DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume **42**, Number **2**, February **2022**

doi:10.3934/dcds.2021138
pp. **815–839**

QUANTITATIVE STATISTICAL STABILITY AND LINEAR RESPONSE FOR IRRATIONAL ROTATIONS AND DIFFEOMORPHISMS OF THE CIRCLE

Stefano Galatolo
Dipartimento di Matematica, Università di Pisa
Largo Bruno Pontecorvo 5, 56127 Pisa, Italy

Alfonso Sorrentino
Dipartimento di Matematica, Università degli Studi di Roma "Tor Vergata"
Via della ricerca scientifica 1, 00133, Roma, Italy

(Communicated by Mark F. Demers)

Abstract. We prove quantitative statistical stability results for a large class of small C^0 perturbations of circle diffeomorphisms with irrational rotation numbers. 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 rotation number. The set of admissible perturbations includes the ones coming from spatial discretization and hence numerical truncation. We also show linear response for smooth perturbations that preserve the rotation number, as well as for more general ones. This is done by means of classical tools from KAM theory, while the quantitative stability results are obtained by transfer operator techniques applied to suitable spaces of measures with a weak topology.

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[1] 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

2020 *Mathematics Subject Classification.* Primary: 37C40; Secondary: 37A30, 37J40, 37A45, 37C05, 37C75, 37E45.

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

[1]The 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 5 or [48]). In other contexts, other kinds of measures might be considered, for example, the so-called measures of maximal entropy.

well as understanding the *regularity* of their behavior, for instance differentiability, Lipschitz or Hölder dependence, etc...

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 $\{(X, T_\delta)\}_{\delta\in[0,\bar{\delta})}$, where $T_\delta : X \to X$ are such that $T_\delta \to T_0$, as $\delta \to 0$, in some suitable topology.

For each $\delta \in [0, \bar{\delta})$ let μ_δ be an invariant Borel probability measure for the system (X, T_δ) (if T_δ is continuous such measures must exist by the classical Krylov-Bogoliubov theorem). We aim to get information on the regularity of this family of measures, by investigating the regularity of the map $\delta \longmapsto \mu_\delta$. 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 $\mathcal{P}(X)$, endowed with a suitable weak norm, see subection 2.1 for more details.

We say that (X, T_0, μ_0) is *statistically stable* (with respect to the considered class of perturbations) if the map $\delta \longmapsto \mu_\delta$ is continuous at $\delta = 0$ (with respect to the chosen topology on the space of measures in which μ_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 smootheness 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 [26] 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] for a general survey and [1], [3] ,[9], [8], [11], [13], [12], [15], [19], [20], [24], [23], [29], [34], [35], [49] for other results about statistical stability for different classes of deterministic 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 happens 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 changes the system into a periodic one, destroying many features of the original dynamics, yet this kind 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 dicretization (see e.g. [28], [30], [31], [37]). We refer to Section 5 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 [23], 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 $\mathcal{P}(X)$, related to the so-called Wassertein 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 14 for the general upper bounds and Proposition 17 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 30 and Corollary 32.
4. We extend these qualitative and quantitative stability results to diffeomorphisms of the circle satisfying suitable assumptions; see Theorems 33 and 35.
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, 13, 14] (see also [10, Section 4]), where one has a smooth behavior for the response of statistical properties of the system to 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 14 – are obtained by means of the so-called Denjoy-Koksma inequality (see Theorem 13).

On the other hand, results in item 3 are obtained as an application of KAM theory for circle maps (see Theorem 27), with a particular focus on the dependence of the KAM-construction on the perturbative parameter. In Section 3 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 diffeomorphims, 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.

2. **Statistical stability of irrational rotations.**

Irrational rotations on the circle (*i.e.*, maps of the circle to itself of the form $x \longmapsto x+\rho$ with $\rho \in \mathbb{R}\setminus\mathbb{Q}$) preserve the Lebesgue measure m on the circle $\mathbb{S}^1 := \mathbb{R}/\mathbb{Z}$ 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 Wassertein distance. This qualitative result might not be surprising for experts, however the construction that we apply also leads to quantitative estimates on the statistical stability, which will be presented in the next subsections.

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 $g : X \longrightarrow \mathbb{R}$ is a Lipschitz function, we denote its (best) Lipschitz constant by $\mathrm{Lip}(g)$, *i.e.*

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

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

$$W(\mu,\nu) := \sup_{\text{Lip}(g)\leq 1, ||g||_\infty \leq 1} \left| \int_{\mathbb{S}^1} g d\mu - \int_{\mathbb{S}^1} g d\nu \right|. \tag{1}$$

We denote

$$\|\mu\|_W := W(0,\mu),$$

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 [4] 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 μ by T and denoted by $T_*\mu$). If 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. We remark that the following result is qualitative, however the general construction that we implement can be exploited to get quantitative estimates too, as it will be shown in the next subsections.

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

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

Suppose μ_δ is an invariant measure[2] of T_δ. 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 $\mathbb{S}^1$ and let L_δ be the transfer operator associated to a measurable map T_δ. Suppose that $\mu_\delta = L_\delta\mu_\delta$. Then, for each $n \geq 1$*

$$\|\mu_\delta - m\|_W \;\leq\; \Big\|m - \frac{1}{n}\sum_{1\leq i\leq n} L^i\mu_\delta\Big\|_W \;+\; \frac{(n-1)}{2}\,\Big\|(L - L_\delta)\mu_\delta\Big\|_W \tag{2}$$

where $L^i := L \circ \ldots \circ L$ (i-times).

