Quantitative Statistical Stability and Linear Response for Irrational Rotations and Diffeomorphisms of the Circle

Stefano Galatolo and Alfonso Sorrentino

Abstract. We show statistical stability for a large class of small $C^0$ perturbations of circle diffeomorphisms with irrational rotation number. We show that if the rotation number is Diophantine the invariant measure varies in a Hölder way under perturbation of the map and the Hölder exponent depends on the Diophantine type of the angle. The perturbations allowed includes the ones coming from spatial discretization and hence numerical truncation. We also show linear response for smooth perturbations that preserve the rotation number and other perturbations. This is done by means of classical tools from KAM theory, while the quantitative stability results are obtained by transfer operator techniques.

1. Introduction

Understanding the statistical properties of a certain dynamical system is of fundamental importance in many problems coming from pure and applied mathematics, as well as in developing applications to other sciences.

In this article, we will focus on the concept of statistical stability of a dynamical system, i.e., how its statistical features change when the systems is perturbed or modified. The interest in this question is clearly motivated by the need of controlling how much, and to which extent, approximations, external perturbations and uncertainties can affect the qualitative and quantitative analysis of its dynamics.

Statistical properties of the long-term evolution of a system are reflected, for instance, by the properties of its invariant measures. When the system is perturbed, it is then useful to understand, and be able to predict, how the relevant invariant measures change by the effect of the perturbation, i.e., what is called the response of the system to the perturbation. In particular, it becomes important to get quantitative estimates on their change by effect of the perturbation, as well as understanding the regularity of their behavior, for instance differentiability, Lipschitz or Hölder dependence, etc...

Date: June 30, 2020.

2010 Mathematics Subject Classification. Primary 37C40 ; Secondary 37A30, 37C40, 37J40 , 37A45, 37C05, 37C75, 37E45 .

Key words and phrases. Linear response, statistical stability, rotations, circle diffeomorphisms, KAM theory.

1The concept of relevant is strictly related to the analysis that is carried out. Hereafter, we will be interested in so called physical measures (see footnote or [25]). In other contexts, other kinds of measures might be considered, for example, the so-called measures of maximal entropy.
These ideas can be applied to many kinds of systems and these concepts can be studied in many different ways. In this paper we will consider *discrete deterministic dynamical systems* and *deterministic perturbations*.

More specifically, we will consider systems of the kind \((X, T_0)\), where \(X\) is a compact metric space and \(T_0 : X \to X\) a map, whose iterations determine the dynamics; we investigate perturbed systems \(\{(X, T_\delta)\}_{\delta \in [0, \overline{\delta})}\), where \(T_\delta : X \to X\) are such that \(T_\delta \to T_0\), as \(\delta \to 0\), in some suitable topology.

For each \(\delta \in [0, \overline{\delta})\) let \(\mu_\delta\) be an invariant Borel probability measure for the system \((X, T_\delta)\); we aim to get information on the regularity of this family of measures, by investigating the regularity of the map \(\delta \mapsto \mu_\delta\). This notion of regularity might depend on the topology with which the space of measures is equipped. In this paper we will be interested in absolutely continuous measures with the \(L^1\) norm, as well as in the whole space of Borel probability measures \(P(X)\), endowed with a suitable weak norm, see subsection 2.1 for more details.

We say that \((X, T_0, \mu_0)\) is *statistically stable* (with respect to the considered class of perturbations) if this map is continuous at \(\delta = 0\) (with respect to the chosen topology on the space of measures in which \(\mu_0\) is perturbed). *Quantitative statistical stability* is provided by quantitative estimates on its modulus of continuity.

Differentiability of this map at \(\delta = 0\) is referred to by saying that the system has *linear response* to a certain class of perturbations. Similarly, higher derivatives and higher degrees of smoothness can be considered.

These questions are by now well understood in the case of uniformly hyperbolic systems, where it has been established Lipschitz and, in some cases, differentiable dependence of the relevant (physical) invariant measures with respect to the considered perturbation (see, for example, [10] for a recent survey on linear response under deterministic perturbations, or the introduction in [30] for a survey focused on higher-order terms in the response and for results in the stochastic setting).

For systems having not a uniformly hyperbolic behavior, in presence of discontinuities, or more complicated perturbations, much less is known and results are limited to particular classes of systems; see, for instance, [2], [4], [8], [15], [17], [18], [9], [12], [17], [22], [23], [24], [33], [28], [27], [39], [35], [45], [42], [59] for other results about statistical stability for different classes of systems. We point out a particular kind of deterministic perturbation which will be considered in this paper: the spatial discretization. In this perturbation, one considers a discrete set in the phase space and replaces the map \(T\) with its composition with a projection to this discrete set. This is what happen for example when we simulate the behavior of a system by iterating a map on our computer, which has a finite resolution and each iterate is subjected to numerical truncation. This perturbation change the system into a periodic one, destroying many features of the original dynamics, yet this kinds of simulations are quite reliable in many cases when the resolution is large enough and are widely used in the applied sciences. Why and under which assumptions these simulations are reliable or not is an important mathematical problem, which is still largely unsolved. Few rigorous results have been found so far about the stability under spatial discretization (see e.g. [13], [22], [34], [35], [46]). We refer to Section 4 for a more detailed discussion on the subject.
The majority of results on statistical stability are established for systems that are, in some sense, chaotic. There is indeed a general relation between the speed of convergence to the equilibrium of a system (which reflects the speed of mixing) and the quantitative aspects of its statistical stability (see [27], Theorem 5).

In this paper we consider a class of systems that are not chaotic at all, namely the diffeomorphisms of the circle. We believe that they provide a good model to start pushing forward this analysis. In particular, we will start our discussion by investigating the case of rotations of the circle, and then explaining how to generalize the results to the case of circle diffeomorphisms (see section 4).

We prove the following results.

(1) The statistical stability of irrational rotations under perturbations that are small in the uniform convergence topology. Here stability is proved with respect to a weak norm on the space $P(X)$, related to the so-called Wasserstein distance; see Theorem 2.

(2) Hölder statistical stability for Diophantine rotations under the same kind of perturbations, where the Hölder exponent depends on the Diophantine type of the rotation number. See Theorem 13 for the general upper bounds and Proposition 15 for examples showing these bounds are in some sense sharp.

(3) Differentiable behavior and linear response for Diophantine rotations, under smooth perturbations that preserve the rotation number; for general smooth perturbations the result still holds, but for a Cantor set of parameters (differentiability in the sense of Whitney); see Theorem 28 and Corollary 30.

(4) We extend these qualitative and quantitative stability results to diffeomorphisms of the circle satisfying suitable assumptions; see Theorems 31 and 33.

(5) We prove the statistical stability of diffeomorphisms of the circle under spatial discretizations and numerical truncations, also providing quantitative estimates on the "error" introduced by the discretization.

We believe that the general statistical stability picture here described for rotations is analogous to the one found, in different settings, for example in [11, 12, 13] (see also [10, Section 4]). We have a smooth behavior of the statistical properties of the system for perturbations not changing the topological class of the system (i.e., changing the system to a topologically conjugated one), while we have less regularity, and in particular Hölder behavior, if the perturbation is allowed to change it. In our case, the rotation number plays the role of determining the topological class of the system.

Some comments on the methodology used to establish these results. As far as items 1 and 2 are concerned, we remark that since rotations are not mixing, the general relation between the speed of convergence to the equilibrium and their statistical stability, that we have recalled above, cannot be applied. However, we can perform some analogous construction considering the speed of convergence to the equilibrium of the Cesàro averages of the iterates of a given measure, which leads to a measure of the speed of convergence of the system to its ergodic behavior (see Lemma 3). Quantitative estimates of this speed of the convergence – and hence
our quantitative stability statement, Theorem 13 – are obtained by means of the so-called Denjoy-Koksma inequality (see Theorem 12).

On the other hand, results in item 3 are obtained as an application of KAM theory for circle maps (see Theorem 25), with a particular focus on the dependence of the KAM-construction on the perturbative parameter. In Section 2, we provide a brief introduction on this subject.

The extension of the statistical stability results established for rotations to circle diffeomorphisms (item 4) is done again by combining our results for irrational rotations with the general theory of linearization of circle diffeomorphisms, including Denjoy theorem, KAM theory and Herman-Yoccoz general theory (see section 3.1).

The final application to spatial discretizations is obtained as corollary of these statements, which – thanks to the rather weak assumptions on the perturbations – are suitable to deal with this particularly difficult kind of setting.

As a final remark, although we have decided to present our results in the framework of circle diffeomorphisms and rotations of the circle, we believe that the main ideas present in our constructions can be naturally applied to extend these results to rotations on higher dimensional tori.

**Organization of the article.** In Section 2 after introducing some tools from number theory and geometric measure theory we prove qualitative and quantitative statistical stability of irrational rotations. The quantitative stability results are proved first by establishing general Hölder upper bounds in subsection 2.2 and then exhibiting particular small perturbations for which we actually have Hölder behavior, hence establishing lower bounds in section 2.3.

In Section 3, after a brief introduction to KAM theory and to the problem of smooth linearization of circle diffeomorphisms, we prove linear response results for suitable deterministic perturbations of Diophantine rotations.

In Section 4, we show how to extend the results of Section 2 to sufficiently smooth circle diffeomorphisms.

Finally, in Section 5, we introduce a class of perturbations coming from spatial discretization and apply our previous results to this kind of perturbations, obtaining some qualitative and quantitative results.

**Acknowledgments.** The authors are grateful to A. Celletti, R. de la Llave, P-A Guiheneuf, C. Liverani, M. Sevryuk for their helpful suggestions. The authors also thank R. Calleja, A. Alessandra and R. de la Llave for sharing with them their results in [19].

S.G. and A.S. have been partially supported by the research project PRIN Project 2017S35EHN “Regular and stochastic behavior in dynamical systems” of the Italian Ministry of Education and Research (MIUR). AS also acknowledges the support of the MIUR Department of Excellence grant CUP E83C18000100006.
2. Statistical stability of irrational rotations

Irrational rotations on the circle preserve the Lebesgue measure \( m \) on the circle \( S^1 := \mathbb{R}/\mathbb{Z} \) and are well known for being uniquely ergodic. It is easy to see that small perturbations of such rotations may have singular invariant measures (i.e., not absolutely continuous with respect to \( m \)), even supported on a discrete set (see examples in Section 2.3). However, we will show that these measures must be close, in some suitable sense, to \( m \).

