$, and w and all of its derivatives up to order m vanish at 0 and 1.

Example 3.11. An example of a C^∞ bump function family is the *exponential weightings*:

$$w^{[p]}(t) = \begin{cases} exp(\frac{-1}{t^p(1-t)^p}) & \text{for } t \in (0,1) \\ 0 & \text{for } t \notin (0,1) \end{cases}$$

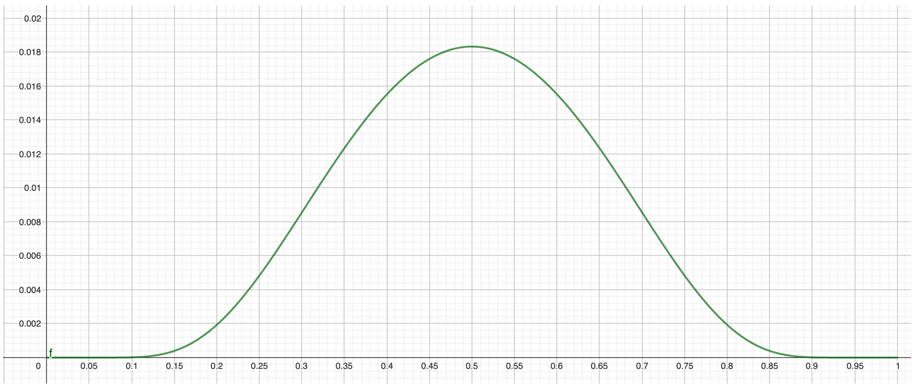


Figure 3.1: Representation of the bump function $w^{[1]}(t)$ for $t \in [0,1]$.

Definition 3.12. Let $(X, \mathcal{B}, \mu)$ be a probability measure space and $F : X \to X$ be a measure-preserving map. Let $n \in \mathbb{N}$ and $f : X \to \mathbb{R}^n$ a function. Let w be a bump function. Given $x_0 \in X$ and $N \in \mathbb{N}$, consider the N-segment of the forward orbit through x_0. We define the *weighted Birkhoff average* of the function f at the point x_0 as the time average of the form:

$$WB_{N,F}(f)(x_0) := \frac{1}{A_N} \sum_{n=0}^{N-1} w(\frac{n}{N}) f(F^n(x_0)), \text{ where } A_N := \sum_{n=0}^{N-1} w(\frac{n}{N})$$

Notice that it is an average of $f(F^n(\theta))$ as $\sum_{n=0}^{N-1} \frac{w(n/N)}{A_N} = 1$. Heuristically, this scheme weights more heavily the "typical" terms in the middle of the sequence, avoiding "boundary effects" due to the fact that we average only a finite orbit segment.

Definition 3.13. An irrational vector $\rho \in \mathbb{R}^d$ is said to be *Diophantine* if for some $\beta > 0$ it is *Diophantine of class* β, i.e. there exists $C_\beta > 0$ such that for every $k \in \mathbb{Z}^d \setminus \{0\}$, and every $n \in \mathbb{Z}$,

$$|k \cdot \rho - n| \geq \frac{C_\beta}{\|k\|^{d+\beta}}$$

In most of the cases, the norm is set to be $\|k\|_1 := k_1 + \ldots + k_d \in \mathbb{R}$. Diophantine vectors are a highly irrational type of vectors. The idea is that they can not be well-approximated by any rational vector. Nonetheless, they are abundant.

Proposition 3.14. (Diophantine vectors are dense in $\mathbb{R}^d$) *For all $\beta > 0$, the set of Diophantine vectors of class β have full Lebesgue measure in $\mathbb{R}^d$.*

Proof. See proof in [8], Proposition 5.4. □

Now we are ready to introduce the Weighted Birkhoff Ergodic theorem, which is the main result that this work is based on.

Theorem 3.15. (Weighted Birkhoff Ergodic Theorem) *Let X be a C^r-manifold and $F : X \to X$ be a C^r-map, which is quasiperiodic on a d-dimensional manifold $X_0 \subseteq X$. Assume F has a Diophantine class β rotation vector ρ. Let μ be the induced measure on X_0 by the C^r-conjugacy between F and an irrational rotation T_ρ. Let $n \in \mathbb{N}$ and let $f : X \to \mathbb{R}^n$ be a C^r-function. Then, for each $m \in \mathbb{N}$ such that $r > d + m(d + \beta)$, there exists a constant $K_m > 0$ independent of $x_0 \in X_0$ that satisfies*

$$|WB_{N,F}(f)(x_0) - \int_{X_0} f d\mu| \leq \frac{K_m}{N^m}, \text{for all } N \in \mathbb{N}$$

where we consider w to be a C^m bump function.

Proof. See following Section 3.4. □

In the proof, we will see that the constant K_m relies on $w(t)$ and its first m derivatives, the function f, and the Diophantine class β of the rotation vector. However, it remains independent of x_0.

Since the limit as $N \to \infty$ is the same for $WB_{N,F}(f)(x_0)$ and $B_{N,F}(f)(x_0)$, the weighted Birkhoff average provides a faster method to compute the space integral. In all 1D quasiperiodic studies shown in [7], convergence using weighted averages to a 30-digits accuracy is achieved well before $N = 10^6$ iterates. While if we had used B_N, we would need $N \sim 10^{30}$ according to Theorem 3.9. Hence, assuming a computation rate of 10^6 iterates per second, WB_N would require 1 second, while

B_N would demand over 1 billion billion years. Therefore, there is a pressing need for fast convergence.

As an observation, no single bump function consistently outperforms others. The optimal choice depends on the specific problem and the desired level of precision. This is illustrated through examples in [7].

Another pertinent observation is that if the ergodic process T were chaotic instead of quasiperiodic, the weighted Birkhoff averages would not provide any advantage over the Birkhoff averages.

It is worth noting that there is a particular case of the theorem that demonstrates superconvergence. In our examples, we will often have this scenario.

Definition 3.16. Let $(a_N)_{N=0}^{\infty}$ be a sequence in a normed vector space such that $a_N \to b$ as $N \to \infty$. We say (a_N) *superconverges* to b if for each $m \in \mathbb{N}$, there exists a constant $K_m > 0$ that satisfies

$$|a_N - b| \leq \frac{K_m}{N^m}, \text{ for all } N \in \mathbb{N}$$

Corollary 3.17. **(C^∞ version of the Weighted Birkhoff Ergodic Theorem)** *Let X be a C^∞-manifold and $F : X \to X$ be a C^∞-map, which is quasiperiodic on a d-dimensional manifold $X_0 \subseteq X$ and has invariant probability measure μ induced by the C^∞-conjugacy between F and an irrational rotation T_ρ. Assume F has a Diophantine rotation vector ρ. Let $n \in \mathbb{N}$ and let $f : X \to \mathbb{R}^n$ be a C^∞-function. Assume w is a C^∞ bump function. Then, for all $x_0 \in X_0$, the weighted Birkhoff average $WB_{N,F}(f)(x_0)$ has superconvergence to $\int_{X_0} f d\mu$. Moreover, the convergence is uniform in x_0.*

Proof. Consider Theorem 3.15 with $r = \infty$. □

3.4 Proof of the Superconvergence of Weighted Birkhoff Averages

To proof Theorem 3.15, we need to previously state two important results.

Lemma 3.18. **(Poisson Summation Formula)** *Let $g : \mathbb{R} \to \mathbb{C}$ be a Schwarz function, i.e. $\forall c > 0$, $n \in \mathbb{N}$ there exists a constant $K_{n,c} > 0$ such that $|g^{(n)}(x)| \leq \frac{K_{n,c}}{|x|^c}$. Let $k \in \mathbb{R}$ and let $\hat{g}(k) := \int_{\mathbb{R}} g(x)e^{-2\pi ikx}dx$ be the Fourier transform. Then,*

$$\sum_{n\in\mathbb{Z}} g(n) = \sum_{k\in\mathbb{Z}} \hat{g}(k)$$

Proof. Let us define $G(x) := \sum_{n\in\mathbb{Z}} g(x+n)$ for $x \in \mathbb{R}$. Notice that $G(x+1) = G(x)$, so G is 1-periodic. The Fourier series of a periodic function G is

$$G(x) = \sum_{k\in\mathbb{Z}} \hat{G}_k e^{2\pi i k x}, \text{ where its Fourier coefficients are } \hat{G}_k := \int_0^1 G(x)e^{-2\pi i k x}dx$$

Then,

$$\begin{aligned}\hat{G}_k &= \int_0^1 \sum_{n\in\mathbb{Z}} g(x+n)e^{-2\pi i k x}dx = \sum_{n\in\mathbb{Z}} \int_0^1 g(x+n)e^{-2\pi i k x}dx \\ &= \sum_{n\in\mathbb{Z}} \int_0^1 g(x+n)e^{-2\pi i k(x+n)}dx = \sum_{n\in\mathbb{Z}} \int_n^{n+1} g(y)e^{-2\pi i k y}dy \\ &= \int_{\mathbb{R}} g(y)e^{-2\pi i k y}dy =: \hat{g}(k)\end{aligned}$$

where we have used that $e^{-2\pi i k n} = 1$ and that g is a Schwarz function, so its Fourier series converges uniformly, to change the order of the integral and the summation.

Therefore, $G(x) = \sum_{k\in\mathbb{Z}} \hat{G}_k e^{2\pi i k x} = \sum_{k\in\mathbb{Z}} \hat{g}(k)e^{2\pi i k x}$. This implies $\sum_{n\in\mathbb{Z}} g(x+n) = \sum_{k\in\mathbb{Z}} \hat{g}(k)e^{2\pi i k x}$. Let us fix $x = 0$, then $\sum_{n\in\mathbb{Z}} g(n) = \sum_{k\in\mathbb{Z}} \hat{g}(k)$ as we wanted. □

Remark 3.19. This result also works for $g \in L^2$, though its proof relies on Distribution Theory.

Lemma 3.20. *$\sum_{k\in\mathbb{Z}^d\setminus\{0\}} \|k\|_1^{-\alpha}$ converges iff $\alpha > d$, where $\|\cdot\|_1 := |k_1| + |k_2| + \ldots + |k_d|$ is the l_1-norm.*

Proof. It suffices to prove the convergence for $\sum_{k\in\mathbb{Z}^d_{\geq 1}} \frac{1}{\|k\|_1^\alpha}$. Observe that,

$$\sum_{k\in\mathbb{Z}^d_{\geq 1}} \frac{1}{\|k\|_1^\alpha} = \sum_{k\in\mathbb{Z}^d_{\geq 1}} \frac{1}{(k_1+\ldots+k_d)^\alpha} = \sum_{S\geq d} \sum_{\{k\in\mathbb{Z}^d_{\geq 1} | k_1+\ldots+k_d = S\}} \frac{1}{S^\alpha}$$

Recall that the number of monomials of d variables and degree n is:

$$\sum_{\{k\in\mathbb{Z}^d_{\geq 0} | k_1+\ldots+k_d = n\}} 1 = \binom{d+n-1}{d-1}$$

and note that

$$\begin{aligned}\sum_{\{k\in\mathbb{Z}^d_{\geq 1} | k_1+\ldots+k_d = S\}} 1 &= \sum_{\{k'\in\mathbb{Z}^d_{\geq 0} | k'_1+\ldots+k'_d = S-d\}} 1 = \binom{S-1}{d-1} \\ &= \frac{(S-1)!}{(d-1)!(S-d)!} = \frac{(S-1)\ldots(S-(d-1))}{(d-1)!}\end{aligned}$$

Therefore,

$$\sum_{S\geq d}\sum_{\{k\in\mathbb{Z}^d_{\geq 1}|k_1+\ldots+k_d=S\}}\frac{1}{S^\alpha}=\sum_{S\geq d}\binom{S-1}{d-1}\frac{1}{S^\alpha}\sim\sum_{S\geq d}\frac{S^{d-1}}{(d-1)!}\frac{1}{S^\alpha}$$

where $\sim$ means that both series either diverge or converge. Hence, the series converges iff $\sum_{S\geq 0}\frac{1}{S^{\alpha+1-d}}$ converges, which happens iff $\alpha-d+1>1$, or equivalently $\alpha>d$. □

We will now give a proof of Theorem 3.15, which is based on the proof in [2].

Proof. As $F: X\to X$ is quasiperiodic on $X_0\subseteq X$, there exists a conjugacy $h:\mathbb{T}^d\to X$ to an irrational rotation T_ρ, such that $h:\mathbb{T}^d\to X_0$ is a C^r-diffeomorphism. Schematically, we have

$$\begin{array}{ccc} X_0\subseteq X & \xrightarrow{F|_{X_0}} & X_0\subseteq X \\ h\uparrow & & h\uparrow \\ \mathbb{T}^d & \xrightarrow[T_\rho]{} & \mathbb{T}^d \end{array}$$

Consider the Lebesgue measure on $\mathbb{T}^d$, denoted by λ_d, and the induced measure on X_0, $\mu(B)=\lambda_d(h^{-1}(B))$, where $B\subseteq X_0$ is a measurable set. Let us take $\theta_0:=h^{-1}(x_0)\in\mathbb{T}^d$ and $g:=f\circ h:\mathbb{T}^d\to\mathbb{R}^n$. Then, doing the change of variables $x=h(\theta)$,

$$\int_{X_0} f d\mu=\int_{\mathbb{T}^d} f\circ h d\lambda_d=\int_{\mathbb{T}^d} g d\lambda_d$$

$$WB_{N,F}(f)(x_0)=\frac{1}{A_N}\sum_{n=0}^{N-1}w\Big(\frac{n}{N}\Big)f(F^n(x_0))=\frac{1}{A_N}\sum_{n=0}^{N-1}w\Big(\frac{n}{N}\Big)g(T_\rho^n(\theta_0))$$

where $A_N:=\sum_{n=0}^{N-1}w(\frac{n}{N})$ and in the second line we used that $(f\circ F^n)(x_0)=(f\circ h\circ T_\rho^n\circ h^{-1})(x_0)=(g\circ T_\rho^n\circ h^{-1})(x_0)=(g\circ T_\rho^n)(\theta_0)$. Hence, it suffices to prove the result for irrational rotations $T_\rho:\mathbb{T}^d\to\mathbb{T}^d$ and functions $g:\mathbb{T}^d\to\mathbb{R}^n$.

Given $m\in\mathbb{N}$ such that $r>d+m(d+\beta)$, our goal is to show that:

$$\big|E_N\big|:=\Big|WB_{N,T_\rho}(g)(\theta_0)-\int_{\mathbb{T}^d} g(\theta)d\theta\Big|\leq\frac{K_m}{N^m}$$

for a certain constant $K_m>0$.

Consider the Fourier series representation of g, $g(\theta)=\sum_{k\in\mathbb{Z}^d}a_k\sigma_k(\theta)$ for $\theta\in$

$\mathbb{T}^d$, where $a_k := \int_{\mathbb{T}^d} g(\theta)\sigma_k(-\theta)d\theta$ and $\sigma_k(\theta) := e^{2\pi i k\cdot\theta}$. Then,

$$\begin{aligned} E_N &:= WB_{N,T_\rho}(g)(\theta_0) - \int_{\mathbb{T}^d} g(\theta)d\theta \\ &= WB_{N,T_\rho}\Big(\sum_{k\in\mathbb{Z}^d} a_k\sigma_k\Big)(\theta_0) - a_0 \\ &= \sum_{k\in\mathbb{Z}^d} a_k WB_{N,T_\rho}(\sigma_k)(\theta_0) - a_0 \\ &= \sum_{k\in\mathbb{Z}^d\setminus\{0\}} a_k WB_{N,T_\rho}(\sigma_k)(\theta_0) \end{aligned}$$

where we used that $a_0 := \int_{\mathbb{T}^d} g(\theta)\sigma_0(-\theta)d\theta = \int_{\mathbb{T}^d} g(\theta)d\theta$, and we took advantage of linearity and the finite summation to rearrange the order of the summations. Implicitly, we are also using that $g \in C^r$, so that its Fourier coefficients decrease. We will show this through the proof.

Now, using that w is defined for all $\mathbb{R}$, thought it is supported in $[0,1]$, $w(0) = w(1) = 0$, and the Poisson Summation Formula 3.18.

$$\begin{aligned} E_N &= \sum_{k\in\mathbb{Z}^d\setminus\{0\}} a_k WB_{N,T_\rho}(\sigma_k)(\theta_0) \\ &= \frac{1}{A_N}\sum_{k\in\mathbb{Z}^d\setminus\{0\}} a_k\Big[\sum_{n=0}^{N-1} w\Big(\frac{n}{N}\Big)\sigma_k(\theta_0+n\rho)\Big] \\ &= \frac{1}{A_N}\sum_{k\in\mathbb{Z}^d\setminus\{0\}} a_k\Big[\sum_{n\in\mathbb{Z}} w\Big(\frac{n}{N}\Big)\sigma_k(\theta_0+n\rho)\Big] \\ &= \frac{1}{A_N}\sum_{k\in\mathbb{Z}^d\setminus\{0\}} a_k\Big[\sum_{n\in\mathbb{Z}}\int_{\mathbb{R}} w\Big(\frac{t}{N}\Big)\sigma_k(\theta_0+t\rho)e^{-2\pi i n t}dt\Big] \\ &= \frac{1}{A_N}\sum_{k\in\mathbb{Z}^d\setminus\{0\}} a_k\Big[e^{2\pi i k\cdot\theta_0}\sum_{n\in\mathbb{Z}}\int_{\mathbb{R}} w\Big(\frac{t}{N}\Big)e^{2\pi i t(k\cdot\rho-n)}dt\Big] \end{aligned}$$

Consider a new variable $s := t/N$, so that $Nds = dt$. Let us start bound E_N:

$$|E_N| \leq \Big(\frac{N}{A_N}\Big)\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\Big|a_k\Big|\sum_{n\in\mathbb{Z}}\Big|\int_{\mathbb{R}} w(s)e^{2\pi i N s(k\cdot\rho-n)}ds\Big|$$

We will use $\|\cdot\|_1$ for vectors and $\|\cdot\|_\infty$ for functions, but to simplify the notation we won't specify it. Let us look at each part separately:

(a) A_N/N converges to $\int_0^1 w(s)ds \neq 0$, thus N/A_N is bounded by some constant $P_w > 0$.

(b) Consider $\Omega := 2\pi N(k \cdot \rho - n)$. Recall w and its first m derivatives vanish at 0 and 1 and that its support is $[0,1]$. Integrating by parts,

$$\int_0^1 w(s)e^{i\Omega s}ds = \frac{w(s)e^{i\Omega s}}{i\Omega}\Big]_0^1 - \int_0^1 \frac{w'(s)e^{i\Omega s}}{i\Omega}ds = \frac{-1}{i\Omega}\int_0^1 w'(s)e^{i\Omega s}ds$$
$$= \cdots = \frac{(-1)^m}{(i\Omega)^m}\int_0^1 w^{(m)}(s)e^{i\Omega s}ds$$

Thus, $\left|\int_0^1 w(s)e^{i\Omega s}ds\right| = \left|\Omega^{-m}\int_0^1 w^{(m)}(s)e^{i\Omega s}ds\right| \leq |\Omega|^{-m}\int_0^1 \left|w^{(m)}(s)e^{i\Omega s}\right|ds \leq |\Omega|^{-m}\|w^{(m)}\|$.

(c) We know $a_k := \int_{\mathbb{T}^d} f(\theta)\sigma_k(-\theta)d\theta$ and that $f \in C^r$. Using that $f(0) = f(1)$ and integrating by parts we have:

$$a_k = \int_0^1 \ldots \Big[\int_0^1 f(\theta)e^{-2\pi i k\cdot\theta}d\theta_1\Big]\ldots d\theta_d$$
$$= \int_0^1 \ldots \Big[\frac{1}{(2\pi i k_1)^{r_1}}\int_0^1 \frac{\partial^{r_1} f(\theta)}{\partial^{r_1}\theta_1}e^{-2\pi i k\cdot\theta}d\theta_1\Big]\ldots d\theta_d$$
$$= \ldots = \frac{1}{(2\pi i k_1)^{r_1}\ldots(2\pi i k_d)^{r_d}}\int_{\mathbb{T}^d}\frac{\partial^{r_1+\ldots+r_d} f(\theta)}{\partial^{r_1}\theta_1\ldots\partial^{r_d}\theta_d}e^{-2\pi i k\cdot\theta}d\theta$$

Therefore, there exist a constant P such that $|a_k| \leq P\|k\|^{-r}\int_{\mathbb{T}^d}\left|\frac{\partial^r f(\theta)}{\partial^r\theta}\right|d\theta \leq P_{f,r}\|k\|^{-r}$ for some constant $P_{f,r} > 0$.

Putting all together we have that,

$$|E_N| \leq \Big(\frac{N}{A_N}\Big)\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\Big|a_k\Big|\sum_{n\in\mathbb{Z}}\Big|\int_{\mathbb{R}} w(s)e^{2\pi i N s(k\cdot\rho-n)}ds\Big|$$
$$< P_w\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\frac{P_{f,r}}{\|k\|^r}\sum_{n\in\mathbb{Z}}\frac{\|w^{(m)}\|}{(2\pi N)^m|k\cdot\rho-n|^m}$$
$$= \frac{P_{w,f,r}}{N^m}\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\frac{1}{\|k\|^r}\sum_{n\in\mathbb{Z}}\frac{1}{|k\cdot\rho-n|^m}$$

where $P_{w,f,r} := P_w P_{f,r}\|w^{(m)}\|(2\pi)^{-m} > 0$ is a constant.

To finish the proof, it is sufficient to see that $\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\|k\|^{-r}\sum_{n\in\mathbb{Z}}|k\cdot\rho-n|^{-m}$ is finite. Consider $n_0 \in \mathbb{Z}$ such that it is the closest to $k\cdot\rho$, so that $k\cdot\rho - n_0 \in$

$(-0.5, 0.5)$. Then,

$$\begin{aligned}\sum_{n\in\mathbb{Z}} |k\cdot\rho - n|^{-m} &= \sum_{n\in\mathbb{Z}} |(k\cdot\rho - n_0) - (n-n_0)|^{-m} \\ &= |k\cdot\rho - n_0|^{-m} + \sum_{n\neq n_0, n\in\mathbb{Z}} |(k\cdot\rho - n_0) - (n-n_0)|^{-m} \\ &= |k\cdot\rho - n_0|^{-m} + \sum_{j\neq 0, j\in\mathbb{Z}} |(k\cdot\rho - n_0) - j|^{-m} \\ &\leq |k\cdot\rho - n_0|^{-m} + 2\sum_{j\in\mathbb{Z}_{>0}} (j-0.5)^{-m} \\ &\leq \frac{\|k\|^{m(d+\beta)}}{P_\beta^m} + P_m\end{aligned}$$

where P_β and P_m are positive constants. To bound the first summand, we used the Definition 3.13 of Diophantine vector of class β. As for the summation, we have used that $\sum_{j\in\mathbb{Z}_{>0}}(j-0.5)^{-m}$ converges as $m > 1$. To finish,

$$\sum_{k\in\mathbb{Z}^d\setminus\{0\}} \|k\|^{-r} \sum_{n\in\mathbb{Z}} \Big|k\cdot\rho - n\Big|^{-m} \leq \sum_{k\in\mathbb{Z}^d\setminus\{0\}} \|k\|^{-r}(P_\beta^{-m}\|k\|^{m(d+\beta)} + P_m)$$

which by Lemma 3.20 is finite iff $m(d+\beta) - r < -d$, or equivalently $r > d - m(d+\beta)$ as we wanted to show. □

Chapter 4

Applications

As an application of the weighted Birkhoff superconvergence method, we return to the main discussion of the paper: solving the Conjugacy Problem. In this chapter, we will explain how this numerical method applies to compute the rotation vector ρ and the Fourier coefficients of the parametrisation h, necessary to establish that a trajectory (x_n) presents quasiperiodic behaviour. Moreover, we will also explain how weighted averages can be used to compute the integral of a periodic function and the Lyapunov exponents of a dynamical system.

4.1 Rotation Vectors

Let us return to the discussion of Section 1.3, though now we will work with an arbitrary dimension d. Let X be a manifold and $F : X \to X$ be a quasiperiodic map on a d-dimensional invariant torus $X_0 \subseteq X$. By definition, F has an associated rotation vector ρ, which coincides with the rotation vector of the irrational rotation T_ρ to which it is conjugated.

However, with this definition the rotation vector of F is not unique. Assume F is conjugated via $h : \mathbb{T}^d \to X$ to a rotation T_ρ, i.e. $F(h(\theta)) = h(T_\rho(\theta))$ for $\theta \in \mathbb{T}^d$. Consider a matrix A with integer entries and $|det(A)| = 1$, so that the map $A : \mathbb{T}^d \to \mathbb{T}^d$, defined as $A(\theta) = A\theta$ for $\theta \in \mathbb{T}^d$, is invertible. Then,

$$F \circ h = h \circ T_\rho \Longleftrightarrow h^{-1} \circ F \circ h = T_\rho \Longleftrightarrow (h \circ A)^{-1} \circ F \circ (h \circ A) = A^{-1} \circ T_\rho \circ A$$

where

$$T_{A\rho} := A^{-1} \circ T_\rho \circ A : \begin{array}{ccc} \mathbb{T}^d & \longrightarrow & \mathbb{T}^d \\ \theta & \longmapsto & A^{-1}(A(\theta) + \rho) = \theta + A\rho \end{array}$$

Therefore, F is also conjugated via $h \circ A : \mathbb{T}^d \to X$ to a rotation $T_{A\rho}$, which has rotation vector $A\rho \neq \rho$. Hence, rotation vectors depend on the choice of the coordinate system and are well-defined except from continuous automorphisms on $\mathbb{T}^d$. Notice that the irrationality of a vector is preserved under this type of transformations.

We now explain how to obtain the rotation vector for F. Converting each finite data of iterates $(x_n = F^n(x_0)) \in X_0$ to an angle data $\phi_n := \phi(x_n) \in \mathbb{T}^d$, our goal is to determine ρ purely from $(\phi_n) \in \mathbb{T}^d$. It is important to stress that the trajectory (ϕ_n) does not have to follow an irrational rotation, it can just be a projection of the dynamics into $\mathbb{T}^d$. Assume the dynamics of (ϕ_n) are described by a map $\Phi : \mathbb{T}^d \to \mathbb{T}^d$.

The definition of lift and its subsequent lemma provided in Section 1.3 can be easily extended to an arbitrary dimension, hence we will take them from granted. As before, we will follow the convention of marking with a hat the lift functions.

Consider $G(\phi) := \hat{\Phi}(\phi) - \phi$ for $\phi \in \mathbb{T}^d$. For $\phi_0 := \phi(x_0) \in \mathbb{T}^d$, the standard approach towards computing the rotation vector ρ is as the limit

$$\begin{aligned}\rho(F) = \rho(\Phi) &:= \lim_{N\to\infty} \frac{1}{N}(\hat{\Phi}^N(\phi_0) - \phi_0) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} (\hat{\Phi}^{n+1}(\phi_0) - \hat{\Phi}^n(\phi_0)) \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} (\hat{\Phi}(\hat{\Phi}^n(\phi_0)) - \hat{\Phi}^n(\phi_0)) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} G(\hat{\Phi}^n(\phi_0)) \\ &= \lim_{N\to\infty} B_{N,\hat{\Phi}}(G)(\phi_0) = \lim_{N\to\infty} WB_{N,\hat{\Phi}}(G)(\phi_0)\end{aligned}$$

Thus, the rotation vector can be computed through weighted Birkhoff averages, which superconverge to $\rho(F)$ provided $\hat{\Phi}$ is C^∞, according to Theorem 2.17.

We will now illustrate a way to do this for $d = 1$, taken from [9]. For higher dimensions, one should take advantage of the *Tubular Neighbourhood Theorem* and construct tubular cylindrical coordinates. Although this is quite difficult and in practice it is more common to use Frequency Analysis. Consider the scenario where X_0 is a 1-dimensional quasiperiodic curve embedded in $X = \mathbb{R}^2$. Let $C := C_B \cup C_U \subset \mathbb{R}^2$ be the complement of X_0, where C_B (resp. C_U) is the bounded (resp. unbounded) component of C. Let $p \in C_B \subseteq \mathbb{R}^2$ and assume the curve is star-shaped with respect to the point $p = (a, b)$. Consider for $z = (x, y) \in X_0 \subset \mathbb{R}^2$,

$$\phi(z) = atan2(x - a, y - b) \in [-\pi, \pi]$$

where *atan*2 refers to the four quadrant arctangent function. Then, take

$$G(\phi) := \begin{cases} \frac{\phi(F(z)) - \phi(z)}{2\pi} & \text{if } \phi(F(z)) - \phi(z) \geq 0 \\ \frac{\phi(F(z)) - \phi(z)}{2\pi} + 1 & \text{otherwise} \end{cases}$$

The following Figure 4.1 tries to explain the nature of this trick.