2.1. Weak statistical stability of irrational rotations. In this section, we aim to prove a statistical stability result for irrational rotations in a weak sense; more specifically, we show that by effect of small natural perturbations, their invariant measures vary continuously with respect to the so-called Wasserstein distance.

Let us first recall some useful notions that we are going to use in the following. Let \((X,d)\) be a compact metric space and let \( \mathcal{M}(X) \) denote the set of signed finite Borel measures on \( X \). If \( g : X \to \mathbb{R} \) is a Lipschitz function, we denote its (best) Lipschitz constant by \( \text{Lip}(g) \), i.e.,

\[
\text{Lip}(g) := \sup_{x,y \in X, x \neq y} \left\{ \frac{|g(x) - g(y)|}{d(x,y)} \right\}.
\]

**Definition 1.** Given \( \mu, \nu \in \mathcal{M}(X) \) we define the Wasserstein-Monge-Kantorovich distance between \( \mu \) and \( \nu \) by

\[
W(\mu, \nu) := \sup_{\text{Lip}(g) \leq 1, M_\mu, M_\nu \leq 1} \left| \int_{S^1} gd\mu - \int_{S^1} gd\nu \right|.
\]

We denote \( \| \cdot \|_W := W(0, \mu) \), where 0 denotes the trivial measure identically equal to zero. \( \| \cdot \|_W \) defines a norm on the vector space of signed measures defined on a compact metric space.

We refer the reader, for example, to [1] for a more systematic and detailed description of these topics.

Let \( T : X \to X \) be a Borel measurable map. Define the linear functional

\[
L_T : \mathcal{M}(X) \to \mathcal{M}(X)
\]

that to a measure \( \mu \in \mathcal{M}(X) \) associates the new measure \( L_T \mu \), satisfying \( L_T \mu(A) := \mu(T^{-1}(A)) \) for every Borel set \( A \subset X \); \( L_T \) will be called transfer operator (observe that \( L_T \mu \) is also called the push-forward of \( \mu \) by \( T \) and denoted by \( T_\ast \mu \)). It follows easily from the definition, that invariant measures correspond to fixed points of \( L_T \), i.e., \( L_T \mu = \mu \).

We are now ready to state our first statistical stability result for irrational rotations.

**Theorem 2** (Weak statistical stability of irrational rotations.). Let \( R_\delta : S^1 \to S^1 \) be an irrational rotation. Let \( \{T_\delta\}_{0 \leq \delta \leq \delta} \) be a family of Borel measurable maps of \( S^1 \) to itself such that

\[
\sup_{x \in S^1} |R_\delta(x) - T_\delta(x)| \leq \delta.
\]
Suppose $\mu_\delta$ is an invariant measure\footnote{In the case when $T_\delta$ is continuous such measures must exist by the Krylov-Bogoliubov theorem \cite{40}. In other cases such measures can be absent, in this case our statement is empty.} of $T_\delta$. Then
\[
\lim_{\delta \to 0} \| m - \mu_\delta \|_W = 0.
\]

Let us start with the following preliminary computation.

**Lemma 3.** Let $L$ be the transfer operator associated to an isometry of $S^1$ and let $L_\delta$ be the transfer operator associated to a measurable map $T_\delta$. Suppose that $\mu_\delta = L_\delta \mu_\delta$. Then, for each $n \geq 1$
\[
\| \mu_\delta - m \|_W \leq \| m - \frac{1}{n} \sum_{1 \leq i \leq n} L^i \mu_\delta \|_W + \frac{(n - 1)}{2} \| (L - L_\delta) \mu_\delta \|_W
\]
where $L^i := L \circ \ldots \circ L$ (i-times).

**Proof.** The proof is a direct computation. Since $\mu_\delta = L_\delta \mu_\delta$ and $m$ is invariant for $L$, then
\[
\| \mu_\delta - m \|_W \leq \| m - \frac{1}{n} \sum_{1 \leq i \leq n} L^i \mu_\delta \|_W + \frac{(n - 1)}{2} \| (L - L_\delta) \mu_\delta \|_W
\]
(3)

Since
\[
L^i - L^i_\delta = \sum_{k=1}^{i} L^{i-k}(L - L_\delta)L_\delta^{k-1}
\]
then
\[
(L^i - L^i_\delta) \mu_\delta = \sum_{k=1}^{i} L^{i-k}(L - L_\delta)L_\delta^{k-1} \mu_\delta = \sum_{k=1}^{i} L^{i-k}((L - L_\delta)\mu_\delta).
\]

Being $L$ is the transfer operator associated to an isometry, then
\[
\| L^{i-k}(L - L_\delta)\mu_\delta \|_W = \|(L - L_\delta)\mu_\delta \|_W
\]
and consequently
\[
\| (L^i - L^i_\delta) \mu_\delta \|_W \leq (i - 1) \| (L - L_\delta) \mu_\delta \|_W.
\]

Substituting in (3), we conclude
\[
\| \mu_\delta - m \|_W \leq \| m - \frac{1}{n} \sum_{1 \leq i \leq n} L^i (m - \mu_\delta) \|_W + \frac{(n - 1)}{2} \| (L - L_\delta) \mu_\delta \|_W.
\]
\[\square\]
Lemma 4. Under the assumptions of Theorem 2 let \( \{\mu_\delta\}_{0 \leq \delta \leq \delta'} \) be a family of Borel measures on \( S^1 \), then

\[
\lim_{n \to \infty} \| m - \frac{1}{n} \sum_{1 \leq i \leq n} L_i \mu_\delta \|_W = 0
\]

uniformly in \( \delta \); namely, for every \( \varepsilon > 0 \) there exists \( \overline{\delta} = \overline{\delta}(\varepsilon) \) such that if \( n \geq \overline{\delta} \) then

\[
\sup_{0 \leq \delta \leq \delta'} \| m - \frac{1}{n} \sum_{1 \leq i \leq n} L_i \mu_\delta \|_W \leq \varepsilon.
\]

Proof. Let \( \delta_{x_0} \) be the delta-measure concentrated at a point \( x_0 \in S^1 \). By unique ergodicity of the system, we get

\[
\lim_{n \to \infty} \| m - \frac{1}{n} \sum_{1 \leq i \leq n} L_i \delta_{x_0} \|_W = 0.
\]

This is uniform in \( x_0 \); in fact, changing \( x_0 \) is equivalent to compose by a further rotation, which is an isometry and hence does not change the \( \| \cdot \|_W \) norm. Any measure \( \mu_\delta \) can be approximated in the \( \| \cdot \|_W \) norm, with arbitrary precision, by a convex combination of delta-measures, i.e., for each \( \varepsilon > 0 \) there are \( x_1, \ldots, x_k \in S^1 \) and \( \lambda_1, \ldots, \lambda_k \geq 0 \), with \( \sum_{i \leq k} \lambda_i = 1 \) such that

\[
\| \mu_\delta - \sum_{1 \leq i \leq k} \lambda_i x_i \|_W \leq \varepsilon.
\]

Since \( R_\alpha \) is an isometry the \( \| \cdot \|_W \) norm is preserved by the iterates of \( L \). Hence for each \( n \geq 0 \), we also have

\[
\| L^n \mu_\delta - L^n \left( \sum_{1 \leq i \leq m} \lambda_i \delta_{x_i} \right) \|_W \leq \varepsilon.
\]

Hence, for any \( n \) we have

\[
\| m - \frac{1}{n} \sum_{1 \leq j \leq n} L^j \left( \sum_{i \leq k} \lambda_i \delta_{x_i} \right) \|_W = \| \sum_{1 \leq i \leq k} \lambda_i m - \sum_{1 \leq i \leq k} \lambda_i \left( \sum_{1 \leq j \leq n} L^j \delta_{x_i} \right) \|_W
\]

and therefore \( \lim_{n \to \infty} \| \sum_{i \leq k} \lambda_i \left( m - \frac{1}{n} \sum_{j \leq n} L^j \delta_{x_i} \right) \|_W = 0 \). From this, the claim of the theorem easily follows. \( \square \)

We can now prove Theorem 2.

Proof of Theorem 2. Let \( L_\delta \) be the transfer operator associated to \( T_\delta \). By Lemma 4 \( \lim_{n \to \infty} \| m - \frac{1}{n} \sum_{1 \leq i \leq n} L^i \mu_\delta \|_W = 0 \) uniformly in \( \delta \). Since

\[
\sup_{x \in S^1} | R_\alpha(x) - T_\delta(x) | \leq \delta,
\]

then \( \| (L - L_\delta) \mu_\delta \|_W \leq \delta \) and

\[
\lim_{\delta \to 0} \| (L - L_\delta) \mu_\delta \|_W = 0.
\]

By Lemma 3 we get that for each \( n \)

\[
\| \mu_\delta - m \|_W \leq \| m - \frac{1}{n} \sum_{1 \leq i \leq n} L^i \mu_\delta \|_W + \frac{(n-1)}{2} \| (L - L_\delta) \mu_\delta \|_W.
\]
It follows from Lemma 4 that we can choose \( n \) such that \( \|m - \frac{1}{n} \sum_{1 \leq i \leq n} L_i \mu_\delta \|_W \) is as small as wanted. Then, using (4), we can choose \( \delta \) sufficiently small so to make \( \frac{(n-1)}{2} \| (L - L_\delta) \mu_\delta \|_W \) as small as needed, hence proving the statement. \( \square \)

2.2. Quantitative statistical stability of Diophantine rotations, upper bounds.

We now consider irrational rotations, for rotation numbers that are “badly” approximable by rationals: the so-called Diophantine numbers. In this case, we can provide a quantitative estimate for the statistical stability of the system by showing that the modulus of continuity of the function \( \delta \mapsto \mu_\delta \) is Hölderian, and that its exponent depends on the Diophantine type of the rotation number.

Let us start by recalling the definition of Diophantine type for a real number (see [41]): this concept expresses quantitatively the rate of approximability of an irrational number by sequences of rationals.

Definition 5. If \( \alpha \) is irrational, the Diophantine type of \( \alpha \) is defined by

\[
\gamma(\alpha) := \sup \{ \gamma \geq 0 : \liminf_{k \to \infty} k^\gamma \| k\alpha \|_Z = 0 \}.
\]

We remark that in some cases \( \gamma(\alpha) = +\infty \). When \( \gamma(\alpha) < +\infty \) we say \( \alpha \) is of finite Diophantine type.