For more in-depth on this topic, in [10] they discuss different ways of obtaining the rotation number and how to tune its approximation.

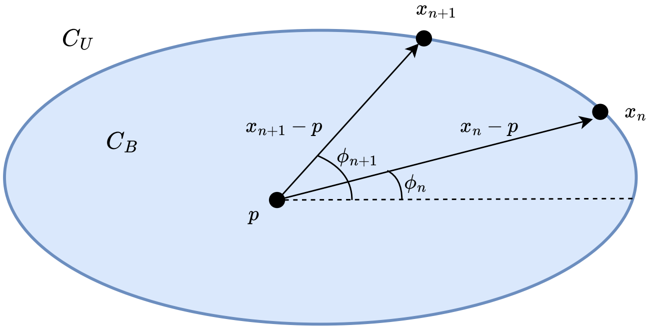


Figure 4.1: Correspondence between x_n and the angle data ϕ_n. It will be used to compute the rotation number of F.

4.2 Computing the Integral of a Periodic C^{∞}-function

Let $f : \mathbb{R}^d \to \mathbb{R}^n$ be a periodic map. In this section, we will see how we can accurately compute its integral with respect to the Lebesgue measure λ_d using weighted Birkhoff averages.

According to [10], it is as simple as following these steps:

(i) Rescale coordinates so that its domain is in $\mathbb{T}^d$.

(ii) Choose any $\rho = (\rho_1, \ldots, \rho_d) \in \mathbb{R}^d$ of Diophantine class $\beta \geq 0$. Recall that by Proposition 3.14, these vectors are dense in $\mathbb{R}^d$, so it is not so far-fetched to find one. For example, for $d = 1$ we could use $\rho = \varphi - 1 \in \mathbb{T}$, where $\varphi = (\sqrt{5} + 1)/2$ is the *golden ratio*.

(iii) Consider an irrational rotation T_ρ on $\mathbb{T}^d$ with rotation vector ρ and the exponential weighting $w^{[1]}$ defined in Example 3.11.

(iv) By Corollary 3.17, $WB_{N,T_\rho}(f)(\theta_0)$ superconverges to $\int_{\mathbb{T}^d} f d\lambda_d$ and the convergence is uniform in $\theta_0 \in \mathbb{T}^d$.

4.3 Fourier Series Coefficients of the Conjugacy

Once the rotation vector is determined, we want to find an approximation of the diffeomorphism $h : \mathbb{T}^d \to X_0$ such that $x_{n+1} = F(x_n)$ is conjugated to $\theta_{n+1} = T_\rho(\theta_n) := \theta_n + \rho \pmod 1$. The Fourier series representation of h is

$$h(\theta) = \sum_{k \in \mathbb{Z}^d} a_k e^{2\pi i k \cdot \theta} \text{ where } a_k := \int_{\mathbb{T}^d} h(\theta) e^{-2\pi i k \cdot \theta} d\theta$$

Recall that by ergodic theorems, it is enough to know a finite number of function's values along a trajectory to approximate its integral. While the map h is not known explicitly, some of its values $x_n = h(n\rho \pmod 1)$ are known. Hence, for every $k \in \mathbb{Z}^d$, we can estimate the *k-th Fourier coefficient of h* using weighted Birkhoff averages as follows:

$$WB_{N,T_\rho}(h(\theta)e^{-2\pi i k\cdot\theta}) = \frac{1}{A_N}\sum_{n=0}^{N-1} w\left(\frac{n}{N}\right) x_n e^{-2\pi i n k\cdot\rho} \longrightarrow a_k$$

which has superconvergence as $N \to \infty$. In the equality, we have used the fact that $x_n = h(\theta_n)$ and $\theta_n = n\rho \pmod 1$.

Therefore, weighted Birkhoff averages are also useful to compute the Fourier coefficients of the conjugacy h. Moreover, the continuity and differentiability of h can be non-rigorously estimated by observing the decay rate of its Fourier series coefficients a_k. For more in-depth on the smoothness of the conjugacy, see [10].

4.4 Lyapunov Exponents

In this section we will show a method that gives superconvergence to the Lyapunov exponents of a quasiperiodic system. We will consider its phase space to be $X = \mathbb{R}^n$, so we can talk about tangent vectors. This discussion is taken from [7].

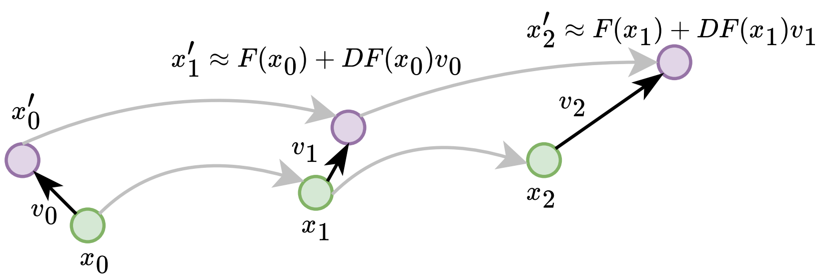


Figure 4.2: Idea of the discrete version of variational equations. We are interested in the average growth of the perturbation in the direction of v_0.

Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a C^∞ map. Consider the initial conditions $x_0, x_0' := x_0 + v_0 \in \mathbb{R}^n$, where $v_0 \in \mathbb{R}^n$ is a vector in the tangent space of x_0 and represents the initial perturbation, as illustrated in Figure 4.2. The following iterates will be $x_1 := F(x_0)$ and $x_1' := F(x_0') \approx F(x_0) + DF(x_0)v_0$, hence we can consider the perturbation of the first iterate to be $v_1 := DF(x_0)v_0$. Repeating this process, we end up with:

$$\begin{cases} x_0 \in \mathbb{R}^n \\ v_0 \in \mathbb{R}^n \\ x_n = F(x_{n-1}) = F^n(x_0) \\ v_n = DF(x_{n-1})v_{n-1} = DF^n(x_0)v_0 \end{cases}$$

where we used

$$\begin{aligned} DF^n(x_0) &= D(F \circ F^{n-1})(x_0) = DF(F^{n-1}(x_0)) \cdot DF^{n-1}(x_0) \\ &= DF(x_{n-1})DF^{n-1}(x_0) = \ldots = DF(x_{n-1})DF(x_{n-2})\ldots DF(x_0) \end{aligned}$$

so that,

$$\begin{aligned} v_n &:= DF(x_{n-1})v_{n-1} = DF(x_{n-1})DF(x_{n-2})v_{n-2} = \ldots = \\ &= DF(x_{n-1})DF(x_{n-2})\ldots DF(x_0)v_0 = DF^n(x_0)v_0 \end{aligned}$$

Now that the perturbation v_n of each iterate x'_n with respect to x_n is known, one may ask which is its average growth. This is usually measured through the Lyapunov multiplier.

Definition 4.1. Using the previous notation, we define the *Lyapunov multiplier in the direction of* v_0 as $\lim_{N\to\infty} \sqrt[N]{\|DF^N(x_0)v_0\|}$.

However, this growth rate is usually measured computing its logarithm.

Theorem 4.2. (Oseledec's Multiplicative Ergodic Theorem) *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a C^∞-map with invariant probability measure μ. There exist numbers $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n \in \mathbb{R}$ such that for μ-a.e. $x_0 \in \mathbb{R}^n$ and vector v_0 in the tangent space of x_0, the limit*

$$\lambda(v_0) := \lim_{N\to\infty} \frac{1}{N} \log \|DF^N(x_0)v_0\|$$

exists and equals one among $\lambda_1, \ldots, \lambda_n$. Moreover, the limit converges to λ_n for μ-a.e. $v_0 \in \mathbb{R}^n$ in the tangent space of x_0.

Proof. See proof in [11]. □

Definition 4.3. Using the notation of the previous Theorem 4.2, we refer to $\lambda_1, \ldots, \lambda_n$ as the *Lyapunov exponents* of F. Moreover, we say λ_n is the *maximal Lyapunov exponent*.

Lyapunov exponents measure the rate at which nearby trajectories diverge or converge and can be used to distinguish between chaos and quasiperiodicity. Consider the case where $F : \mathbb{R}^n \to \mathbb{R}^n$ is quasiperiodic on a d-dimensional subset $X_0 \cong \mathbb{R}^d \subset \mathbb{R}^n$. Then, on X_0 there are d Lyapunov exponents that are zero, which correspond to the d directions tangents to X_0. The existence of a positive Lyapunov exponent that corresponds to any of this tangent directions would be a signature of chaos.

The maximal Lyapunov exponent is often computed iteratively using a method analogous to the *power method* (which aims to find the dominant eigenvalue of a matrix). Given $x_0 \in \mathbb{R}^n$ and $v_0 \in \mathbb{R}^n$ in the tangent space of x_0, this version of the power method is defined recurrently as:

$$\begin{cases} w_0' := v_0 \\ l_0 := \|w_0'\| \\ w_0 := \frac{w_0'}{l_0} \end{cases} \qquad \begin{cases} x_n := F(x_{n-1}) \\ w_n' := DF(x_{n-1})w_{n-1} \\ l_n := \|w_n'\| \\ w_n := \frac{w_n'}{l_n} \end{cases}$$

One can check that $\|w_n\| = 1$ and $v_n = \Lambda_n w_n$, where $\Lambda_n := l_n l_{n-1} \dots l_0$. The idea is that, by iterating several times, we will be considering the most dominant direction. Hence, this will help us to find the maximal Lyapunov exponent. If we are considering $x_0 \in X_0$ such that it lies in a quasiperiodic orbit, the maximal Lyapunov exponent can be computed as an average over the trajectory (x_n):

$$\begin{aligned} \lambda_n : &= \lim_{N\to\infty} \frac{1}{N} \log \|DF^N(x_0)v_0\| = \lim_{N\to\infty} \frac{1}{N} \log \|v_n\| \\ &= \lim_{N\to\infty} \frac{1}{N} \log \Lambda_n = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \log l_n \end{aligned}$$

Hence, we can compute λ_n using weighted Birkhoff averages. Thus, we have obtained a method that superconverges to the maximal Lyapunov exponent.

Remark 4.4. For $n = 2$, the second Lyapunov exponent can be determined, although we will not provide a proof here. Once λ_2 is known, to obtain λ_1 we compute the sum $\lambda_1 + \lambda_2$, which can be expressed as a weighted Birkhoff average when x_0 is in a quasiperiodic trajectory:

$$\begin{aligned} \lambda_1 + \lambda_2 &= \lim_{N\to\infty} \frac{1}{N} \log |\det DF^N(x_0)| = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \log |\det DF(x_n)| \\ &= \lim_{N\to\infty} B_{N,F}(\log |\det DF(x_0)|) = \lim_{N\to\infty} WB_{N,F}(\log |\det DF(x_0)|) \end{aligned}$$

In addition, note that if F is a 2-dimensional map that satisfies the measure-preserving condition, i.e. $|det(DF)| = 1$, then the sum of the Lyapunov exponents is zero, $\lambda_1 + \lambda_2 = 0$. Therefore, on a quasiperiodc curve, implying that one Lyapunov exponent is zero, it follows that the other Lyapunov exponent must also be zero.

4.5 Machine Limitations

It is crucial to note that machines cannot perform computations with the exact Diophantine vector, as all the operands involved would be truncated so they are well-approximated by rationals. Consequently, an error arises from the truncation when performing numerical computations. However, as long as we have enough bits, we can always use an approximation of the Diophantine vector that achieves the desired precision.

Another problem we face when doing numerical approximations is how to choose the initial condition $x_0 \in \mathbb{R}^n$. There is no method to find a point such that it follows a quasiperiodic orbit. A simple way, which is surprisingly useful in practice, is to see if the plotted curves appear to densely fill a simple curve in $\mathbb{R}^n$. A more sophisticated approach is to consider an increasing finite sequence of orbits lengths $(N_i)_{i\in I}$, where $I \subset \mathbb{N}$ is finite, and compute the correspondent ρ_{N_i} for each $i \in I$. If ρ_{N_i} seems to converge as $i \in I$ grows, it is likely that x_0 is sampled from a quasiperiodic orbit. If instead it seems to oscillate, x_0 probably comes from a chaotic zone.

Remark 4.5. Ideally, when we do the numerical computations we can choose an increasing sequence of orbit lengths $(N_i)_{i\in I}$ such that the distance between x_{N_i} and x_0 is decreasing, so that we are considering iterates each time closer to x_0. This can be done thanks to the quasiperiodicity condition.

This in turn leads to an error, as we can only intuitively see that an initial condition gives way to a quasiperiodic trajectory. Moreover, we end up just knowing that the trajectory we are working with is near to a quasiperiodic orbit.

Chapter 5

Persistence of Quasiperiodic Motions under small Perturbations

The necessity of introducing Kolmogorov-Arnold-Moser (KAM) Theory arises from the fact that, in most of the cases, we will be working with a trajectory near to a quasiperiodic orbit. Thus, a natural question is to ask if quasiperiodicity persists for small perturbations. We won't go into detail on KAM Theory as it is a very complicated study and one should dedicate far more than an entire bachelor thesis to explain it properly. Hence, we will just quickly state the results that have an important role in our work.

5.1 Informal Explanation of the KAM Theorem

KAM Theory is generally studied for dynamical systems that are close to integrable systems, whose phase space is foliated by invariant d-dimensional tori. On each torus, the dynamical system acts as a rotation with an associated rotation vector, that varies throughout the family of invariant tori. KAM Theory shows that, under suitable regularity assumptions, most (but not all) of such tori are deformed and survive under small perturbations, i.e. there is a map from the original manifold to the deformed one that is continuous in the perturbation. Tori that survive have "sufficiently" irrational rotation vectors. This implies that the motion on the deformed tori continues to be quasiperiodic. The union of persistent n-dimensional tori, known as the *Kolmogorov set*, tend to fill the whole phase space as the strength of the perturbation is decreased.

When we defined Diophantine vectors, we mentioned that they are the "most irrational" type among irrational vectors. Thus, trajectories characterised by a Diophantine irrational vector, as well as trajectories sufficiently close to them, even without an irrational rotation vector, will present quasiperiodic behaviour. This justifies studying quasiperiodicity also in trajectories close to (pure) quasiperiodic ones.

Remark 5.1. In the previous chapter, we talked about Lyapunov exponents and how can they can quantify the sensibility of initial conditions. In particular, we mentioned that for d-dimensional quasiperiodic tori, the d Lyapunov exponents corresponding to the tangent vectors to the tori are zero.

In the following chapter, we will see how all the concepts introduced throughout the work can be numerically computed through the Arnold, Standard and Henon map. As previously mentioned, we won't go into the details of KAM Theory. However, we give informal statements that apply to these maps.

Informal Statement 1. (Arnold's Theorem) Let $f : \mathbb{R} \to \mathbb{R}$ be a 1-periodic C^∞-function. For $\alpha \in \mathbb{R}$, consider the circle diffeomorphism:

$$\begin{array}{rccl} A_{\alpha,\epsilon}: & \mathbb{T} & \longrightarrow & \mathbb{T} \\ & \theta & \longmapsto & \theta + \alpha + \epsilon f(\theta) \text{ (mod 1)} \end{array}$$

where $\epsilon \in \mathbb{R}$ is a parameter. Assume its rotation number ρ is Diophantine. Then, if ϵ is sufficiently small, $A_{\alpha,\epsilon}$ conjugates to an irrational rotation T_ρ with rotation number ρ. This result is also known as the *KAM Theorem for Circle Diffeomorphisms*.

The theorem can also be stated in the analytic category and in the finite differentiable case of f. The following version of the KAM Theorem applies for area-preserving maps on the cylinder.

Informal Statement 2. (Moser's twist Theorem) Let $F_0 : \mathbb{T} \times \mathbb{R} \to \mathbb{T} \times \mathbb{R}$ be an integrable map, so that $F_0(x, y) = (x + w(y) \text{ (mod 1)}, y)$, where $w : \mathbb{R} \to \mathbb{R}$ is a 1-periodic C^∞-function and satisfies the *twist condition* $\frac{dw}{dy} \neq 0$. Consider an area-preserving map of the form

$$\begin{array}{rccl} F_\epsilon: & \mathbb{T} \times \mathbb{R} & \longrightarrow & \mathbb{T} \times \mathbb{R} \\ & (x, y) & \longmapsto & (x + w(y) + \epsilon f(x, y) \text{ (mod 1)}, y + \epsilon g(x, y)) \end{array}$$

where $\epsilon \in \mathbb{R}$ is a parameter, and g and f are C^∞-functions. For sufficiently small ϵ, there exist many deformed tori that are close to the unperturbed invariant tori $\{y = \textit{constant}\}$. Moreover, the survived tori have Diophantine rotation number. This theorem is also known as the *KAM Theorem for area-preserving maps on the Cylinder*.

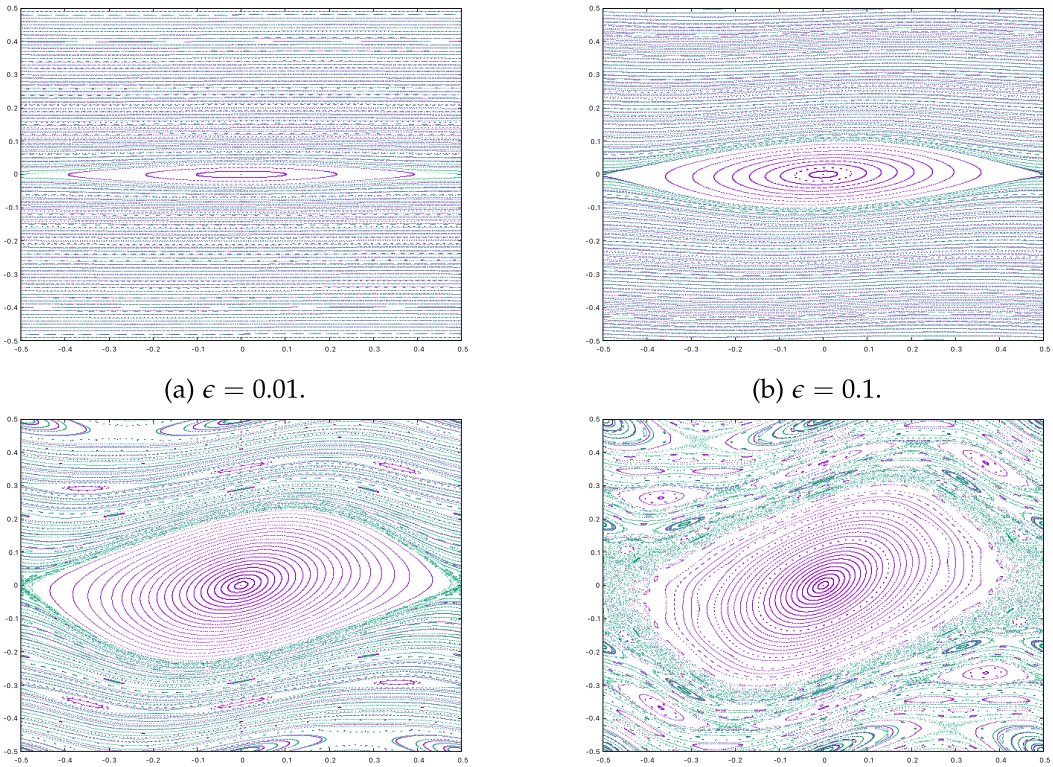
In a neighbourhood of an elliptic fixed point, changing to angle coordinates, the previous theorem also applies.

Informal Statement 3. (Stability of elliptic fixed points) Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be an area-preserving map. Let $(x^*, y^*) \in \mathbb{R}^2$ be an elliptic fixed point such that its eigenvalues $\lambda, 1/\lambda$ satisfy $\lambda^3 \neq 1$ and $\lambda^4 \neq 1$. Assume, in addition, that a condition similar to the twist condition that depends on the derivatives of F up to order 3 is satisfied. Then, the elliptic fixed point is surrounded by quasiperiodic invariant curves.

Example 5.2. This is very clear with the *Standard Map* F_ϵ, used also in Example 2.12, and defined as:

$$\begin{array}{rccl} F_\epsilon : & \mathbb{T} \times \mathbb{R} & \longrightarrow & \mathbb{T} \times \mathbb{R} \\ & (x, y) & \longmapsto & (x + y - \frac{\epsilon}{2\pi} sin(2\pi x), y - \frac{\epsilon}{2\pi} sin(2\pi x)) \end{array}$$

When the perturbation is zero, the phase space is foliated with horizontal invariant tori. If we allow ϵ to be small though non-zero, the nonlinear part begins to play a role, and many of these tori are slightly deformed, as depicted in Figure 5.1. As we keep increasing ϵ, the perturbation grows stronger and most tori are destroyed, leading to regions of chaotic behaviour bounded by the surviving KAM tori. Notice that we have a stable elliptic fixed point at $(0, 0)$ and an unstable hyperbolic fixed point at $(0.5, 0)$. In the vicinity of the first, the phase space exhibits concentric invariant tori, while in the second, the intersection between stable and unstable manifolds, leads to chaos. Hence, in a neighbourhood of the elliptic point we can apply the result mentioned above. The invariant tori that are homotopic to the curve $\{y = 0\}$ are known as *primary tori*, while the ones that are contractible and arise from the perturbation are known as *secondary tori*.



(a) $\epsilon = 0.01$.

(b) $\epsilon = 0.1$.

(c) $\epsilon = 0.5$.

(d) $\epsilon = 0.9$.

Figure 5.1: Orbits of the Standard Map for different values of ϵ. The initial conditions are (x_0, y_0), where $x_0 \in \{0, 0.5\}$ and $y_0 = 0.01k \in [-0.5, 0.5]$ for $k \in \mathbb{Z}$. We have set the orbit length at $N = 500$. This drawing identifies $\mathbb{T} = [-0.5, 0.5]/-0.5 \sim 0.5$ so that the elliptic fixed point $(0,0)$ is in the middle.

Chapter 6

Numerical Examples

In this chapter, we use the superconvergence method of weighted Birkhoff averages to numerically explore the concepts explained in the precedent chapters through worked examples. You can find all the coding used to obtain the graphics of this chapter in Appendix A.

For the *Arnold Circle map*, we compute approximations of its rotation number and Lyapunov exponent. In fact, we plot the *devil's staircase* and the Lyapunov exponent for different values of the perturbation ϵ to understand the complicity between them. Using the *Standard map*, we show the improvement on the speed of convergence of weighted Birkhoff averages with respect to the (unweighted) Birkhoff averages. Finally, we provide a numerical recipe taken from [9] that explains in detail each of the steps needed to follow to numerically compute the rotation number ρ and the conjugacy h for a quasiperiodic curve. The procedure is illustrated through the *Henon map*.

6.1 Application to the Arnold Circle Map

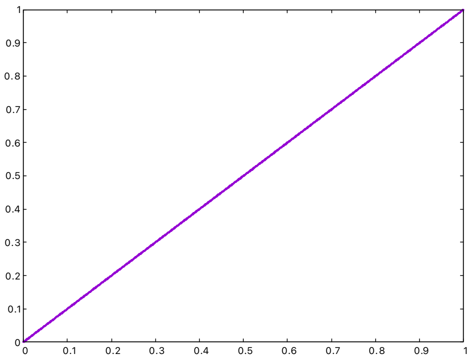
Let us consider again the *Arnold Circle map*, used in Example 1.13, defined as:

$$\begin{array}{rccl} A_{\alpha,\epsilon}: & \mathbb{T} & \longrightarrow & \mathbb{T} \\ & \theta & \longmapsto & \theta + \alpha - \frac{\epsilon}{2\pi} sin(2\pi\theta) \text{ (mod 1)} \end{array}$$

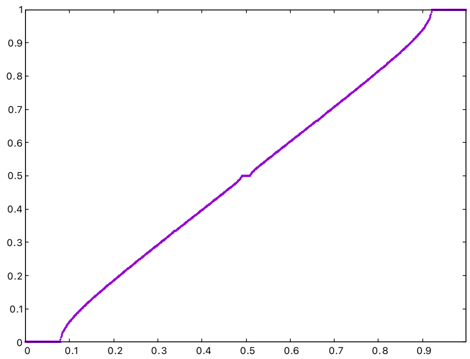
where α is a constant and ϵ is a small non-negative constant that represents the nonlinear perturbation.

We know that KAM Theory is also studied for Circle Diffeomorphisms and one can see the consequences of Arnold's Theorem, stated in the previous chapter, by prefixing different values of ϵ and plotting the *devil's staircase*, i.e. the rotation number against the value of α, as depicted in Figure 6.1. As expected, when $\epsilon = 0$,

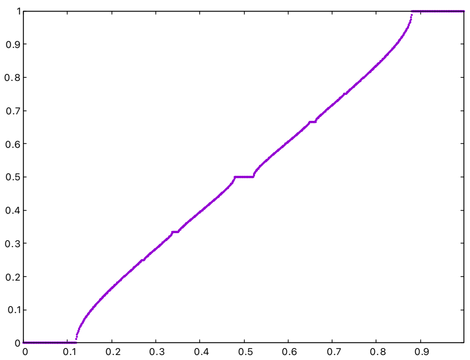
$A_{\alpha,0}$ is integrable and its dynamics are a (pure) rotation by $\rho = \alpha$. As we increase ϵ, stair-step patterns begin to emerge. While there exists a positive measure set of values for α for which the dynamics of $A_{\alpha,\epsilon}$ are quasiperiodic, periodic orbits start to appear. Subsets of the parameter space $\{\alpha, \epsilon\}$ where the system locks in a periodic orbit are known as *Arnold tongues*. As ϵ increases, these tongues become wider, indicating a larger region of parameter space where periodic orbits exist. At each platform, the rotation number is rational, i.e. $\rho = \frac{p}{q} \in \mathbb{Q}$ with $p, q \in \mathbb{N}$ coprime, and there is a couple of periodic orbits of period q, one attracting and the other repelling, that are born and die at the extrema of the platform, where the nonlinear perturbation is most pronounced, through saddle-node bifurcations.



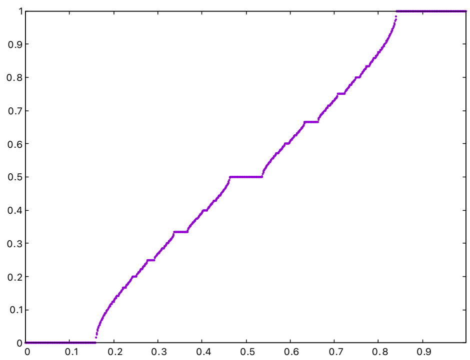
(a) $\epsilon = 0$.



(b) $\epsilon = 0.5$.