Remark 6. The Diophantine type of \( \alpha \) can be also defined by

\[
\gamma(\alpha) := \inf \left\{ \gamma \geq 0 : \exists c > 0 \text{ s.t. } \| k\alpha \|_Z \geq c \cdot |k|^{-\gamma} \forall k \in \mathbb{Z} \setminus \{0\} \right\}.
\]

In the light of this last remark on the Diophantine type of a number, we recall the definition of Diophantine number as it very commonly stated in the literature.

Definition 7. Given \( c > 0 \) and \( \tau \geq 0 \), we say that a number \( \alpha \in (0, 1) \) is \( (c, \tau) \)-Diophantine if

\[
|\alpha - \frac{p}{q}| > \frac{c}{|q|^{1+\tau}} \quad \forall \quad \frac{p}{q} \in \mathbb{Q} \setminus \{0\}.
\]

We denote by \( D(c, \tau) \) the set of of \( (c, \tau) \)-Diophantine numbers and by \( D(\tau) := \cup_{c>0} D(c, \tau) \).

Remark 8. Comparing with Definition 5, it follows that every \( \alpha \in D(\tau) \) has finite Diophantine type \( \gamma(\alpha) \leq \tau \). On the other hand, if \( \alpha \) has finite Diophantine type, then \( \alpha \in D(\tau) \) for every \( \tau > \gamma(\alpha) \).

Remark 9. Let us point out the following properties (see [51, p. 601] for their proofs):

- if \( \tau < 1 \), the set \( D(\tau) \) is empty;
- if \( \tau > 1 \) the set \( D(\tau) \) has full Lebesgue measure;
- if \( \tau = 1 \), then \( D(\tau) \) has Lebesgue measure equal to zero, but it has Hausdorff dimension equal to 1 (hence, it has the cardinality of the continuum).
See also [37, Section V.6] for more properties.

Now we introduce the notion of discrepancy of a sequence $x_1, ..., x_N \in [0, 1]$. This is a measure of the equidistribution of the points $x_1, ..., x_N$. Given $x_1, ..., x_N \in [0, 1]$ we define the discrepancy of the sequence by

$$D_N(x_1, ..., x_N) := \sup_{\alpha \leq \beta, \alpha, \beta \in [0, 1]} \left| \frac{1}{N} \sum_{1 \leq i \leq N} 1_{[\alpha, \beta]}(x_i) - (\beta - \alpha) \right|$$

it can be proved (see [41, Theorem 3.2, page 123]) that the discrepancy of sequences obtained from orbits of and irrational rotation is related to the Diophantine type of the rotation number.

**Theorem 10.** Let $\alpha$ be an irrational of finite Diophantine type. Let us denote by $D_{N, \alpha}(0)$ the discrepancy of the sequence $\{x_i\}_{0 \leq i \leq N} = \{\alpha i - \lfloor \alpha i \rfloor\}_{0 \leq i \leq N}$ (where $\lfloor \cdot \rfloor$ stands for the integer part). Then:

$$D_{N, \alpha}(0) = O(N^{-\frac{1}{\gamma(\alpha)} + \varepsilon}) \quad \forall \varepsilon > 0.$$

From the definition of discrepancy, Theorem 10, and the fact that the translation is an isometry, we can deduce the following corollary.

**Corollary 11.** Let $x_0 \in S^1$, let us denote by $D_{N, \alpha}(x_0)$ the discrepancy of the sequence $\{x_i\}_{1 \leq i \leq N} = \{x_0 + \alpha i - \lfloor x_0 + \alpha i \rfloor\}_{0 \leq i \leq N}$. Then Theorem 10 holds uniformly for each $x_0$, namely for every $\varepsilon > 0$ there exists $C = C(\varepsilon) \geq 0$ such that for each $x_0$ and $N \geq 1$

$$D_{N, \alpha}(x_0) \leq CN^{-\frac{1}{\gamma(\alpha)} + \varepsilon}.$$

The discrepancy is also related to the speed of convergence of Birkhoff sums of irrational rotations. The following is known as the Denjoy-Kocsma inequality (see [41, Theorem 5.1, page 143 and Theorem 1.3, page 91]).

**Theorem 12.** Let $f$ be a function of bounded variation, that we denote by $V(f)$. Let $x_1, ..., x_N \in [0, 1]$ be a sequence with discrepancy $D_N(x_1, ..., x_N)$. Then

$$\left| \frac{1}{N} \sum_{1 \leq i \leq N} f(x_i) - \int_{[0, 1]} f \, dx \right| \leq V(f) D_N(x_1, ..., x_N).$$

We can now prove a quantitative version of our stability result.

**Theorem 13** (Quantitative statistical stability of Diophantine rotations). Let $R_\alpha : S^1 \to S^1$ be an irrational rotation. Suppose $\alpha$ has finite Diophantine type $\gamma(\alpha)$. Let $\{T_\delta\}_{0 \leq \delta \leq \varepsilon}$ be a family of Borel measurable maps of the circle such that

$$\sup_{x \in S^1} |R_\alpha(x) - T_\delta(x)| \leq \delta.$$

Suppose $\mu_\delta$ is an invariant measure of $T_\delta$. Then, for each $\ell < \frac{1}{\gamma(\alpha) + 1}$ we have:

$$||m - \mu_\delta||_W = O(\delta^\ell).$$
Let us first prove some preliminary result.

**Lemma 14.** Under the assumptions of Theorem 13, let \( \{\mu_{\delta}\}_{0 \leq \delta \leq \bar{\delta}} \) be a family of Borel probability measures on \( S^1 \). Then, for every \( \epsilon > 0 \)
\[
\left\| m - \frac{1}{n} \sum_{1 \leq i \leq n} L_i \mu_{\delta} \right\|_W = O(n^{-\frac{1}{\gamma(\alpha)} + \epsilon})
\]
uniformly in \( \delta \); namely, for every \( \epsilon > 0 \), there exist \( C = C(\epsilon) \geq 0 \) such that for each \( \delta \) and \( n \geq 1 \)
\[
\left\| m - \frac{1}{n} \sum_{1 \leq i \leq n} L_i \mu_{\delta} \right\|_W \leq C n^{-\frac{1}{\gamma(\alpha)} + \epsilon}.
\]

**Proof.** Let us fix \( \epsilon > 0 \). By Theorem 12 and Corollary 11 we have that there is \( C \geq 0 \) such that for each Lipschitz function \( f \) with Lipschitz constant 1, and for each \( x_0 \in S^1 \) we have
\[
\left| \frac{1}{n} \sum_{1 \leq i \leq n} f(R_{\alpha}^i(x_0)) - \int_{[0,1]} f \,dx \right| \leq C n^{-\frac{1}{\gamma(\alpha)} + \epsilon} \quad \forall \, n \geq 1.
\]
Let \( \delta_{x_0} \) be the delta-measure concentrated at a point \( x_0 \in S^1 \). By definition of \( \| \cdot \|_W \), we conclude that
\[
\left\| m - \frac{1}{n} \sum_{1 \leq i \leq n} L_i \delta_{x_0} \right\|_W \leq C n^{-\frac{1}{\gamma(\alpha)} + \epsilon}.
\]
Now, as in the proof of Lemma 3 any measure \( \mu_{\delta} \) can be approximated, arbitrary well, in the \( \| \cdot \|_W \) norm by a convex combination of delta-measures and we obtain (7) from (8), exactly in the same way as done in the proof of Lemma 3.}

**Proof of Theorem 13.** Let \( L_{\delta} \) be the transfer operator of \( T_{\delta} \). Let us fix \( \epsilon > 0 \); without loss of generality we can suppose \( \epsilon < \frac{1}{\gamma(\alpha)} \). By Lemma 14 we have that
\[
\left\| m - \frac{1}{n} \sum_{1 \leq i \leq n} L_i \mu_{\delta} \right\|_W \leq C n^{-\frac{1}{\gamma(\alpha)} + \epsilon}.
\]
By Lemma 3 we get that for each \( n \geq 1 \)
\[
\| \mu_{\delta} - m \|_W \leq \left\| m - \frac{1}{n} \sum_{1 \leq i \leq n} L_i \mu_{\delta} \right\|_W + \frac{(n-1)}{2} \| (L - L_{\delta}) \mu_{\delta} \|_W.
\]
Hence
\[
\| \mu_{\delta} - m \|_W \leq C n^{-\frac{1}{\gamma(\alpha)} + \epsilon} + \frac{(n-1)}{2} \| (L - L_{\delta}) \mu_{\delta} \|_W \leq C n^{-\frac{1}{\gamma(\alpha)} + \epsilon} + \frac{(n-1)}{2} \delta,
\]
where we have used that, since \( \sup_{x \in S^1} |R_{\alpha}(x) - T_{\delta}(x)| \leq \delta \), then
\[
\| (L - L_{\delta}) \mu_{\delta} \|_W \leq \delta.
\]
Proposition 15. Let us consider the rotation \( R_a \) with rotation number \( a \) of Diophantine type \( 1 < \gamma(a) \leq +\infty \), there exist perturbations of “size \( \delta \)”, for which the unique invariant measure varies in a Hölder way.

In this subsection we discuss that the upper bound on the statistical stability obtained in Theorem 13 is essentially optimal. We show that for a rotation \( R_a \) with rotation number \( \alpha \) of Diophantine type \( 1 < \gamma(\alpha) \leq +\infty \), there exist perturbations of “size \( \delta \)”, for which the unique invariant measure varies in a Hölder way.

More specifically, for any \( r \geq 0 \) we will construct a sequence \( \delta_n \to 0 \) and \( C^\infty \)-maps \( T_n \) such that: \( \| R_\alpha - T_n \|_{C^r} \leq \delta_n \), \( T_n \) has a unique invariant probability measure \( \mu_n \) and \( \| \mu_n - m \|_W \geq C \delta_n^p \) for some \( C \geq 0 \) and \( p > 1 \).

2.3. Quantitative statistical stability of Diophantine rotations, lower bounds.

In this subsection we discuss that the upper bound on the statistical stability obtained in Theorem 13 is essentially optimal. We show that for a rotation \( R_\alpha \) with rotation number \( \alpha \) of Diophantine type \( 1 < \gamma(\alpha) \leq +\infty \), there exist perturbations of “size \( \delta \)”, for which the unique invariant measure varies in a Hölder way.