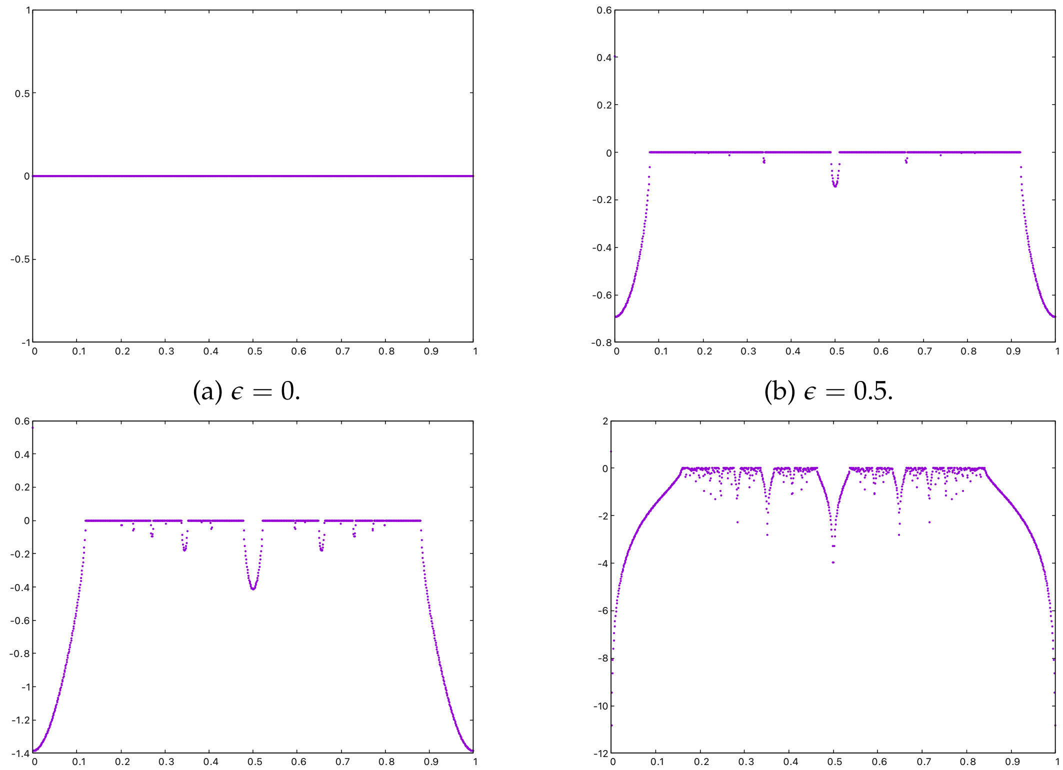
(c) $\epsilon = 0.75$.



(d) $\epsilon = 1$.

Figure 6.1: Devil's staircase for different values of ϵ against α. We have set the initial condition at $x_0 = 0.5$ and the orbit length at $N = 10,000$.

One can also compute the corresponding Lyapunov exponent of the Arnold map. Note the significance relationship between the plots of the devil's staircase and those of the Lyapunov exponents in Figure 6.2. As the steps of the staircase become larger, the corresponding Lyapunov exponent decreases. Moreover, we can see that when the dynamics are quasiperiodic the associated Lyapunov exponents is zero. While when the dynamics are periodic, the Lyapunov exponent is negative. This is due to the fact that the orbits are being attracted to the attracting periodic orbit.



(a) $\epsilon = 0$.

(b) $\epsilon = 0.5$.

(c) $\epsilon = 0.75$.

(d) $\epsilon = 1$.

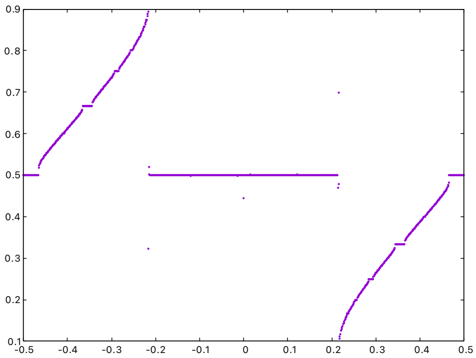
Figure 6.2: Lyapunov exponents for different values of ϵ against α. We have set the initial condition at $x_0 = 0.5$ and the orbit length at $N = 10,000$.

6.2 Application to the Standard Map

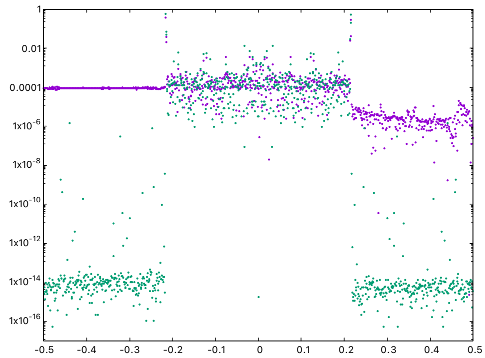
The *Standard map* F_ϵ has already come out several times during the work, see Example 2.12 for its definition. In this section, we will use it to compare the improvement of the weighted Birkhoff averages with respect to the unweighted ones. To do so, we will plot the difference between the rotation number of primary tori $|\rho_{N_2} - \rho_{N_1}|$ for different $N_1, N_2 \in \mathbb{N}$, using both methods, against the second coordinate of the initial condition $y_0 \in \mathbb{R}$.

It is interesting to compare the following Figure 6.3 with the orbits of the *Standard map*, depicted in Figure 5.1. Notice that, for the invariant tori that have survived, weighted Birkhoff averages converge significantly faster than the normal ones. Conversely, in the zone where these tori have been destroyed and we do not longer have quasiperiodic motions on the first coordinate, weighted Birkhoff averages provide no advantage.

As suggested by Theorem 3.9 and illustrated in Figure 6.3, with $N_1 = 10,000$ and $N_2 = 100,000$ the difference $|\rho_{N_2} - \rho_{N_1}|$ computed using (unweighted) Birkhoff averages is about 10^{-4}. In contrast, when using weighted Birkhoff averages for quasiperiodic trajectories, the difference is significantly smaller. In this case, it is on the order of 10^{-14}.



(a) $\epsilon = 0$.



(b) $\epsilon = 0.5$.

Figure 6.3: We have set $\epsilon = 0.5$ and $x_0 = 0$. In (a), you can see the rotation number of primary tori against the second coordinate of the initial condition $y_0 \in [-0.5, 0.5]$, computed with $N = 100,000$ and using weighted Birkhoff averages. Plot (b) depicts the difference between $|\rho_{N_2} - \rho_{N_1}|$, with $N_1 = 10,000$ and $N_2 = 100,000$, against the second coordinate of the initial condition $y_0 \in [-0.5, 0.5]$. We have used the logarithm scale in the vertical axis. In purple you can see the results obtained using (unweighted) Birkhoff averages, while in green the ones obtained using weighted Birkhoff averages.

6.3 Application to the Henon Map

In this section, we give a numerical recipe taken from [9] such that, for any quasiperiodic trajectory (x_n) in an invariant curve, explains in detail each of the steps needed to follow to numerically compute its rotation number ρ and the conjugacy h. The following steps constitute the main steps of our algorithm.

Step 1: Choose $x_0 \in X$ and $N \in \mathbb{N}$. Compute the orbit segment $O_N = \{x_n = F^n(x_0)\}_{n=0}^N$ and see if it appears to be sampled from a quasiperiodic zone, following the instructions on Section 4.5. If it seems to come from a chaotic zone, repeat this step choosing another x_0.

Step 2: Compute the rotation number ρ using weighted Birkhoff averages, as explained in Section 4.1. It is crucial to compute it as accurate as possible, as ρ will be used in subsequent numerical computations and even a small error in ρ can lead to significant noise in these computations. To improve the precision, increase the value of N.

Step 3: Compute the Fourier coefficients of h using weighted Birkhoff averages, as described in Section 4.3, though with a shorter sample size. Use N much smaller than in the rotation number computation and consider only some of these values, such as $k = 10i$ for $i \in \mathbb{N}$. This approach allows us to numerically determine $N_0 \in \mathbb{N}$ such that the norm of the Fourier coefficients $\|(a_k, b_k)\| < \epsilon_{machine}$ for all $|k| > N_0$.

Step 4: Consider the truncation index of the Fourier coefficients as $N_F \in \mathbb{N}$ roughly the 10% or 20% of N_0. Then, for all $|k| < N_F$ compute (a_k, b_k) through weighted Birkhoff averages with a moderate accuracy, i.e. take N larger than in Step 3 but smaller than in Step 2. Let us denote the h_0 the N_F Fourier polynomial of h, the *numerical defect* is

$$\epsilon_0 = \sup_{\theta \in \mathbb{T}} \Big| F(h_0(\theta)) - h_0(\theta + \rho) \Big|$$

If ϵ_0 is smaller than some tolerance, the initial guess is considered good. Otherwise, we repeat this step increasing N_F.

In Chapter 7, we explain a way to obtain an even better approximation of h by running the Newton iteration scheme once the approximation h_0 is known.

We will show how to apply the previous numerical recipe with a worked example. Consider the *Henon map* defined as:

$$\begin{array}{rccc} H_\alpha : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & \begin{pmatrix} x \\ y \end{pmatrix} & \longmapsto & \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y - x^2 \end{pmatrix} \end{array}$$

where the matrix represents a rotation by α. It is straightforward to verify that the determinant of the Jacobian matrix of H_α

$$DH_\alpha \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -2x & 1 \end{pmatrix}$$

is one for all $(x, y) \in \mathbb{R}^2$, indicating that the system is area-preserving with respect to the Lebesgue measure in $\mathbb{R}^2$.

Note that the origin is an elliptic fixed point, so in the absence of perturbations, nearby trajectories rotate around it with quasiperiodic motion. However, the perturbation due to the quadratic nonlinearity $y - x^2$, deforms or destroys some of these curves. In a small neighbourhood of the origin $(0, 0)$, where the perturbation is small, KAM Theory guarantees that those with a rotation number $\alpha/2\pi$ that is Diophantine survive.

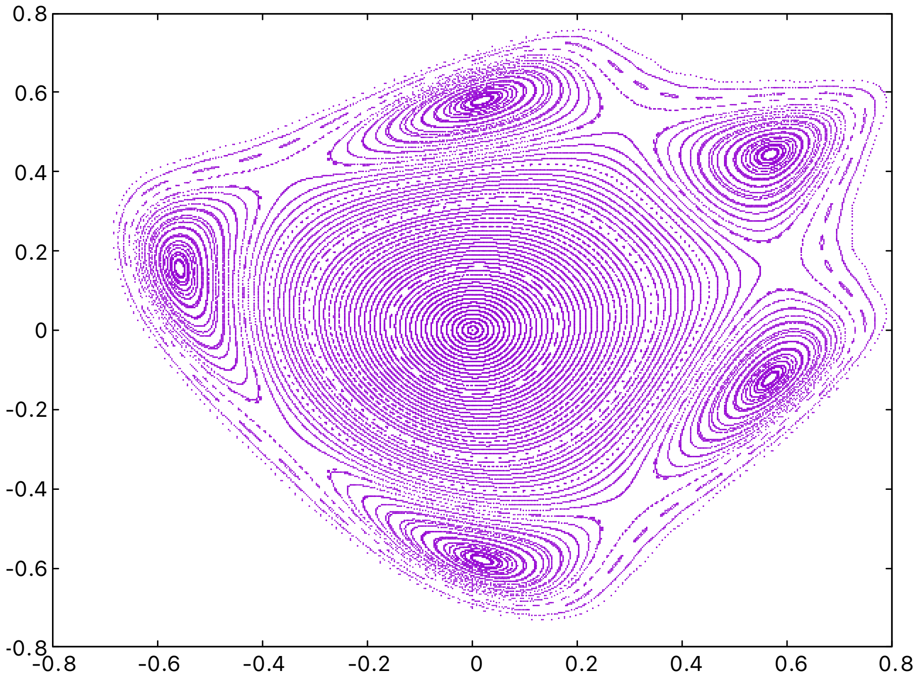


Figure 6.4: Orbits of the Henon map with $N = 1,000$ iterates, $\alpha = \arccos(0.24)$ and initial conditions $(0, y_0)$, where $y_0 \in [-0.7, 0.7]$ varies with an increment of 0.01. Note that near the origin, the dynamics are close to a rotation and we see a large set of invariant torus. As we move apart, they get more distorted and eventually there appears a family of 5-period systems. Further from the origin, the dynamics appear chaotic.

Throughout this example, we will assume $\alpha = \arccos(0.24)$.

Step 1. Consider the initial condition $x_0 = (0.4, 0) \in \mathbb{R}^2$. Let us compute the approximation of the rotation number for different values of N. See Figure 6.5. The computations suggest that it comes from a quasiperiodic trajectory, as the numerical approximations of the rotation number seem to converge. Then, we can move on to the next step.

Remark 6.1. If instead we had chosen $x_0' = (0.3, -0.44) \in \mathbb{R}^2$, we would not have seen any sign of convergence in the numerical computations of the rotation number, hence we would deduce it comes from a chaotic zone.

N	$\rho(x_1)$	$\rho(x_2)$
100	0.206164038365342	0.196863099485937
500	0.206174513248940	0.197503558666674
1000	0.206174514865070	0.197628415757003
5000	0.206174514865715	0.199431995293399
10000	0.206174514865712	0.199737097322017
50000	0.206174514865718	0.199823145343572
100000	0.206174514865710	0.199984739391916
150000	0.206174514865705	0.199998753698169
200000	0.206174514865702	0.199999888701773

Figure 6.5: Numerically computed values of the rotation number for different values of N and initial conditions $x_0 = (0.4, 0)$ and $x_0' = (0.3, -0.44) \in \mathbb{R}^2$.

Step 2. Based on the previous step, we deduce that the rotation number associated to the orbit of x_0 is $\rho \approx 0.206174514865704$, which is likely correct except possibly the last decimal digit.

Step 3. Note that for $N = 5,000$ we already have twelve correct digits for the rotation number, therefore we will compute Fourier coefficients with this orbit length. Our aim is to find the index $N_0 \in \mathbb{N}$ such that for all $|k| \geq N_0$, the k-th Fourier coefficient has norm smaller than a prefixed tolerance. To get the result faster, we will only consider $k = 10i$ for some $i \in \mathbb{N}$. Fixing the tolerance at 10^{-10}, we get that the norm of the 40^{th}-Fourier coefficient is smaller than the tolerance. Then, take $N_0 = 40$.

Step 4. Considering $N_F = 0.15 \cdot N_0 = 6$, a good approximation of h is

$$h_0(\theta) := \sum_{k=-N_F}^{N_F} \begin{pmatrix} a_k \\ b_k \end{pmatrix} e^{2\pi i k \theta}$$

where we compute the Fourier coefficients with better accuracy than in the previous step, i.e. with N larger than $5,000$. In our computations, we use $N = 100,000$. To measure how good is the approximation one can compute its initial defect:

$$\epsilon_0 = \sup_{\theta \in \mathbb{T}} \Big| F(h_0(\theta)) - h_0(\theta + \rho) \Big|$$

which, thanks to our codes that you can find in the appendix, we can say that it is around $9.58 \cdot 10^{-4}$. If we want a more precise approximation h_0, we can repeat this using a greater N_F. For instance, taking $N_F = N_0 = 40$, we get that the error is around $5.258 \cdot 10^{-12}$. Therefore, with this procedure we can get very good approximations of the parametrisation h.

There is another way, besides repeating Step 4 with a greater N_F, to tune the approximation h_0 of h. This is explained in the following Chapter 7.

Remark 6.2. Figure 6.4 also shows a group of five disjoint invariant circles, which have the property that each iterate jumps from one torus to the next one. Taking F^5, this reasoning applies. For a more detailed explanation of how this method applies for general K-periodic systems of quasiperiodic invariant sets, refer to [9].

Chapter 7

Delving Deeper into the Conjugacy

Consider a manifold X and map $F : X \to X$ that is quasiperiodic on a d-dimensional invariant set $X_0 \subseteq X$ with a rotation vector ρ. By definition, there exists a diffeomorphism $h : \mathbb{T}^d \to X_0$ that conjugates F to an irrational rotation $T_\rho : \mathbb{T}^d \to \mathbb{T}^d$, i.e. $F(h(\theta)) = h(T_\rho(\theta))$ for all $\theta \in \mathbb{T}^d$.

$$\begin{array}{ccc} X_0 \subseteq X & \xrightarrow{F|_{X_0}} & X_0 \subseteq X \\ h \uparrow & & h \uparrow \\ \mathbb{T}^d & \xrightarrow[T_\rho]{} & \mathbb{T}^d \end{array}$$

In previous chapters, we explained how to obtain a good approximation h_0 of h using weighted Birkhoff averages. Now, we will show a way to refine this approximation h_0. We will focus on the case where $X = \mathbb{R}^2$, and thus X_0 is a curve on the plane and $T_\rho : \mathbb{T} \to \mathbb{T}$ is defined as $T(\theta) := \theta + \rho \pmod 1$ for $\theta \in \mathbb{T}$. We will assume F is an area-preserving map, where invariant tori with irrational dynamics are natural, as shows the KAM Theorem. Exploiting the fact that h must satisfy the conjugacy equation $F(h(\theta)) = h(T_\rho(\theta))$ for all $\theta \in \mathbb{T}$, the idea is to perform the well-known Newton's iterative method to find a better approximation of h.

7.1 Parametrisation Method for an Invariant Tori

We will assume that the rotation number ρ (or a good approximation of it) is known. We are looking for h such that it satisfies the following equation

$$F(h(\theta)) = h(\theta + \rho) \text{ for all } \theta \in \mathbb{T}$$

To ensure the uniqueness of the solution, we need to fix a phase condition. We will assume $h(0)$ lies in a particular line in the plane by fixing two vectors $p, \eta \in \mathbb{R}^2$ and ask that the inner product $(p - h(0)) \cdot \eta = 0$.

Hence, we want to solve the resulting system of equations:

$$\begin{cases} F(h(\theta)) = h(\theta + \rho) \\ (p - h(0)) \cdot \eta = 0 \end{cases}$$

These are two equations for only one unknown h. To balance the system, we introduce a scalar unfolding parameter β. Thus, we now seek to solve:

$$\begin{cases} F(h(\theta)) = (1 + \beta)h(\theta + \rho) \\ (p - h(0)) \cdot \eta = 0 \end{cases}$$

These are two equations in two unknowns, h and β. The following lemma shows that the solutions of the unfolded system satisfy the original equations.

Lemma 7.1. Let $h_* \in C_p^k(\mathbb{R}) := \{h \in C^k | h \text{ is 1-periodic}\}$, and $\beta_* \in \mathbb{R}$. Consider

$$\begin{array}{rccl} \Psi\colon & \mathbb{R} \times C_p^k(\mathbb{R}) & \longrightarrow & \mathbb{R} \times C_p^k(\mathbb{R}) \\ & (\beta, h) & \longmapsto & ((p - h(0)) \cdot \eta, F(h(\theta)) - (1 + \beta)h(\theta + \rho)) \end{array}$$

If $\Psi(\beta_*, h_*) = 0$, then $\beta_* = 0$ and h conjugates the dynamics generated by F to the rotation map T_ρ.

Proof. This proof comes from [9]. Let $\beta_* \in \mathbb{R}$ and $h_* \in C_p^k(\mathbb{R})$ be a zero of Ψ. Then,

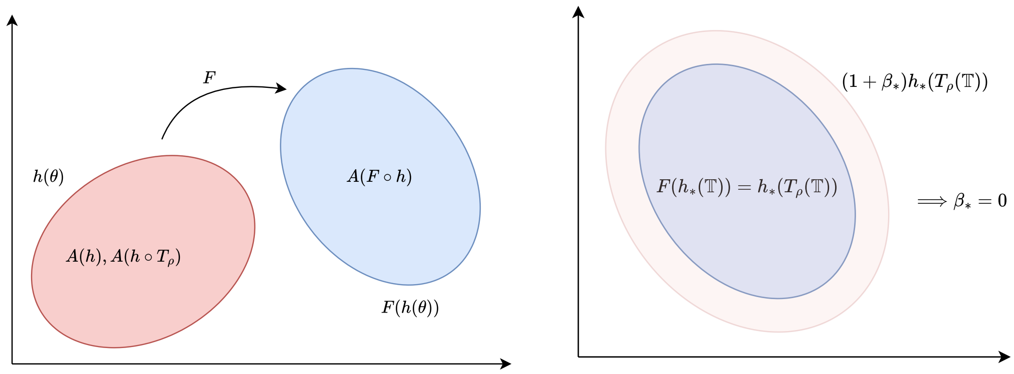
$$F(h_*(\theta)) = (1 + \beta_*)h_*(\theta + \rho)$$

Let Γ be the curve parametrised by h_*, and $\hat{\Gamma} := F \circ \Gamma$ be the curve parametrised by $F \circ h_*$. Note that $h_*(\theta + \rho)$ is just a reparametrisation with different phase of Γ.

For an arbitrary $h \in C_p^k(\mathbb{R})$, consider the area enclosed by the simple closed curves Γ (with both parameterisations) and $\hat{\Gamma}$:

$$A(h) = \frac{1}{2} \int_\Gamma h_1 dy - h_2 dx = \frac{1}{2} \int_0^1 \Big(h_1(\theta) \frac{d}{d\theta} h_2(\theta) - h_2(\theta) \frac{d}{d\theta} h_1(\theta) \Big) d\theta$$

$$\begin{aligned} A(h \circ T_\rho) &= \frac{1}{2} \int_\Gamma (h_1 \circ T_\rho) dy - (h_2 \circ T_\rho) dx \\ &= \frac{1}{2} \int_0^1 \Big(h_1(\theta + \rho) \frac{d}{d\theta} h_2(\theta + \rho) - h_2(\theta + \rho) \frac{d}{d\theta} h_1(\theta + \rho) \Big) d\theta \end{aligned}$$



$$\begin{aligned} A(F \circ h) &= \frac{1}{2} \int_{\hat{\Gamma}} (F \circ h)_1 dy - (F \circ h)_2 dx \\ &= \frac{1}{2} \int_0^1 \Big(F(k(\theta))_1 \frac{d}{d\theta} F(k(\theta))_2 - F(k(\theta))_2 \frac{d}{d\theta} F(k(\theta))_1 \Big) d\theta \end{aligned}$$

We know that $A(h) = A(h \circ T_\rho)$, as they are computed over the same curve Γ. Due to the assumption that F is a diffeomorphism, $\hat{\Gamma}$ is diffeomorphic to Γ. Since, F is a measure-preserving map, $A(h) = A(F \circ h)$.

However, plugging $F(h_*(\theta)) = (1 + \beta_*)h_*(\theta + \rho)$ into the areas formulae, one gets that $A(F \circ h_*) = (1 + \beta_*)A(h_* \circ T_\rho)$. As we have just seen that $A(h_*) = A(h_* \circ T_\rho) = A(F \circ h_*)$, it follows that $\beta_* = 0$. From this, we obtain that $F(h_*(\theta)) = h_*(\theta + \rho)$; hence h_* conjugates the dynamics of F to an irrational rotation T_ρ. □

7.2 Newton Scheme in Fourier Coefficient Space

By the previous Lemma 7.1, we now seek to solve $\Psi(\beta, h) = 0$. Assume h_0 is an approximate zero of the equation and choose $\beta_0 = 0$. The Newton sequence is

$$\begin{pmatrix} \beta_{n+1} \\ h_{n+1} \end{pmatrix} = \begin{pmatrix} \beta_n \\ h_n \end{pmatrix} + \begin{pmatrix} \delta_n \\ \Delta_n \end{pmatrix}$$

where $(\delta_n, \Delta_n)^T$ satisfies $D\Psi(\beta_n, h_n)(\delta_n, \Delta_n)^T = -\Psi(\beta_n, h_n)$. We refer to $D\Psi(\beta, h)$ as the Frechet derivative of Ψ, thus for $\beta, \delta \in \mathbb{R}$ and $h, \Delta \in C_p^k(\mathbb{R})$,

$$D\Psi(\beta, h) \begin{pmatrix} \delta \\ \Delta \end{pmatrix} = \begin{pmatrix} -\Delta(0) \cdot \eta \\ -\delta h(\theta + \rho) + DF(h(\theta))\Delta(\theta) - (1 + \beta)\Delta(\theta + \rho) \end{pmatrix}$$

This refinement of the approximation of h allows us to add an extra step to the numerical recipe we gave in Section 6.3 while working with the Henon map.

Step 5: Perform the Newton method with an initial condition $\beta_0 = 0$ and h_0 defined as in the previous step. Iterate until the defect

$$\epsilon_n = \sup_{\theta \in \mathbb{T}} \left| F(h_n(\theta)) - h_n(\theta + \rho) \right|$$

is either smaller than some tolerance or saturates. Then, we consider the conjugacy h we were seeking as the last iterate h_n of the Newton scheme. If convergence is not achieved, return to Step 4 and refine N_F.

We won't do this step in our examples, as it requires some further studies to implement the corresponding code. Interested readers can find a detailed explanation in [9]. From the theoretical point of view, this justifies that close to the numerical approximation there is a "true" invariant torus.

Recall h is 1-periodic and thus $h(\theta) = \sum_{n\in\mathbb{Z}} \binom{a_n}{b_n} e^{2\pi i n\theta}$. Using this assumption about the form of the unknown function facilitates finding the solution. In fact, one can notice that the translation by ρ is a diagonal operator in Fourier space as

$$h \circ T_\rho(\theta) = h(\theta + \rho) = \sum_{n\in\mathbb{Z}} \begin{pmatrix} a_n \\ b_n \end{pmatrix} e^{2\pi i n(\theta+\rho)} = \sum_{n\in\mathbb{Z}} e^{2\pi i n\rho} \begin{pmatrix} a_n \\ b_n \end{pmatrix} e^{2\pi i n\theta}$$

The phase condition can also be written in terms of the Fourier series

$$(p - h(0)) \cdot \eta = p_1\eta_1 + p_2\eta_2 - \sum_{n\in\mathbb{Z}} (a_n\eta_1 + b_n\eta_2)$$

In addition, assuming $F \in C^k$, since $h \in C^k_p$, then $F \circ h \in C^k_p$. Hence,

$$F \circ h(\theta) = \sum_{n\in\mathbb{Z}} (F \circ h)_n e^{2\pi i n\theta}$$

where its Fourier coefficients $(F \circ h)_n$ depend in a nonlinear way on a_n and b_n, since $F \circ h(\theta) = h \circ T_\rho(\theta)$.

In conclusion, writing h using its Fourier series representation makes $h \circ T_\rho$ very easy to compute, as a translation by ρ is a diagonal operator in the Fourier space. The equation that determines the phase condition can also be managed using this representation of h. The map $F \circ h$ may be difficult to compute, but we can exploit the fact that $h \circ T_\rho = F \circ h$ to obtain its Fourier series. Therefore, solving $\Psi(\beta, h) = 0$ using the Newton scheme is usually done in the Fourier space (truncating up to a certain order).

Appendix A

Codes used to produce the graphics

A.1 Standard Map

The following code asks the user to introduce the number of iterates $N \in \mathbb{N}$, the perturbation $\epsilon \in \mathbb{R}$, and the first coordinate of the initial condition $x_0 \in \mathbb{T}$. It then computes the forward orbit segment of length N for the *Standard map* given the initial condition (x_0, y_0), where $y_0 = 0.01k \in [-0.5, 0.5]$ and $k \in \mathbb{Z}$. The resulting data is saved to a file so the orbits can be later plotted. This code has been used to generate Figure 5.1.

```
/* Orbits of the Standard map */

#include <stdio.h>
#include <math.h>
#include <stdlib.h>

#define PI 4*atan(1.)

void standard_map (double z[], double EPSILON);


int main (void) {
  double iy = 0.01, x0, y0, z[2];
  double EPSILON;
  FILE *out;
  int n, N;

  /* User data */
  printf ("\nGive me the number of iterates:\n");
  scanf (" %d", &N);
```

```c
  if (N<=0) {
    printf ("It must be positive. End\n\n");
    return 1;
  }
  printf ("Give me the perturbation (<1):\n");
  scanf (" %le", &EPSILON);
  if (EPSILON>=1) {
    printf ("This perturbation is too big.\n\n");
    return 1;
  }
  printf ("\nGive me x0 for the initial condition (which will be of the form (x0,y0)):\n");
  scanf (" %le", &x0);
  x0=x0-floor(x0);
  if (x0>=0.5) {
    x0-=1;
  }

  /* File */
  out=fopen("orbits_standardKAM.data", "w");
  if (out==NULL) {
    printf ("Problems with the file. End.\n\n");
    exit(1);
  }
  fprintf (out, "#x0%13cy0\n", ' ');

  /* Orbits */
  printf ("\n -------- computing orbits of the standard map --------\n");
  for (y0=-0.5; y0<=0.5; y0 += iy) {
    z[0]=x0;
    z[1]=y0;
    for (n=0; n<N; n++) {
      fprintf (out, "%10.8le\t%10.8le\n", z[0], z[1]);
      standard_map(z, EPSILON);
    }
  }

  fclose (out);
  printf ("Results can be found in orbits_standardKAM.data.\n\n");
  return 0;
}


/* Standard Map */
void standard_map (double z[], double EPSILON) {
  z[1]=z[1]-EPSILON/(2*PI)*sin(2*PI*z[0]);
  z[0]=z[0]+z[1];
  z[0]=z[0]-floor(z[0]);
```

```
  z[1]=z[1]-floor(z[1]);
  if (z[0]>=0.5) {
    z[0]-=1;
  }
  if (z[1]>=0.5) {
    z[1]-=1;
  }
}
```