More specifically, for any \( r \geq 0 \) we will construct a sequence \( \delta_n \to 0 \) and \( C^\infty \)-maps \( T_n \) such that: \( \| R_\alpha - T_n \|_{C^r} \leq \delta_n \), \( T_n \) has a unique invariant probability measure \( \mu_n \) and \( \| \mu_n - m \|_W \geq C \delta_n^p \) for some \( C \geq 0 \) and \( p > 1 \).

Proposition 15. Let us consider the rotation \( R_\alpha : S^1 \to S^1 \), where \( \alpha \) is an irrational number with \( 1 < \gamma(\alpha) \leq +\infty \). For each \( r \geq 0 \) and \( \gamma' < \gamma(\alpha) \) there exist a sequence of numbers \( \delta_j > 0 \) and \( C^\infty \)-diffeomorphisms \( T_j : S^1 \to S^1 \) such that \( \| T_j - R_\alpha \|_{C^r} \leq 2 \delta_j \) and

\[
\| m - \mu_j \|_W \geq \frac{1}{2} \delta_j^{1+r}.
\]

for every \( j \in \mathbb{N} \) and for every \( \mu_j \) invariant measure of \( T_j \).

Proof. We remark the unique invariant measure for \( R_\alpha \) is the Lebesgue measure \( m \). Let us choose \( \gamma' < \gamma(\alpha) \); it follows from the definition of \( \gamma(\alpha) \) that there are infinitely many integers \( k_j \in \mathbb{N} \) and \( p_j \in \mathbb{Z} \) such that

\[
|k_j \alpha - p_j| \leq \frac{1}{k_j^{\gamma'}} \quad \iff \quad |\alpha - \frac{p_j}{k_j}| \leq \frac{1}{k_j^{\gamma'+r}}.
\]

Let us set \( \delta_j := -\alpha + \frac{p_j}{k_j} \). Clearly, \( |\delta_j| \leq \frac{1}{k_j^{\gamma'+r}} \to 0 \) as \( j \to \infty \).

Consider \( T_j \) defined as \( T_j(x) = R_{\alpha + \delta_j}(x) \); for each \( r \geq 0 \) we have that \( \| T_j - R_\alpha \|_{C^r} \approx 2|\delta_j| \). Since \( \delta_j + \alpha = \frac{p_j}{k_j} \) is rational, every orbit is \( k_j \)-periodic. Let us consider the orbit starting at 0 and denote it by \( y_0 := 0, y_1 := \delta_j, \ldots, y_{k_j-1} := 1 - \delta_j, y_{k_j} := 0 \pmod{\mathbb{Z}} \).
Consider the measures
\[ \mu_j = \frac{1}{k_j} \sum_{0 \leq i < k_j} \delta_{y_i}, \]
where \( \delta_{y_i} \) is the delta-measure concentrated at \( y_i \). The measure \( \mu_j \) is clearly invariant for the map \( T_j \) and it can be directly computed that
\[ \| m - \mu_j \|_W \geq \frac{1}{2k_j}. \]

Observe that \( |\delta_j| \leq \frac{1}{k_j^{1+r}} \), hence we get \( |\delta_j|^{-\frac{1}{1+r}} \leq \frac{1}{k_j} \); then
\[ \| m - \mu_j \|_W \geq \frac{1}{2}|\delta_j|^{-\frac{1}{1+r}}. \]

This example can be further improved by perturbing the map \( \tilde{T}_j = R_{\alpha+\delta_j} \) to a new map \( T_j \) in a way that the measure \( \mu_j \) (supported on the attractor of \( T_j \)) and the measure \( \mu_j + \frac{k_j}{2} \) (supported on the repeller of \( T_j \)) are the only invariant measures of \( T_j \), and \( \mu_j \) is the unique physical measure for the system. This can be done by making a \( C^\infty \) perturbation on \( \tilde{T}_j = R_{\alpha+\delta_j} \), as small as wanted in the \( C^r \)-norm. In fact, let us denote, as before, by \( \{y_k\}_k \) the periodic orbit of 0 for \( R_{\alpha+\delta_j} \). Let us consider a \( C^\infty \) function \( g : [0, 1] \to [0, 1] \) such that:

- \( g \) is negative on the each interval \([y_i, y_i + \frac{1}{2k_j}]\) and positive on each interval \([y_i + \frac{1}{2k_j}, y_{i+1}]\) (so that \( g(y_i + \frac{1}{2k_j}) = 0 \));
- \( g' \) is positive in each interval \([y_i + \frac{1}{3k_j}, y_{i+1} - \frac{1}{3k_j}]\) and negative in \([y_i, y_{i+1}]\) - \([y_i + \frac{1}{3k_j}, y_{i+1} - \frac{1}{3k_j}]\).

Considering \( D_\delta : S^1 \to S^1 \), defined by \( D_\delta(x) := x + \delta g(x) \) (mod. \( \mathbb{Z} \)), it holds that the iterates of this map send all the space, with the exception of the set \( \Gamma_{\text{rep}} := \{y_i + \frac{1}{3k_j} : 0 \leq i < k_j\} \) (which is a repeller), to the set \( \Gamma_{\text{att}} := \{y_i : 0 \leq i < k_j\} \) (the attractor). Then, define \( T_j \) by composing \( R_{\alpha+\delta_j} \) and \( D_\delta \), namely
\[ T_j(x) := D_\delta(x + (\delta_j + \alpha)). \]

The claim follows by observing that for the map \( T_j(x) \), both sets \( \Gamma_{\text{att}} \) and \( \Gamma_{\text{rep}} \) are invariant and, in particular, the whole space \( S^1 - \Gamma_{\text{rep}} \) is attracted by \( \Gamma_{\text{att}} \). \( \square \)

The construction done in the previous proof can be extended to show Hölder behavior for the average of a given fixed regular observable. We show an explicit example of such an observable, with a particular choice of rotation number \( \alpha \).

**Proposition 16.** Consider a rotation \( R_\alpha \) with rotation angle \( \alpha := \sum_{i=1}^{\infty} 2^{-2^i} \). Let \( T_j \) be its perturbations as constructed in Proposition [7] and let \( \mu_j \) denote their invariant measures; recall that \( \| T_j - R_\alpha \|_{C^\ell} \leq 2|\delta_j| = 2 \sum_{i=1}^{\infty} 2^{-2^i} \).

Then, there is an observable \( \psi : S^1 \to \mathbb{R} \), with derivative in \( L^2(S^1) \), and \( C \geq 0 \) such that
\[ \left| \int_{S^1} \psi dm - \int_{S^1} \psi d\mu_j \right| \geq C \sqrt{|\delta_j|}. \]

3The translated measure is defined as follows: \( |\mu_j + \frac{1}{2k_j}|(A) := \mu_j(A - \frac{1}{2k_j}) \) for each measurable set \( A \) in \( S^1 \), where \( A - \frac{1}{2k_j} \) is the translation of the set \( A \) by \( -\frac{1}{2k_j} \).
Proof. Comparing the series with a geometric one, we get that
\[ \sum_{n=1}^{\infty} 2^{-2^{2n}} \leq 2^{-2^{2(n+1)+1}}. \]
By this, it follows
\[ \|2^{2^{2n}} \theta\| \leq 2^{-2^{2(n+1)+1}} = \frac{1}{2(2^{2^{2n+2}})} = \frac{1}{2(2^{2^{2n}})}^2. \]
Since it also holds that \( \|2^{2^{2n}} \theta\| \geq 2^{-2^{2(n+1)+1}} \), we conclude that \( \gamma(\alpha) = 4 \).
Following the construction in the proof of Proposition 15, we have that with a perturbation of size less than
\[ \frac{1}{2(2^{2^{2}+2^{2}})2^{2}} \]
the angles \( \alpha_j := \alpha - \delta_j = \sum_{i=1}^{j} 2^{-2^{2i}} \) generate orbits of period \( 2^{2^{2}} \).
Now let us construct a suitable observable which can “see” the change of the invariant measure under this perturbation. Let us consider
\[ (11) \psi(x) := \sum_{i=1}^{\infty} \frac{1}{(2^{2^{2}})^2} \cos(2^{2^{2}} 2\pi x) \]
and denote by \( \psi_k(x) := \sum_{i=1}^{k} \frac{1}{(2^{2^{2}})^2} \cos(2^{2^{2}} 2\pi x) \) its truncations. Since for the observable \( \psi \), the \( i \)-th Fourier coefficient decreases like \( i^{-2} \), then \( \psi \) has derivative in \( L^2(S^1) \). Let \( \{x_i\} \) be the periodic orbit of 0 for the map \( R_{\alpha_j} \) and let \( \mu_j := \frac{1}{2^{2^{2}}} \sum_{i=0}^{\alpha_j} \delta_{x_i} \) be the physical measure supported on it. Since \( 2^{2^{2}} \) divides \( 2^{2^{2}}(j+1) \) then \( \sum_{i=1}^{2^{2^{2}}} \psi_k(x_i) = 0 \) for every \( k < j \), thus \( \int_{S^1} \psi_{j-1} \ d\mu_j = 0 \). Then
\[ v_j := \int_{S^1} \psi \ d\mu_j \geq \frac{1}{(2^{2^{2}})^2} - \sum_{j+1}^{\infty} \frac{1}{(2^{2^{2}})^2} \]
\[ \geq 2^{-2^{2j+1}} - 2^{-2^{2(j+1)+1}}. \]
For \( j \) big enough
\[ 2^{-2^{2j+1}} - 2^{-2^{2(j+1)+1}} \geq \frac{1}{2}(2^{-2^{2}})^2. \]
Summarizing, with a perturbation of size
\[ \delta_j = \sum_{j=1}^{\infty} 2^{-2^{2i}} \leq 2 \cdot 2^{-2^{2(j+1)}} = 2^{-2^{2(j+1)}} = 2(2^{-2^{2}})^4 \]
we get a change of average for the observable \( \psi \) from \( \int_{S^1} \psi \ dm = 0 \) to \( v_n \geq \frac{1}{2}(2^{-2^{2}})^2 \).
Therefore, there is \( C \geq 0 \) such that with a perturbation of size \( \delta_j \), we get a change of average for the observable \( \psi \) of size bigger than \( C \sqrt{\delta_j} \).
\[ \square \]

Remark 17. Using in (11) \( \frac{1}{(2^{2^{2}})^2} \) instead of \( \frac{1}{(2^{2^{2}})^2} \), we can obtain a smoother observable. Using rotation angles with bigger and bigger Diophantine type, it is possible to obtain a dependence of the physical measure on the perturbation with worse and worse Hölder exponent. Using angles with infinite Diophantine type it is possible to have a behavior whose modulus of continuity is worse than the Hölder one.
3. Linear response and KAM theory

In this section, we would like to discuss differentiable behavior and linear response for Diophantine rotations, under suitable smooth perturbations. In particular, we will obtain our results by means of the so-called KAM theory.

Let us first start by explaining more precisely, what linear response means. Let \((T_\delta)_{\delta \geq 0}\) be a one parameter family of maps obtained by perturbing an initial map \(T_0\). We will be interested on how the perturbation made on \(T_0\) affects some invariant measure of \(T_0\) of particular interest. For example its physical measure. Suppose hence \(T_0\) has a physical measure \(\mu_0\) and let \(\mu_\delta\) be physical measures of \(T_\delta\).

The linear response of the invariant measure of \(T_0\) under a given perturbation is defined, if it exists, by the limit

\[
\dot{\mu} := \lim_{\delta \to 0} \frac{\mu_\delta - \mu_0}{\delta}
\]

where the meaning of this convergence can vary from system to system. In some systems and for a given perturbation, one may get \(L^1\)-convergence for this limit; in other systems or for other perturbations one may get convergence in weaker or stronger topologies. The linear response to the perturbation hence represents the first order term of the response of a system to a perturbation and when it holds, a linear response formula can be written as:

\[
\mu_\delta = \mu_0 + \dot{\mu} \delta + o(\delta)
\]

which holds in some weaker or stronger sense.

We remark that given an observable function \(c : X \to \mathbb{R}\), if the convergence in \((12)\) is strong enough with respect to the regularity\(^4\) of \(c\), we get

\[
\lim_{t \to 0} \frac{\int_{S^1} c \, d\mu_\delta - \int_{S^1} c \, d\mu_0}{t} = \int_{S^1} c \, d\dot{\mu}
\]

showing how the linear response of the invariant measure controls the behavior of observable averages.

3.1. Conjugacy theory for circle maps. Let us recall some classical results on smooth linearization of circle diffeomorphisms and introduce KAM theory.

Let \(\text{Diff}_r(S^1)\) denote the set of orientation preserving homeomorphism of the circle of class \(C^r\) with \(r \in \mathbb{N} \cup \{+\infty, \omega\}\). Let \(\text{rot}(f) \in S^1\) denote the rotation number of \(f\) (see, for example, [77] Section II.2) for more properties on the rotation number).

\(^4\) An invariant measure \(\mu\) is said to be physical if there is a positive Lesbegue measure set \(B\) such that for each continuous observable \(f\)

\[
\int_{S^1} f \, d\mu = \lim_{n \to \infty} \frac{f(x) + f(T(x)) + ... + f(T^n(x))}{n + 1}
\]

for each \(x \in B\) (see [52]).

\(^5\) For example, \(L^1\) convergence in \((12)\) allows to control the behavior of \(L^\infty\) observables in \((14)\), while a weaker convergence in \((12)\), for example in the Wasserstein norm (see definition\(^1\)) allows to get information on the behavior of Lipschitz observable.
A natural question is to understand when a circle diffeomorphism is conjugated to a rotation with the same rotation number, namely whether there exists a homeomorphism $h : \mathbb{S}^1 \to \mathbb{S}^1$ such that the following diagram commutes:

$$
\begin{array}{ccc}
\mathbb{S}^1 & \xrightarrow{f} & \mathbb{S}^1 \\
\uparrow h & & \uparrow h \\
\mathbb{S}^1 & \xrightarrow{\text{rot}(f)} & \mathbb{S}^1
\end{array}
$$

i.e., $h^{-1} \circ f \circ h = \text{R}_{\text{rot}(f)}$. Moreover, whenever this conjugacy exists, one would like to understand what is the best regularity that one could expect.

**Remark 18.** Observe that if $h$ exists, then it is essentially unique, in the sense that if $h_i : \mathbb{S}^1 \to \mathbb{S}^1$, $i = 1, 2$, are homeomorphisms conjugating $f$ to $\text{R}_{\text{rot}(f)}$, then $h_1 \circ h_2^{-1}$ must be a rotation itself: $h_1 \circ h_2^{-1} = R_\beta$ for some $\beta \in \mathbb{S}^1$ (see [37, Ch. II, Proposition 3.3.2]).

This question has attracted a lot of attention, dating back, at least, to Henri Poincaré.

Let us start by recalling the following result due to Denjoy [21] shows that diffeomorphisms with irrational rotation number and satisfying some extra mild regularity assumption (for example, $C^2$ diffeomorphisms do satisfy it) are conjugated to irrational rotations by an homeomorphism.

**Theorem 19** (Denjoy). Let $T$ be an orientation preserving diffeomorphism of the circle with an irrational rotation number $\alpha$ and such that $\log(T')$ has bounded variation. Then there exists a homeomorphism $h : \mathbb{S}^1 \to \mathbb{S}^1$ such that

$$T \circ h = h \circ \text{R}_\alpha.$$

**Remark 20.** Denjoy constructed diffeomorphisms $T$ only of class $C^1$ that are not conjugated to rotations (i.e., such that the support of their invariant measure $\mu$ is not the whole $\mathbb{S}^1$). These are usually called in the literature Denjoy-type diffeomorphisms.

Some of the first contributions about smooth linearization (i.e., obtaining a conjugacy of higher regularity) were due to V.I. Arnol’d [9] and J. Moser [47]. These results are in the perturbative setting and are generally referred to as KAM theory. Namely, they consider perturbations of Diophantine rotations

$$f_\varepsilon(x) = R_\alpha + \varepsilon u(x, \varepsilon)$$

and prove that, under suitable regularity assumptions on $u$, there exists $\varepsilon_0 > 0$ (depending on the properties of $\alpha$ and $u$) such that if $|\varepsilon| < \varepsilon_0$, then $f_\varepsilon$ is conjugated to a $\text{R}_{\text{rot}(f_\varepsilon)}$. See below for a more precise statement.

**Remark 21.** Observe that $f_\varepsilon$ has not necessarily rotation number $\alpha$, even if one asks that $u(\cdot, \varepsilon)$ has zero average.

**Remark 22.** In the analytic setting, KAM theorem for circle diffeomorphisms was firstly proved by Arnol’d (see [5, Corollary to Theorem 3, p. 173]), showing that the conjugation is analytic. In the smooth case, it was proved by Moser [47] under the assumption that $u$ is sufficiently smooth (the minimal regularity needed was later improved by Rüssmann [50]). The literature on KAM theory and its recent developments is huge and we do not aim to provide an accurate account here; for
reader’s sake, we limit ourselves to mentioning some recent articles and surveys, like [17, 20, 25, 43, 44, 54] and references therein.

Later, Herman [37] and Yoccoz [56, 57] provided a thorough analysis of the situation in the general (non-perturbative) context. Let us briefly summarize their results (see also [26] for a more complete account).

**Theorem 23** (Herman [37], Yoccoz [56, 57]).

- Let $f \in \text{Diff}_r^+(S^1)$ and $\text{rot}(f) \in D(\tau)$. If $r > \max\{3, 2\tau - 1\}$, then there exists $h \in \text{Diff}_{r-\tau-\varepsilon}^+(S^1)$, for every $\varepsilon > 0$, conjugating $f$ to $R_{\text{rot}(f)}$.
- Let $f \in \text{Diff}_\infty^+(S^1)$ and $\text{rot}(f) \in D(\tau)$. Then, there exists $h \in \text{Diff}_\infty^+(S^1)$ conjugating $f$ to $R_{\text{rot}(f)}$.
- Let $f \in \text{Diff}_\omega^+(S^1)$ and $\text{rot}(f) \in D(\tau)$. Then, there exists $h \in \text{Diff}_\omega^+(S^1)$ conjugating $f$ to $R_{\text{rot}(f)}$.

**Remark 24.** The above results can be generalized to larger classes of rotation number, satisfying a weaker condition than being Diophantine. Optimal conditions were studied by Yoccoz and identified in Brjuno numbers for the smooth case and in those satisfying the so-called $H$-condition (named in honour of Herman); we refer to [56, 57] for more details on these classes of numbers.

### 3.2. Linear response for Diophantine circle rotations.

In this subsection we describe how, as a corollary to KAM theory, one can prove the existence of linear response for Diophantine rotations.

Let us state the following version of KAM theorem, whose proof can be found in [53, Theorem 9.0.4] (cfr also [17, Theorem 2] and [19]).

**Theorem 25** (KAM Theorem for circle diffeomorphisms). Let $\alpha \in D(\tau)$, with $\tau > 1$ and let us consider a smooth family of circle diffeomorphisms

$$f_\varepsilon(x) = R_\alpha + \varepsilon u(x, \varepsilon) \quad |\varepsilon| < 1$$

with

1. $u(x, \varepsilon) \in C^\infty(S^1)$ for every $|\varepsilon| < 1$;
2. the map $\varepsilon \rightarrow u(\cdot, \varepsilon)$ is $C^\infty$;
3. $\int_{S^1} u(x, \varepsilon) dx = A\varepsilon^m + o(\varepsilon^m)$, where $A \neq 0$ and $m \geq 0$.

Then, there exists a Cantor set $C \subset (-1, 1)$ containing 0, such that for every $\varepsilon \in C$ the map $f_\varepsilon$ is smoothly conjugated to a rotation $R_{\alpha_\varepsilon}$, with $\alpha_\varepsilon \in D(\tau)$. More specifically, there exists

$$h_\varepsilon(x) = x + \varepsilon v(x, \varepsilon) \in C^\infty(S^1)$$

such that

$$S^1 \xrightarrow{f_\varepsilon} S^1 \xrightarrow{h_\varepsilon} S^1$$

$$\xrightarrow{h_\varepsilon} \xrightarrow{h_\varepsilon} \iff f_\varepsilon \circ h_\varepsilon = h_\varepsilon \circ R_{\alpha_\varepsilon}.$$

In particular:
• the maps $\varepsilon \mapsto h_\varepsilon$ and $\varepsilon \mapsto \alpha_\varepsilon$ are $C^\infty$ on the Cantor set $C$, in the sense of Whitney;
• $\alpha_\varepsilon = \alpha + A\varepsilon^{m+1} + o(\varepsilon^{m+1})$.