The following code computes the rotation number for the primary tori of the *Standard map* via weighted Birkhoff averages. At the beginning, the user must introduce the number of iterates $N \in \mathbb{N}$, the perturbation $\epsilon \in \mathbb{R}$, and the first coordinate of the initial condition $x_0 \in \mathbb{T}$. The resulting data is saved in a file for subsequent plotting. This is used to generate Figure 6.3 (a).

```
/* weighted Birkhoff averages to compute the rotation vector of the Standard map */

#include <stdio.h>
#include <math.h>
#include <stdlib.h>

#define PI 4*atan(1.)

void standard_map (double z[], double EPSILON);
double exp_weights (double t);
double weighted_birkhoff (double z[], int N, double EPSILON);


int main (void) {
  double z[2], wb, x0, y0, iy=0.001, EPSILON;
  int N;
  FILE *out;

  /* User data */
  printf ("\nGive me the number of iterates:\n");
  scanf (" %d", &N);
  if (N<=0) {
    printf ("It must be positive. End\n\n");
    return 1;
  }
  printf ("Give me the perturbation (<1):\n");
  scanf (" %le", &EPSILON);
  if (EPSILON>=1) {
    printf ("This perturbation is too big.\n\n");
    return 1;
  }
  printf ("\nGive me x0 for the initial condition (which will be of the
```

```
    form (x0,y0)):\n");
  scanf (" %le", &x0);
  x0=x0-floor(x0);


  /* File */
  out=fopen("wb_standardmap.data", "w");
  if (out==NULL) {
    printf ("Problems with the file. End.\n\n");
    exit(1);
  }
  fprintf (out, "#y0%14cwb[0]\n", ' ');


  /* weighted Birkhoff */
  printf ("\n ---------- computing the Weighted Birkhoff average ----------\n");
  for (y0=-0.5; y0<=0.5; y0+=iy) {
    z[0]=x0;
    z[1]=y0;
    wb=weighted_birkhoff(z, N, EPSILON);
    fprintf (out, "%10.8le\t%+14.8le\n", y0, wb);
  }

  fclose (out);
  printf ("Results can be found in wb_standardmap.data.\n\n");

  return 0;
}


/* Standard Map */
void standard_map (double z[], double EPSILON) {
  z[1]=z[1]-EPSILON/(2*PI)*sin(2*PI*z[0]);
  z[0]=z[0]+z[1];
}


/* Exponential weighting */
double exp_weights (double t) {
  if (t<=0||t>=1) {
     return 0.;
    }
  return exp(-1/(t*(1-t)));
}


/* Weighted Birkhoff average */
double weighted_birkhoff (double z0[], int N, double EPSILON) {
```

```
  int n;
  double z1[2], weight, aN=0., wb=0.;
  for (n=0; n<N; n++) {
    z1[0]=z0[0];
    z1[1]=z0[1];
    standard_map(z1, EPSILON);
    weight=exp_weights((double)n/N);
    wb+=weight*(z1[0]-z0[0]);
    aN+=weight;
    z0[0]=z1[0]-floor(z1[0]);
    z0[1]=z1[1]-floor(z1[1]);
  }
  return wb/aN;
}
```

The following code computes the difference between the rotation number of primary tori of the *Standard map* $|\rho_{N_2} - \rho_{N_1}|$, for different $N_1, N_2 \in \mathbb{N}$. At the beginning, the user must introduce a pair of numbers $N_1, N_2 \in \mathbb{N}$, the perturbation $\epsilon \in \mathbb{R}$ and the first coordinate of the initial condition $x_0 \in \mathbb{T}$. The resulting data is saved in two separate files so it can be later plotted. This is used to generate Figure 6.3 (b).

```
/* Comparision of the speed of convergence with B and WB averages */
#include <stdio.h>
#include <math.h>
#include <stdlib.h>

#define PI 4*atan(1.)

void standard_map (double z[], double EPSILON);
double exp_weights (double t);
void rotation_number(double z[2], double EPSILON, int N1, int N2, double
    rho[4]);


int main (void) {
  double x0, y0, iy=0.001, EPSILON;
  double z[2], rho[4];
  int N1, N2;
  FILE *out1, *out2;

  /* User data */
  printf ("\nGive me the number of iterates N1 < N2 for which you want to
     compare the result:\n");
  scanf (" %d %d", &N1, &N2);
  if (N1>=N2) {
    printf ("Error: N1>=N1\n\n");
    return 1;
```

```
  }
  printf ("\nGive me the perturbation (<1):\n");
  scanf (" %le", &EPSILON);
  if (EPSILON>=1) {
    printf ("This perturbation is too big.\n\n");
    return 1;
  }
  printf ("\nGive me x0 for the initial condition (which will be of the form (x0,y0)):\n");
  scanf (" %le", &x0);
  x0=x0-floor(x0);

  /* Files */
  out1=fopen("speedB.data", "w");
  if (out1==NULL) {
    printf ("Problems with the file. End.\n\n");
    exit(1);
  }
  fprintf (out1, "#x0%14difference\n", ' ');
  out2=fopen("speedWB.data", "w");
  if (out2==NULL) {
    printf ("Problems with the file. End.\n\n");
    exit(1);
  }
  fprintf (out2, "#x0%14cdifference\n", ' ');

  /* Birkhoff */
  printf ("\n ------- computing normal and weighted Birkhoff averages -------\n");
  for (y0=-0.5; y0<0.5; y0+=iy) {
    z[0]=x0;
    z[1]=y0;
    rotation_number(z, EPSILON, N1, N2, rho);
    fprintf (out1, "%+14.8le\t%+14.8le\n", y0, fabs(rho[1]-rho[0]));
    fprintf (out2, "%+14.8le\t%+14.8le\n", y0, fabs(rho[3]-rho[2]));
  }

  fclose (out1);
  fclose (out2);
  printf ("Results can be found in speedB.data.\n");
  printf ("Results can be found in speedWB.data.\n\n");
  return 0;
}


/* Standard Map */
void standard_map (double z[], double EPSILON) {
  z[1]=z[1]-EPSILON/(2*PI)*sin(2*PI*z[0]);
```

```c
  z[0]=z[0]+z[1];
}


/* exponential weightings */
double exp_weights (double t) {
  if (t<=0||t>=1) {
     return 0.;
    }
  return exp(-1/(t*(1-t)));
}

/* Rotation Number through weighted Birkhoff averages*/
void rotation_number(double z0[2], double EPSILON, int N1, int N2, double
     rho[4]) {
  int n;
  double weight1, weight2, aN1=0., aN2=0., z1[2], diff;
  rho[0]=0.;
  rho[2]=0.;
  rho[3]=0.;
  for (n=0; n<N1; n++) {
    z1[0]=z0[0];
    z1[1]=z0[1];
    standard_map(z1, EPSILON);
    weight1=exp_weights((double)n/N1);
    weight2=exp_weights((double)n/N2);
    diff=z1[0]-z0[0];
    rho[0]+=diff;
    rho[2]+=weight1*diff;
    rho[3]+=weight2*diff;
    aN1+=weight1;
    aN2+=weight2;
    z0[0]=z1[0]-floor(z1[0]);
    z0[1]=z1[1]-floor(z1[1]);
  }
  rho[1]=rho[0];
  rho[0]=rho[0]/N1;
  rho[2]=rho[2]/aN1;
  for (n=N1; n<N2; n++) {
    z1[0]=z0[0];
    z1[1]=z0[1];
    standard_map(z1, EPSILON);
    weight2=exp_weights((double)n/N2);
    diff=z1[0]-z0[0];
    rho[1]+=diff;
    rho[3]+=weight2*diff;
    aN2+=weight2;
    z0[0]=z1[0]-floor(z1[0]);
```

```
    z0[1]=z1[1]-floor(z1[1]);
  }
  rho[1]=rho[1]/N2;
  rho[3]=rho[3]/aN2;
}
```

A.2 Arnold Map

Given the perturbation $\epsilon \in \mathbb{R}$, the orbit length $N \in \mathbb{N}$, and an initial condition $x_0 \in \mathbb{T}$; the following code computes the rotation number and the Lyapunov exponent of the *Arnold map* for different values of $\alpha \in [0,1]$. The resulting data is saved in two separate files, so we can later plot them. This is used to generate Figure 6.1 and Figure 6.2.

```
/* Devil's staircase and Lyapunov exponent of the Arnold map */

#include <stdio.h>
#include <math.h>
#include <stdlib.h>

double EPSILON;
#define PI 4*atan(1.)

double diff_arnold_map (double x, double ALPHA);
double exp_weights (double t);
void weighted_birkhoff(double x, double ALPHA, int N, double wb[2]);

int main (void) {
  double x0, ALPHA, wb[2];
  int N;
  FILE *out1, *out2;

  /* User data */
  printf ("\nGive me the number of iterates:\n");
  scanf(" %d", &N);
  if (N<=0) {
    printf ("It must be positive. End.\n\n");
    return 1;
  }
  printf ("\nGive me epsilon:\n");
  scanf (" %le", &EPSILON);
  printf ("\nGive me the initial condition x0:\n");
  scanf (" %le", &x0);

  /* Files */
  out1=fopen("devil_arnoldmap.data", "w");
```

```
  if (out1==NULL) {
    printf ("Problems with the file. End.\n\n");
    exit(1);
  }
  out2=fopen("lyapunov_arnoldmap.data", "w");
  if (out2==NULL) {
    printf ("Problems with the file. End.\n\n");
    exit(1);
  }

  /* weighted Birkhoff */
  printf("\n ------ Computing the rotation number and Lyapunov exponents ------\n");
  for (ALPHA=0; ALPHA<1; ALPHA+=0.001) {
    weighted_birkhoff(x0, ALPHA, N, wb);
    fprintf(out1, "%10.8le\t%10.8le\n", ALPHA, wb[0]);
    fprintf(out2, "%10.8le\t%10.8le\n", ALPHA, wb[1]);
  }

  fclose(out1);
  fclose(out2);
  printf ("Results can be found in devil_arnoldmap.data.\n");
  printf ("Results can be found in lyapunov_arnoldmap.data.\n\n");

  return 0;
}


/* difference between iterates of the Arnold Circle map */
double diff_arnold_map (double x, double ALPHA) {
  return ALPHA-EPSILON/(2*PI)*sin(2*PI*x);
}


/* differencial of the Arnold Map */
double deriv_arnold_map(double x) {
  return 1-EPSILON*cos(2*PI*x);
}


/* exponential weightings */
double exp_weights (double t) {
  if (t<=0||t>=1) {
     return 0.;
    }
  return exp(-1/(t*(1-t)));
}
```

```
/* weighted Birkhoff */
void weighted_birkhoff(double x, double ALPHA, int N, double wb[2]) {
  double diff, aN=0., weight;
  int n;
  wb[0]=0.; /* rotation number */
  wb[1]=0.; /* lyapunov exponent */
  for (n=0; n<N; n++) {
    x=x-floor(x);
    diff=diff_arnold_map(x, ALPHA);
    weight=exp_weights((double)n/N);
    wb[0]+=weight*diff;
    wb[1]+=weight*log(fabs(deriv_arnold_map(x)));
    aN+=weight;
    x+=diff;
  }
  wb[0]/=aN;
  wb[1]/=aN;
}
```

A.3 Henon Map

Analogous to the code previously presented for the *Standard map*, this code aims to plot the orbits of the *Henon map*. Given a pre-fixed $\alpha \in \mathbb{R}$, the following code asks the user to introduce the orbit length $N \in \mathbb{N}$ and the first coordinate of an initial condition $x_0 \in \mathbb{T}$. It then computes the forward orbit segment of length N of the *Henon map* with the initial condition (x_0, y_0), where $y_0 = 0.01k \in [-0.7, 0.7]$ and $k \in \mathbb{Z}$. It saves the resulting data in a file so the orbits can later be plotted. This code has been used to produce Figure 6.4.

```
/* Orbits of the Henon map */

#include <stdio.h>
#include <math.h>
#include <stdlib.h>

#define PI 4*atan(1.)
#define ALPHA acos(0.24)

void henon_map (double z[2]);
void orbit (double x0, double y0);

int main (void) {
  double iy = 0.01, x0, y0, z[2];
  int n, N;
  FILE *out;
```

```c

  /* User data */
  printf ("\nWe have set alpha = %.16le.\n", ALPHA);
  printf ("Give me the number of iterates:\n");
  scanf(" %d", &N);
  if (N<=0) {
    printf ("It must be positive. End.\n\n");
    return 1;
  }
  printf ("\nGive me x0 for the initial condition (which will be of the form (x0,y0)):\n");
  scanf (" %le", &x0);
  x0=x0-floor(x0);

  /* File */
  out=fopen("orbits_henonmap.data", "w");
  if (out==NULL) {
    printf ("Problems with the file. End.\n\n");
    exit(1);
  }
  fprintf (out, "#x0%13cy0\n", ' ');

  /* Orbits */
  printf ("\n --------- computing orbits of the henon map ---------\n");
  for (y0=-0.7; y0<=0.7; y0 += iy) {
    z[0]=x0;
    z[1]=y0;
    for (n=0; n<N; n++) {
      fprintf (out, "%10.8le\t%10.8le\n", z[0], z[1]);
      henon_map(z);
    }
  }

  fclose (out);
  printf ("Results can be found in orbits_henonmap.data.\n\n");
  return 0;
}


/* Henon map */
void henon_map (double z[2]) {
    double bz[2];
    bz[0] = z[0]*cos(ALPHA) - (z[1]-z[0]*z[0])*sin(ALPHA);
    bz[1] = z[0]*sin(ALPHA) + (z[1]-z[0]*z[0])*cos(ALPHA);
    z[0]= bz[0];
    z[1]= bz[1];
}
```