Remark 26. Observe that $f_\varepsilon$ does not have necessarily rotation number $\alpha$. In particular, the map $\text{rot} : \text{Diff}_0^1(\mathbb{S}^1) \to \mathbb{S}^1$ is continuous with respect to the $C^0$-topology (see for example [37, Ch. II, Proposition 2.7]).

Remark 27.

(i) Theorem 25 is proved in [53] in a more general form, considering also the cases of $u(x, \varepsilon)$ being analytic or just finitely differentiable (in this case, there is a lower bound on the needed differentiability, cfr Theorem 24). In particular, the proof of the asymptotic expansion of $\alpha_\varepsilon$ appears on [53, p. 149].

(ii) One could provide an estimate of the size of this Cantor set: there exist $M > 0$ and $r_0 > 0$ such that for all $0 < r < r_0$ the set $(-r, r) \cap C$ has Lebesgue measure $\geq Mr^{1/4}$ (see [53, formula (9.2)])

(iii) A version of this theorem in the analytic case, can be also found in [5, Theorem 2]; in particular, in [5, Sections 8] it is discussed the property of monogenically dependence of the conjugacy and the rotation number on the parameter.

These results can be extended to arbitrary smooth circle diffeomorphisms with Diophantine rotation numbers and to higher dimensional tori (see [53]).

Let us discuss how to deduce from this result the existence of linear response for the circle diffeomorphisms $f_\varepsilon$.

Theorem 28. Let $\alpha \in \mathcal{D}(\tau)$, with $\tau > 1$ and let us consider a family of circle diffeomorphisms obtained by perturbing the rotation $R_\alpha$ in the following way:

$$f_\varepsilon(x) = R_\alpha + \varepsilon u(x, \varepsilon) \quad |\varepsilon| < 1,$$

where $u(x, \varepsilon) \in C^\infty(\mathbb{S}^1)$, for every $|\varepsilon| < 1$, and the map $\varepsilon \mapsto u(\cdot, \varepsilon)$ is $C^\infty$.

Then, the circle rotation $R_\alpha$ admits linear response, in the limit as $\varepsilon$ goes to 0, by effect of this family of perturbations.

More precisely, there exists a Cantor set $C \subset (-1, 1)$ such that

$$\lim_{\varepsilon \in C, \varepsilon \to 0} \frac{\mu_\varepsilon - m}{\varepsilon} = 2\pi i \sum_{n \in \mathbb{Z} \setminus \{0\}} \frac{n \hat{u}(n)}{1 - e^{2\pi in\alpha}} e^{2\pi inx} \quad (\text{in the } L^1\text{-sense})$$

where $\mu_\varepsilon$ the unique invariant probability measure of $f_\varepsilon$, for $\varepsilon \in C$, and $\{\hat{u}(n)\}_{n \in \mathbb{Z}}$ the Fourier coefficients of $u(x, 0)$.

Remark 29. In this article we focus on the circle; however, a similar result could be proved for rotations on higher dimensional tori, by using analogous KAM results in that setting (see for example [53]).
As we have already observed in Remark 26, the rotation number of $f_\varepsilon$ varies continuously with respect to the perturbation, from here the need of taking the limit in (17) on a Cantor set of parameters (corresponding to certain Diophantine rotation numbers for which the KAM algorithm can be applied). Under the assumption that the perturbation does not change the rotation number, and this is Diophantine, then the KAM algorithm can be applied for all values of the parameters $\varepsilon$, hence $C$ coincides with the whole set of parameters; therefore the limit in (17) can be taken in the classical sense.

Corollary 30. Under the same hypotheses and notation of Theorem 28, if in addition we have that $\text{rot}(f_\varepsilon) = \alpha$ for every $|\varepsilon| < 1$, then there exists linear response without any need of restricting to a Cantor set and it is given by

\begin{equation}
\lim_{\varepsilon \to 0} \frac{\mu_\varepsilon - m}{\varepsilon} = 2\pi i \sum_{n \in \mathbb{Z} \setminus \{0\}} \left( \frac{n \hat{u}(n)}{1 - e^{2\pi i n \alpha}} \right) e^{2\pi i n x} \quad \text{(in the $L^1$-sense)}.
\end{equation}

Proof. (Corollary 30). As we have remarked above, this corollary easily follows from Theorem 28 by observing that $\text{rot}(f_\varepsilon) = \alpha \in \mathcal{D}(\tau)$ for every $|\varepsilon| < 1$, hence $C \equiv (-1,1)$. In fact, this follows from [53, Section 9.2, pp. 147-148]: in their notation our parameter $\varepsilon$ corresponds to $\mu$ and their $a(\mu)$ corresponds to our $\text{rot}(f_\varepsilon)$. In particular, they define the Cantor set as $\mathcal{C}_F = v^{-1} (\mathcal{D}_F)$ (see [53, p.148]): in our notation this corresponds to the values of $\varepsilon \in (-1,1)$ for which $\text{rot}(f_\varepsilon)$ belongs to the a certain set of Diophantine numbers that includes $\alpha$. Since, by hypothesis, $\text{rot}(f_\varepsilon) \equiv \alpha$, it follows that $C \equiv (-1,1)$ and, in particular, the limit in (17) is meant in the classical sense. □

Let us now prove Theorem 28.

Proof. (Theorem 28). First of all, applying Theorem 25, it follows that for every $\varepsilon \in C$, the map $f_\varepsilon := R_\alpha + \varepsilon u(x, \varepsilon)$ possesses a unique invariant probability measure given by

$$
\mu_\varepsilon = h_\varepsilon \ast m
$$

where $m$ denotes the Lebesgue measure on $S^1$ and $h_\varepsilon \ast m$ denotes the push-forward by $h_\varepsilon$; in particular, $\mu_0 = m$. This measure is absolutely continuous with respect to $m$ and its density is given by

\begin{equation}
\frac{d\mu_\varepsilon}{dx}(x) = \frac{1}{\partial_x h_\varepsilon(h_\varepsilon^{-1}(x))}.
\end{equation}

In fact, if $A$ is a Borel set in $S^1$, then

$$
\mu_\varepsilon(A) = \int_A \mu_\varepsilon(dy) = \int_{h_\varepsilon(A)} \partial_x(h_\varepsilon^{-1})(x) dx = \int_{h_\varepsilon(A)} \frac{dx}{\partial_x h_\varepsilon(h_\varepsilon^{-1}(x))}.
$$

Hence, it follows from (19) that

\begin{equation}
\frac{d\mu_\varepsilon}{dx}(x) = \frac{1}{\partial_x h_\varepsilon(h_\varepsilon^{-1}(x))} = \frac{1}{1 + \varepsilon \partial_x v(h_\varepsilon^{-1}(x), 0) + o(\varepsilon)}
\end{equation}

\begin{equation}
= \frac{1}{1 + \varepsilon \partial_x v(x, 0) + o(\varepsilon)} = 1 - \varepsilon \partial_x v(x, 0) + o(\varepsilon),
\end{equation}
where \( o_C(\varepsilon) \) denotes a term that goes to zero faster than \( \varepsilon \in C \), uniformly in \( x \).

Then the linear response is given by

\[
\dot{\mu} = \lim_{\varepsilon \in C, \varepsilon \to 0} \frac{\mu - \mu_0}{\varepsilon} = \lim_{\varepsilon \in C, \varepsilon \to 0} \frac{\mu - \mu_0}{\varepsilon}
\]

which, passing to densities and using (20), correspond to

\[
\lim_{\varepsilon \in C, \varepsilon \to 0} \frac{1}{\varepsilon} (1 - \varepsilon \partial_x v(x, 0) + o(\varepsilon) - 1) = -\partial_x v(x, 0).
\]

Giving a formula for the response

\[
\frac{d\dot{\mu}}{dx}(x) = -\partial_x v(x, 0).
\]

Moreover, we can find a more explicit representation formula. In fact, it follows from (18) that \( f_\varepsilon \circ h_\varepsilon = h_\varepsilon \circ R_{\alpha_\varepsilon} \):

\[
x + \varepsilon v(x, \varepsilon) + \alpha + \varepsilon u(x + \varepsilon v(x, \varepsilon), \varepsilon) = x + \alpha_\varepsilon + \varepsilon v(x + \alpha_\varepsilon, \varepsilon).
\]

Recall, from the statement of Theorem 25 that

\[
\alpha_\varepsilon = \alpha + A\varepsilon^{m+1} + o(\varepsilon^{m+1}),
\]

where \( m \) and \( A \) are defined by (see item (ii) in Theorem 25)

\[
\langle u(\cdot, \varepsilon) \rangle := \int_{S^1} u(x, \varepsilon) dx = A\varepsilon^m + o(\varepsilon^m).
\]

Hence, expanding equation (22) in terms of \( \varepsilon \) and equating the terms of order 1, we obtain the following (observe that \( \alpha_\varepsilon \) will contribute to the first order in \( \varepsilon \) only if \( m = 0 \) and, therefore, \( A = \langle u(\cdot, 0) \rangle := \int_{S^1} u(x, 0) dx \neq 0 \)):

\[
v(x + \alpha, 0) - v(x, 0) = u(x, 0) - \langle u(\cdot, 0) \rangle \quad \forall x \in S^1,
\]

the so-called homological equation.

Observe that it makes sense that we need to subtract to \( u(x, 0) \) its average, if this is not zero. In fact, in order for (23) to have a solution, its right-hand side must have zero average: to see this, it is sufficient to integrate both sides and use that the Lebesgue measure is invariant under \( R_{\alpha_\varepsilon} \):

\[
\int_{S^1} u(x, 0) dx = \int_{S^1} v(x + \alpha, 0) dx - \int_{S^1} v(x, 0) dx = 0.
\]

Let us now find an expression for \( v(x, 0) \) in Fourier series. In fact, let us consider:

\[
v(x, 0) := \sum_{n \in \mathbb{Z}} \hat{v}(n)e^{2\pi inx} \quad \text{and} \quad u(x, 0) := \sum_{n \in \mathbb{Z}} \hat{u}(n)e^{2\pi inx}.
\]

In Fourier terms, (23) becomes:

\[
\sum_{n \in \mathbb{Z}} \hat{v}(n) (e^{2\pi in\alpha} - 1) e^{2\pi inx} = \sum_{n \in \mathbb{Z}\backslash\{0\}} \hat{u}(n)e^{2\pi inx}
\]

and therefore for \( n \neq 0 \)

\[
\hat{v}(n) = \frac{\hat{u}(n)}{e^{2\pi in\alpha} - 1};
\]

we do not determine \( \hat{v}(0) \), as it should be expected, since \( v \) is determined by (23) only up to constants.
Substituting in (21), we conclude:
\[
\frac{d\hat{\mu}}{dx}(x) = -\partial_x v(x,0) = -2\pi i \sum_{n \in \mathbb{Z}} n \hat{v}(n)e^{2\pi inx}
\]
\[
= 2\pi i \sum_{n \in \mathbb{Z}\backslash\{0\}} \left( \frac{n \hat{u}(n)}{1 - e^{2\pi in\alpha}} \right) e^{2\pi inx}.
\]

\[\square\]

4. Beyond rotations: the case of circle diffeomorphisms

In this section, we want to describe how it is possible to extend our previous results from irrational rotations to diffeomorphisms of the circle having irrational rotation number.

We prove the following:

**Theorem 31.** Let \( T_0 \) be an orientation preserving diffeomorphism of the circle with an irrational rotation number \( \alpha \) and such that \( \log(T') \) has bounded variation (for example \( f \) is of class \( C^2 \)). Let \( \mu_0 \) be its unique invariant (absolutely continuous) probability measure (see Theorem 19). Let \( \{T_\delta\}_{0 \leq \delta \leq \delta} \) be a family of Borel measurable maps of the circle such that
\[
\sup_{x \in \mathbb{S}^1} |T_0(x) - T_\delta(x)| \leq \delta.
\]
Suppose that for each \( 0 \leq \delta \leq \delta \), \( \mu_\delta \) is an invariant measure of \( T_\delta \). Then
\[
\lim_{\delta \to 0} \int_{\mathbb{S}^1} f \, d\mu_\delta = \int_{\mathbb{S}^1} f \, d\mu_0
\]
for all \( f \in C^0(\mathbb{S}^1) \).

The proof will follow by combining Theorem 2 with Denjoy Theorem 19.

**Proof of Theorem 31.** By Theorem 19 we can conjugate \( T_0 \) with the rotation \( R_\alpha \). We apply the same conjugation to \( T_\delta \) for each \( \delta > 0 \) obtaining a family of maps \( U_\delta := h \circ T_\delta \circ h^{-1} \). We summarize the situation in the following diagram
\[
\begin{array}{cccccc}
\mathbb{S}^1 & \xrightarrow{T_0} & \mathbb{S}^1 & \xrightarrow{T_\delta} & \mathbb{S}^1 \\
\downarrow h & & \downarrow h & & \downarrow h \\
\mathbb{S}^1 & \xrightarrow{R_\alpha} & \mathbb{S}^1 & \xrightarrow{U_\delta} & \mathbb{S}^1
\end{array}
\]

Since \( h \) is an homeomorphism of a compact space it is uniformly continuous. This implies that
\[
\lim_{\delta \to 0} \sup_{x \in \mathbb{S}^1} |R_\alpha(x) - U_\delta(x)| = 0.
\]
Let \( \overline{\mu}_\delta := h_\ast \mu_\delta \). These measures are invariant for \( U_\delta \). Then, by Theorem 2 we get
\[
\lim_{\delta \to 0} ||\overline{\mu}_\delta - m||_V = 0.
\]
This implies (uniformly approximating any continuous function with a sequence of Lipschitz ones) that for each \( g \in C^0(\mathbb{S}^1) \)
\[
\lim_{\delta \to 0} \int_{\mathbb{S}^1} g \, d\overline{\mu}_\delta = \int_{\mathbb{S}^1} g \, dm.
\]
Now consider \( f \in C^0(S^1) \) and remark that (using the definition of push-forward of a measure)

\[
\int_{S^1} f \, d\mu_\delta = \int_{S^1} f \circ h^{-1} \circ h \, d\mu_\delta = \int_{S^1} f \circ h^{-1} \, d\mu_0,
\]

\[
\int_{S^1} f \, d\mu_0 = \int_{S^1} f \circ h^{-1} \, d\mu_0.
\]

By \([25]\) considering \( g = f \circ h^{-1} \) this shows

\[
\lim_{\delta \to 0} \int_{S^1} f \, d\mu_\delta = \int_{S^1} f \, d\mu_0.
\]

\[\square\]

Similarly, one can extend the quantitative stability results proved in Theorem \([17]\) to smooth diffeomorphisms of the circle.

**Remark 32.** We point out that the following theorem holds under much less regularity for \( T_0 \) (the proof remains the same). In fact, it is enough that \( T_0 \in C^r(S^1) \) with \( r \) sufficiently big so that the conjugation \( h \) is bi-Lipschitz; compare with Theorem \([23]\).

**Theorem 33.** Let \( T_0 \) be a \( C^\infty \) diffeomorphism of the circle with Diophantine rotation number \( \alpha \in \mathcal{D}(\tau) \). Let \( \{T_\delta\}_{0 \leq \delta \leq \overline{\delta}} \) be a family of Borel measurable maps of the circle such that

\[
\sup_{x \in S^1} |T_0(x) - T_\delta(x)| \leq \delta.
\]

Suppose that for each \( 0 \leq \delta \leq \overline{\delta} \), \( \mu_\delta \) is an invariant measure of \( T_\delta \). Then, for each \( \ell < \frac{1}{\gamma(\alpha)+1} \) we have:

\[
||m - \mu_\delta||_W = O(\delta^\ell).
\]

**Proof.** By Theorem \([23]\), there exists \( h \in \text{Diff}^\infty(S^1) \) conjugating \( T_0 \) with the rotation \( R_\alpha \). We apply the same conjugation to \( T_\delta \) for each \( \delta > 0 \) obtaining a family of maps \( U_\delta \). The situation is still summarized by \([24]\). Since \( h \) is a bilipschitz map we have

\[
\lim_{\delta \to 0} \sup_{x \in S^1} |R_\alpha(x) - U_\delta(x)| = 0
\]

and there is a \( C \geq 1 \) such that for any pair of probability measures \( \mu_1, \mu_2 \)

\[
C^{-1} ||\mu_1 - \mu_2||_W \leq ||h_*^{-1}\mu_1 - h_*^{-1}\mu_2||_W \leq C||\mu_1 - \mu_2||_W
\]

(and the same holds for \( h_* \)). Let \( \overline{\mu}_\delta := h_*(\mu_\delta) \). These measures are invariant for \( U_\delta \).

By Theorem \([13]\) we then get that for each \( \ell < \frac{1}{\gamma(\alpha)+1} \) we have:

\[
||m - \overline{\mu}_\delta||_W = O(\delta^\ell).
\]

This imply

\[
||\mu_0 - \mu_\delta||_W = ||h_*^{-1}m - h_*^{-1}\overline{\mu}_\delta||_W = O(\delta^\ell).
\]

\[\square\]
Finally, one can also extend the existence of linear response, along the same lines of Theorem 28 and Corollary 30. In fact, as observe in Remark 27 (iii), KAM theorem can be extended to sufficiently regular diffeomorphisms of the circle (one can prove it either directly, e.g., [5, 17, 47, 51, 53], or by combining the result for rotations of the circle, with Theorem 23). Since the proof can be adapted mutatis mutandis, we omit further details.

5. Stability under discretization and numerical truncation

As an application of what discussed in this section we want to address the following question:

**Question:** Why are numerical simulations generally quite reliable, in spite of the fact that numerical truncations are quite bad perturbations, transforming the system into a piecewise constant one, having only periodic orbits?

Let us consider the uniform grid $E_N$ on $S^1$ defined by

$$E_N = \left\{ \frac{i}{N} \in \mathbb{R}/\mathbb{Z} : 1 \leq i \leq N \right\}.$$

In particular when $N = 10^k$ the grid represents the points which are representable with $k$ decimal digits. Let us consider the projection $P_N : S^1 \to E_N$ defined by

$$P_N(x) = \frac{\lfloor Nx \rfloor}{N},$$

where $\lfloor \cdot \rfloor$ is the floor function.

Given a map $T : S^1 \to S^1$ and let $N \in \mathbb{N}$; we define its $N$-discretization $T_N : S^1 \to S^1$ by

$$T_N(x) := P_N(T(x)).$$

This is an idealized representation of what happens if we try to simulate the behavior of $T$ on a computer, having $N$ points of resolution. Of course the general properties of the systems $T_N$ and $T$ are a priori completely different, starting from the fact that $T_N$ is forced to be periodic. Still these simulations gives in many cases quite a reliable picture of many aspects of the behavior of $T$, which justifies why these naive simulations are still much used in many applied sciences.

Focusing on the statistical properties of the system and on its invariant measures, one can investigate whether the invariant measures of the system $T_N$ (when they exist) converge to the physical measure of $T$, and in general if they converge to some invariant measure of $T$. In this case, the statistical properties of $T$ are in some sense robust under discretization. Results of this kind have been proved for some classes of piecewise expanding maps (see [18], [32]) and for topologically generic diffeomorphisms of the torus (see [33], [35], [46]).

Since the discretization is a small perturbation in the uniform convergence topology, a direct application of Theorem 31 gives

**Corollary 34.** Let $T_0$ be an orientation preserving diffeomorphism of the circle with an irrational rotation number $\alpha$ and such that $\log(T_0)$ has bounded variation and let
$N \geq 1$. Let $T_N = P_N \circ T_0$ be the family of maps given by its $N$–discretizations. Suppose $\mu_N$ is an invariant measure of $T_N$. Then
\[
\lim_{N \to \infty} \int_{S^1} f \, d\mu_N = \int_{S^1} f \, d\mu_0
\]
for all $f \in C^0(S^1)$.