The following code aims to compute the rotation number of an invariant tori of the *Henon map*, with a pre-fixed α, given an initial condition $(x_0, y_0) \in \mathbb{R}$ in a neighbourhood of the elliptic fixed point $(0,0)$. This is used in the Step 1 and Step 2 of the numerical recipe in Section 6.3.

```
/* Rotation number for the Henon map. STEP 1 and STEP 2*/
#include <stdio.h>
#include <math.h>
#include <stdlib.h>

#define PI 4*atan(1.)
#define ALPHA acos(0.24)

void henon_map (double z[]);
double exp_weights (double t);
double WB_rotation_number(double z0[], int N);

int main (void) {
    double z[2], rho;
    int N;

    /* User data */
    printf("\nWe have set alpha = %le\n", ALPHA);
    printf ("Give me the number of iterates:\n");
    scanf(" %d", &N);
    if (N<=0) {
     printf ("It must be positive. End.\n\n");
     return 1;
    }
    printf ("\nGive me the initial condition (x0, y0) close to (0,0):\n");
    scanf ("%le %le", &z[0], &z[1]);

    /* weighted Birkhoff */
    printf("\n——Computing the rotation number——\n");
    rho = WB_rotation_number(z, N);
    printf("The rotation number of the orbit is %.15le\n", rho);
    printf("The rotation number of the origin is %.15le\n\n", ALPHA/(2*PI));

    return 0;
}


/* Henon map*/
void henon_map (double z[]) {
    double aux[2];
    aux[0] = z[0];
    aux[1] = z[1];
    z[0] = aux[0]*cos(ALPHA)-(aux[1]-aux[0]*aux[0])*sin(ALPHA);
```

```
    z[1] = aux[0]*sin(ALPHA)+(aux[1]-aux[0]*aux[0])*cos(ALPHA);
}


/* weightings */
double exp_weights (double t) {
    if (t<=0||t>=1) {
     return 0.;
    }
    return exp(-1/(t*(1-t)));
}


/* wieghted birkhoff average */
double WB_rotation_number(double z[], int N) {
    double theta1, theta0, sum=0., weight, An=0., ang;
    int n;
    theta0=atan2(z[1],z[0]);
    for (n=0; n<N; n++) {
     henon_map(z);
     theta1=atan2(z[1],z[0]);
     ang=theta1-theta0;
     if (ang<0) {
         ang+=2*PI;
       }
     weight=exp_weights((double) n/N);
     sum+=weight*ang;
     An+=weight;
     theta0=theta1;
    }
    return sum/An/(2*PI);
}
```

The subsequent code aims to compute the ideal truncation index of the Fourier series approximation of the conjugacy. A priori, one must fix the tolerance $tol \in \mathbb{R}$ and the parameter $\alpha \in \mathbb{R}$ of the *Henon map*. The code asks the user to introduce the desired orbit length $N \in \mathbb{N}$, an initial condition $(x_0, y_0) \in \mathbb{R}^2$, and the rotation number ρ. Recall that, for this to work, the initial condition must lie in an invariant tori with an associated rotation number equal to the one that is given. Then, the code searches for the first index such that the norm of the corresponding Fourier coefficient of h is smaller than the pre-established tolerance. This is used in Step 3 of the numerical recipe in Section 6.3.

```
/* Truncation of the Fourier series representation. STEP 3 */
#include <stdio.h>
#include <math.h>
#include <stdlib.h>
```

```c

#define tol 1.0e-10
#define PI 4*atan(1.)
#define ALPHA acos(0.24)

void henon_map (double z[]);
double exp_weights (double t);
void fourier_coefficients(double z[2], int k, int N, double rho, double az[2], double bz[2]);

int main (void) {
  double z[2], az[2],bz[2];
  double norm1=10, norm2, rho;
  int k, N;

  /* User data */
  printf("\nWe have set alpha = %.15le\n", ALPHA);
  printf ("Give me the number of iterates:\n");
  scanf(" %d", &N);
  if (N<=0) {
    printf ("It must be positive. End.\n");
    return 1;
  }
  printf ("\nGive me the initial condition (x0, y0) that follows a quasiperiodic trajectory:\n");
  scanf (" %le %le", &z[0], &z[1]);
  printf ("Give me its rotation number:\n");
  scanf(" %le", &rho);

  /* weighted Bitkhoff */
  printf("\n ------- Computing the ideal truncation of the Fourier series -------\n");
  for (k=0; k<N && norm1>tol; k+=10) {
    fourier_coefficients(z, k, N, rho, az, bz);
    norm1=sqrt(az[0]*az[0]+bz[0]*bz[0]);
    norm2=sqrt(az[1]*az[1]+bz[1]*bz[1]);
    if (norm2>norm1) {
      norm1=norm2;
    }
    printf("Norm[%d]= %.15le\n", k, norm1);
  }
  if (k==N) {
    printf("\nFourier coefficients have not reach 0. Try to increment N.\n\n");
  } else {
    printf ("\nWe can truncate at N0=%d.\n\n", k);
  }

```

```
  return 0;
}

void henon_map (double z[]) {
  double aux[2];
  aux[0] = z[0];
  aux[1] = z[1];
  z[0] = aux[0]*cos(ALPHA)-(aux[1]-aux[0]*aux[0])*sin(ALPHA);
  z[1] = aux[0]*sin(ALPHA)+(aux[1]-aux[0]*aux[0])*cos(ALPHA);
}

double exp_weights (double t) {
  if (t<=0||t>=1) {
     return 0.;
    }
  return exp(-1/(t*(1-t)));
}


/* Given an index k, we compute its Fourier coefficient (az[0]+ibz[0],az
   [1]+ibz[1]) */
void fourier_coefficients(double z[2], int k, int N, double rho, double
    az[2], double bz[2]) {
  int n;
  double weight, An;
  az[0]=0.;
  az[1]=0.;
  bz[0]=0.;
  bz[1]=0.;
  An=0.;
  /* Computation of each coefficient */
  for (n=0; n<N; n++) {
    weight=exp_weights((double)n/N);
    az[0]+=weight*z[0]*cos(2*PI*k*n*rho);
    az[1]+=weight*z[1]*cos(2*PI*k*n*rho);
    bz[0]+=weight*z[0]*sin(2*PI*k*n*rho);
    bz[1]+=weight*z[1]*sin(2*PI*k*n*rho);
    An+=weight;
    henon_map(z);
  }
  if (k==0) {
    az[0]=az[0]/An;
    az[1]=az[1]/An;
    bz[0]=bz[0]/An;
    bz[1]=bz[1]/An;
  } else {
    az[0]=2*az[0]/An;
    az[1]=2*az[1]/An;
```

```
    bz[0]=2*bz[0]/An;
    bz[1]=2*bz[1]/An;
  }
}
```

Finally, the next code computes an initial approximation h_0 of h and its initial defect. Similar to the previous code, the parameter $\alpha \in \mathbb{R}$ of the *Henon map* must be predefined. Then, the user is asked to introduce the orbit length N, the truncation of the Fourier coefficients N_F, and the initial condition (x_0, y_0) along with its rotation number ρ. Typically, N_F is chosen about %15 of the N_0 obtained from the previous code. After that, the program proceeds to compute the Fourier approximation h_0 and its initial defect. If the initial defect is not sufficiently small, the process can be repeated with a larger N_F. This is used in Section 6.3, Step 4 of the Numerical Recipe.

```
/* Initial guess of h0 for the Henon Map. STEP 4 */
#include <stdio.h>
#include <math.h>
#include <stdlib.h>

#define PI 4*atan(1.)
#define ALPHA acos(0.24)
#define EXP 2.718281828459045

void henon_map (double z[]);
double exp_weights (double t);
void fourier_coefficients(double **z, int k, int N, double rho, double **az, double **bz);
void evaluate_h (double theta, int NF, double **az, double **bz, double hz[2]);
double error_conjugacy (double theta, int NF, double rho, double **a, double **b);

int main (void) {
  double **z, **az, **bz;
  double rho, error, theta, defect=0.;
  int k, n, N, NF;

  /* User data */
  printf("\nWe have set alpha = %.15le\n", ALPHA);
  printf ("Give me the number of iterates N and the truncation of the Fourier Series NF:\n");
  scanf(" %d %d", &N, &NF);
  if (N<=0 ||NF<=0||NF>N) {
    printf ("They must be positive. End.\n");
    return 1;
  }
```

```
az=(double**)calloc(NF,sizeof(double*));
bz=(double**)calloc(NF,sizeof(double*));
z=(double**)calloc(N,sizeof(double*));
if (az==NULL || bz==NULL) {
  printf ("Error with calloc. End.\n");
  exit(1);
}
for(k=0; k<NF; k++) {
  az[k]=(double*)calloc(2,sizeof(double));
  bz[k]=(double*)calloc(2,sizeof(double));
  if (az[k]==NULL || bz[k]==NULL) {
    printf ("Error with calloc. End.\n");
    exit(1);
  }
}
for(n=0; n<N; n++) {
  z[n]=(double*)calloc(2,sizeof(double));
  if (z[n]==NULL) {
    printf ("Error with calloc. End.\n");
    exit(1);
  }
}
printf ("Give me the initial condition (x0, y0):\n");
scanf (" %le %le", &z[0][0], &z[0][1]);
printf ("Give me its rotation number:\n");
scanf(" %le", &rho);

/* Computing all the objects needed */
printf("\n------ Computing iterates ------\n");
for (n=1; n<N; n++) {
  z[n][0]=z[n-1][0];
  z[n][1]=z[n-1][1];
  henon_map(z[n]);
}
printf("\n------ Computing Fourier coefficients ------\n");
for (k=0; k<NF; k++) {
  fourier_coefficients(z,k,N,rho,az,bz);
}

/* Looking if it's good */
printf("\n------ Looking if it is a good initial condition ------\n");
for (theta=0; theta<1; theta+=0.001) {
  error=error_conjugacy(theta,NF,rho,az,bz);
  if (error>defect) {
    defect=error;
  }
}
printf ("\nThe initial deffect is of %.15le\n\n", defect);
```

```


  for (k=0; k<NF; k++) {
    free(az[k]);
    free(bz[k]);
  }
  for(n=0;n<N;n++) {
    free(z[n]);
  }
  free(az);
  free(bz);
  free(z);
  return 0;
}


/* computes the following iterate of the Henon map */
void henon_map (double z[]) {
  double aux[2];
  aux[0]=z[0];
  aux[1]=z[1];
  z[0]=aux[0]*cos(ALPHA)-(aux[1]-aux[0]*aux[0])*sin(ALPHA);
  z[1]=aux[0]*sin(ALPHA)+(aux[1]-aux[0]*aux[0])*cos(ALPHA);
}


/* weightings */
double exp_weights (double t) {
  if (t<=0||t>=1) {
      return 0.;
    }
  return exp(-1/(t*(1-t)));
}


/* saves in az and bz the fourier coefficients*/
void fourier_coefficients(double **z, int k, int N, double rho, double **az, double **bz) {
  int n;
  double weight, An=0.;
  az[k][0]=0.;
  az[k][1]=0.;
  bz[k][0]=0.;
  bz[k][1]=0.;
  /* Computation of each coefficient*/
  for (n=0; n<N; n++) {
    weight=exp_weights((double)n/N);
    az[k][0]+=weight*z[n][0]*cos(2*PI*k*n*rho);
```

```
    az[k][1]+=weight*z[n][1]*cos(2*PI*k*n*rho);
    bz[k][0]+=weight*z[n][0]*sin(2*PI*k*n*rho);
    bz[k][1]+=weight*z[n][1]*sin(2*PI*k*n*rho);
    An+=weight;
  }
  if (k==0) {
    az[0][0]=az[0][0]/An;
    az[0][1]=az[0][1]/An;
    bz[0][0]=bz[0][0]/An;
    bz[0][1]=bz[0][1]/An;
  } else {
    az[k][0]=2*az[k][0]/An;
    az[k][1]=2*az[k][1]/An;
    bz[k][0]=2*bz[k][0]/An;
    bz[k][1]=2*bz[k][1]/An;
  }
}


/* evaluates the series approximation of h at theta */
void evaluate_h (double theta, int NF, double **az, double **bz, double
    h_theta[2]) {
  int k;
  double sinus, cosinus;
  h_theta[0]=az[0][0];
  h_theta[1]=az[0][1];
  for (k=1; k<NF; k++) {
    sinus=sin(2*PI*k*theta);
    cosinus=cos(2*PI*k*theta);
    h_theta[0]+=az[k][0]*cosinus+bz[k][0]*sinus;
    h_theta[1]+=az[k][1]*cosinus+bz[k][1]*sinus;
  }
}


/* returns the defect of the approximation for theta */
double error_conjugacy (double theta, int NF, double rho, double **az,
    double **bz) {
  double h_theta[2], h_thetarho[2];
  double norm;
  evaluate_h(theta, NF, az, bz, h_theta);
  henon_map(h_theta);
  evaluate_h(theta+rho, NF, az, bz, h_thetarho);
  h_theta[0]-=h_thetarho[0];
  h_theta[1]-=h_thetarho[1];
  norm=sqrt(h_theta[0]*h_theta[0]+h_theta[1]*h_theta[1]);
  return norm;
}
```

Remark A.1. The codes use the real Fourier series expression for ease of implementation, instead of the complex one.

Bibliography

[1] E. Sander, and J.A. Yorke. (2015). The many facets of chaos. *International Journal of Bifurcation and Chaos*, 25(04), 1530011.

[2] S. Das, and J.A. Yorke. (2018). Super convergence of ergodic averages for quasiperiodic orbits. *Nonlinearity*, 31(2), 491.

[3] A. Katok, and B. Hasselblatt. (1995). *Introduction to the modern theory of dynamical systems* (No. 54). Cambridge university press.

[4] C. Ulcigrai. (2019). *Dynamical Systems and Ergodic Theory*. Part III.

[5] U. Krengel. (1978). On the speed of convergence in the ergodic theorem. *Monatshefte für Mathematik*, 86(1), 3-6.

[6] V.I. Arnol'd. (2013). *Mathematical methods of classical mechanics* (Vol. 60). Springer Science & Business Media.

[7] S. Das, Y. Saiki, E. Sander, and J.A. Yorke. (2017). Quantitative Quasiperiodicity. *Nonlinearity*, 30(11), 4111.

[8] M.R. Herman. (1979). Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations. *Publications Mathématiques de l'IHÉS*, 49, 5-233.

[9] D. Blessing, and J.D. James. (2023). Weighted Birkhoff Averages and the Parameterization Method. *arXiv preprint arXiv:2306.16597*.

[10] S. Das, Y. Saiki, E. Sander, and J.A. Yorke. (2016). Quasiperiodicity: rotation numbers. *The Foundations of Chaos Revisited: From Poincaré to Recent Advancements*, 103-118.

[11] M. S. Raghunathan. (1979). A proof of Oseledec's multiplicative ergodic theorem. *Israel Journal of Mathematics*, 32, 356-362.