**Proof.** The statement follows by Theorem 3.1 noticing that
\[
\sup_{x \in S^1} |T_0(x) - T_N(x)| \leq \frac{1}{N}.
\]
\[\square\]

We think this result is very similar to the one shown in Proposition 8.1 of [46]. Comparing this kind of results with the ones in [34], we point out that in this statement we do not suppose the system to be topologically generic and that the convergence is proved for all discretizations, while in [34] the convergence is proved for a certain sequence of finer and finer discretizations.

As an application of our quantitative stability result (Theorem 13 and 33), we can also provide a quantitative estimate for the speed of convergence of the invariant measure of the $N$-discretized system to the original one. We remark that as far as we know, there are no other similar quantitative convergence results of this kind in the literature.

**Corollary 35.** Let $T_0$ be a $C^\infty$ diffeomorphism of the circle with Diophantine rotation number $\alpha \in D(\tau)$. Let $T_N = P_N \circ T_0$ be the family of its $N$-discretizations. Suppose $\mu_N$ is an invariant measure of $T_N$. Then, for each $\ell < \frac{1}{\gamma(\alpha)+1}$
\[
\|m - \mu_N\|_W = O(N^{-\ell}).
\]

The proof of Corollary 35 is essentially the same as the one of Corollary 34.

**References**

[1] L. Ambrosio, N. Gigli, G. Savaré. Gradient flows in metric spaces and in the space of probability measures (Second edition). Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2008. x+334 pp.

[2] J. Alves, **Strong statistical stability of non-uniformly expanding maps.** Nonlinearity, 17, 4, 1109-1215 (2004).

[3] J. F. Alves, M. Soufi *Statistical Stability in Chaotic Dynamics* Progress and Challenges in Dyn. Sys. Springer Proc. in Math. & Statistics V. 54, 2013, pp 7-24

[4] J. Alves, M. Viana *Statistical stability for robust classes of maps with non-uniform expansion.* Ergodic Theory and Dynam. Systems, 22 , 1, 1-32 (2002).

[5] V.I.Arnold. Small divisors I: On mappings of the circle onto itself. Izvestiya Akad. Nauk SSSR, Ser. Mat. 25 (1961), 21–86 (in Russian); English translation: Amer. Math. Soc. Transl., Ser. 2 46 (1965), 213–284; Erratum: Izvestiya Akad. Nauk SSSR, Ser. Mat. 28 (1964), 479–480 (in Russian).

[6] W. Bahsoun, S. Galatolo, I. Nisoli, X. Niu *A Rigorous Computational Approach to Linear Response* Nonlinearity, Volume 31, Number 3 pp. 1073–1109 (2018)

[7] W. Bahsoun, M. Ruziboev, B. Saussol *Linear response for random dynamical systems* arXiv:1710.03706

[8] W. Bahsoun, S. Vaienti *Metastability of certain intermittent maps* Nonlinearity 25, 1, 107, (2012)
[9] W. Bahsoun, B. Saussol Linear response in the intermittent family: differentiation in a weighted $C^0$-norm. [arXiv:1512.01080]
[10] V. Baladi Linear response, or else ICM Seoul 2014 talk. [arXiv:1408.2937]
[11] V. Baladi, M. Benedicks, N. Schnellmann Whitney-Hölder continuity of the SRB measure for transversal families of smooth unimodal maps. Invent. Math. 201 773-844 (2015)
[12] Baladi, V., Smania, D., Linear response formula for piecewise expanding unimodal maps, Nonlinearity 21 (2008), 677â–$\$711. (Corrigendum, Nonlinearity 25 (2012), 2203â–$\$2205.)
[13] Baladi, V., Smania, D., Linear response for smooth deformations of generic nonuniformly hyperbolic unimodal maps, Ann. Sci. Ec. Norm. Sup. 45 (2012), 861â–$\$926.
[14] V. Baladi, T. Kuna and V. Lucarini Linear and fractional response for the SRB measure of smooth hyperbolic attractors and discontinuous observables Nonlinearity 30 1204-1220 (2017)
[15] V. Baladi, M. Todd Linear response for intermittent maps Comm. in Math. Phys. V 347, n 3, pp 857–874 (2016)
[16] M. Blank, G. Keller Random perturbations of chaotic dynamical systems: stability of the spectrum Nonlinearity, Volume 11, pp. 1351–1364 (1998)
[17] H.W. Broer, Mikhail B. Sevryuk. KAM Theory: Quasi-periodicity in Dynamical Systems Handbook of Dynamical Systems Vol. 3 (2010), Elsevier/North-Holland, Amsterdam, pp xi–+543, Editors: Editors: H. Broer F. Takens B. Hasselblatt.
[18] A. Boyarsky. Computer orbits. Comput. Math. Appl. Ser. A 12(10) (1986), 1057–1064.
[19] R. Calleja, A. Celletti, R. de la Llave. Whitney regularity and monogenicity of quasi-periodic solutions in KAM theory: a simple approach based on a-posteriori theorems. Preprint 2020.
[20] R. de la Llave. A tutorial on KAM theory. Smooth ergodic theory and its applications (Seattle, WA, 1999), 175–192, Proc. Sympos. Pure Math., 69, Amer. Math. Soc., Providence, RI, 2001.
[21] A. Denjoy Sur les courbes définies par les équations différentielles a la surface du tore.,J. Math. Pures et Appl. 11, serie 9, 333-375, (1932)
[22] D. Dolgopyat. On dynamics of mostly contracting diffeomorphisms , Comm. in Math. Physics , 213 (2000) 181-201.
[23] D. Dolgopyat On differentiability of SRB states for partially hyperbolic systems , Invent. Math. 155 (2004) 389–449.
[24] D. Dolgopyat, Prelude to a kiss, Modern dynamical systems (ed. M. Brin , B.Hasselblatt and Ya. Pesin ), Cambridge Univ Press, 2004, 313-324.
[25] H. S. Dumas. The KAM story. A friendly introduction to the content, history, and significance of classical Kolmogorov-Arnold-Moser theory. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, xvi+361 pp., 2014.
[26] H. Eliasson, B. Fayad, R. Krikorian. Jean-Christophe Yoccoz and the theory of circle diffeomorphisms. La gazette des mathématiciens, Société mathématiques de France (Jean-Christophe Yoccoz - numéro spécial Gazette, April 2018): 55–66.
[27] S. Galatolo Quantitative statistical stability and speed of convergence to equilibrium for partially hyperbolic skew products . J. Éc. Pol. Math., 5, 377–405 (2018)
[28] S. Galatolo Quantitative statistical stability and convergence to equilibrium. An application to maps with indifferent fixed points Chaos, Solitons & Fractals V. 103, pp. 596-601 (2017).
[29] S. Galatolo, P. Giulietti A Linear Response for dynamical systems with additive noise Nonlinearity, 32, n. 6, pp. 2269-2301 (2019)
[30] S. Galatolo, J. Sedro Quadratic response of random and deterministic dynamical systems. Chaos 30 (2020), no. 2, 023113, 15 pp.
[31] M. Ghil and V. Lucarini The Physics of Climate Variability and Climate Change arXiv preprint. [arXiv:1910.00583] (2019)
[32] P. Gora and A. Boyarsky. Why computers like Lebesgue measure. Comput. Math. Appl. 16(4) (1988), 321–329.
[33] S. Gouëzel, Liverani, C. Banach spaces adapted to Anosov systems. Ergodic Theory and Dynamical Systems, 26(1), 189-217 (2006)
[34] P.-A. Guihéneuf Physical measures of discretizations of generic diffeomorphisms Ergod. Th. & Dynam. Sys. (2018), 38, 1422–1458
[35] P.-A. Guihéneuf. Discretisations spatiales de systèmes dynamiques génériques. PhD Thesis, Université Paris-Sud, 2015.
[36] M. Hairer, A. Majda A simple framework to justify linear response theory Nonlinearity, 23, 909–922, (2010)
[37] M. Herman. Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations. Inst. Hautes Etudes Sci. Publ. Math. No. 49 (1979), 5–233.

[38] G. Keller, C. Liverani Stability of the spectrum for transfer operators Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 28 no. 1, 141–152 (1999).

[39] A. Korepanov Linear response for intermittent maps with summable and nonsummable decay of correlations. arXiv:1508.06571

[40] N. Kryloff and N. Bogoliouboff. La théorie générale de la mesure dans son application à l’étude des systèmes dynamiques de la mécanique non linéaire. Ann. of Math. (2), 38 (1):65–113, 1937.

[41] L. Kuipers and H. Niederreiter. Uniform distribution of sequences. Pure and Applied Mathematics.Wiley-Interscience, 1974.

[42] K. K. Lin Convergence of invariant densities in the small-noise limit Nonlinearity 18, 2 pp. 659-683 (2005)

[43] J. Massetti. A normal form à la Moser for diffeomorphisms and a generalization of Rîssmann’s translated curve theorem to higher dimensions. Anal. PDE 11 (1), 149–170, 2018

[44] J.N. Mather, G. Forni. Action minimizing orbits in Hamiltonian systems. In: Graffi (ed) Transition to Chaos in Classical and Quantum Mechanics. Springer LNM 1589 (1992): 92–188.

[45] R. J. Metzger Stochastic Stability for Contracting Lorenz Maps and Flows Comm. Math. Phys. 212, pp. 277–296 (2000)

[46] Miernowski, T.. DiscrÃľtisations des homÃľomorphismes du cercle. Erg. Th. Dyn. Sys. 26(6) (2006), 1867تا؛1903.

[47] J. K. Moser On invariant curves of area-preserving mappings of an annulus. Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II (1962), 1–20.

[48] M. Pollicott, P. Vytynova, Linear response and periodic points. Nonlinearity 29, no. 10, 3047–3066 (2016)

[49] J. A. Vano. A Nash-Moser Implicit Function Theorem with Whitney Regularity and Applications. Ph. D. Dissertation, The University of Texas at Austin (2002), pp. 172 (downloadable from https://web.ma.utexas.edu/mp_arc/c/02/02-276.pdf).

[50] C. E. Wayne. An introduction to KAM theory. Dynamical systems and probabilistic methods in partial differential equations (Berkeley, CA, 1994), 3–29, Lectures in Appl. Math., 31, Amer. Math. Soc., Providence, RI, 1996.

[51] C. Wormell; G. Gottwald On the validity of linear response theory in high-dimensional deterministic dynamical systems. J. Stat. Phys. 172, no. 6, 1479–1498. (2018)