[2] We remark that we are not supposing μ_δ being the unique invariant measure of T_δ. When the invariant measure is not unique, the statement hence holds for every such measure. On the other hand, if there is no invariant measure for T_δ, then the statement is empty.

Proof. The proof is a direct computation. Since $\mu_\delta = L_\delta \mu_\delta$ and m is invariant for L, then

$$\begin{aligned} \|\mu_\delta - m\|_W &\leq \Big\|\frac{1}{n}\sum_{1\leq i\leq n} L_\delta^i \mu_\delta - \frac{1}{n}\sum_{1\leq i\leq n} L^i m\Big\|_W \\ &\leq \Big\|\frac{1}{n}\sum_{1\leq i\leq n} L^i(m-\mu_\delta)\Big\|_W + \Big\|\frac{1}{n}\sum_{1\leq i\leq n}(L^i - L_\delta^i)\mu_\delta\Big\|_W. \end{aligned} \quad (3)$$

Since

$$L^i - L_\delta^i = \sum_{k=1}^{i} L^{i-k}(L-L_\delta)L_\delta^{k-1}$$

then

$$\begin{aligned} (L^i - L_\delta^i)\mu_\delta &= \sum_{k=1}^{i} L^{i-k}(L-L_\delta)L_\delta^{k-1}\mu_\delta \\ &= \sum_{k=1}^{i} L^{i-k}(L-L_\delta)\mu_\delta. \end{aligned}$$

Being L is the transfer operator associated to an isometry, then

$$\|L^{i-k}(L-L_\delta)\mu_\delta\|_W \leq \|(L-L_\delta)\mu_\delta\|_W \quad (4)$$

and consequently

$$\|(L^i - L_\delta^i)\mu_\delta\|_W \leq (i-1)\|(L-L_\delta)\mu_\delta\|_W.$$

Substituting in (3), we conclude

$$\|\mu_\delta - m\|_W \leq \Big\|\frac{1}{n}\sum_{1\leq i\leq n} L^i(m-\mu_\delta)\Big\|_W + \frac{(n-1)}{2}\|(L-L_\delta)\mu_\delta\|_W.$$

□

Lemma 4. *Under the assumptions of Theorem 2, let $\{\mu_\delta\}_{0\leq\delta\leq\overline{\delta}}$ be a family of Borel probability measures on $\mathbb{S}^1$, then*

$$\lim_{n\to\infty}\Big\|m - \frac{1}{n}\sum_{1\leq i\leq n} L^i\mu_\delta\Big\|_W = 0$$

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

$$\sup_{0\leq\delta\leq\overline{\delta}}\Big\|m - \frac{1}{n}\sum_{1\leq i\leq n} L^i\mu_\delta\Big\|_W \leq \varepsilon.$$

Proof. Let δ_{x_o} be the delta-measure concentrated at a point $x_0 \in \mathbb{S}^1$. 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 μ_δ 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 $x_1, ..., x_k \in \mathbb{S}^1$and $\lambda_1, ..., \lambda_k \geq 0$, with $\sum_{i\leq k}\lambda_i = 1$ such that

$$\Big\|\mu_\delta - \sum_{1\leq i\leq k}\lambda_i\delta_{x_i}\Big\|_W \leq \varepsilon.$$

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

$$\left\|L^n\mu_\delta - L^n\big(\sum_{1\leq i\leq k}\lambda_i\delta_{x_i}\big)\right\|_W \leq \varepsilon,$$

which implies

$$\left\|m - L^n\mu_\delta\right\|_W \leq \varepsilon + \left\|m - L^n\big(\sum_{1\leq i\leq k}\lambda_i\delta_{x_i}\big)\right\|_W$$

and

$$\left\|m - \frac{1}{n}\sum_{1\leq i\leq n}L^i\mu_\delta\right\|_W \leq \varepsilon + \left\|m - \frac{1}{n}\sum_{1\leq j\leq n}L^j\big(\sum_{i\leq k}\lambda_i\delta_{x_i}\big)\right\|_W.$$

We estimate now the behavior of the right hand side of the last inequality as $n \to \infty$. For any n we have

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

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

We can now prove Theorem 2.

Proof of Theorem 2. Let L_δ be the transfer operator associated to T_δ. 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 δ. Since

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

then $\|(L - L_\delta)\mu_\delta\|_W \leq \delta$; indeed for each g such that $\mathrm{Lip}(g) \leq 1$, we get

$$\begin{aligned}\left|\int_{\mathbb{S}^1} g\ d[(L_\delta - L_0)(\mu_\delta)]\right| &= \left|\int_{\mathbb{S}^1}(g\circ T_0 - g\circ T_\delta)d\mu_\delta\right| \\ &\leq \sup_{x\in\mathbb{S}^1}|T_0(x) - T_\delta(x)|\ \mu_\delta(\mathbb{S}^1).\end{aligned}$$

Therefore, we conclude that

$$\lim_{\delta\to 0}\|(L - L_\delta)\mu_\delta\|_W = 0. \tag{5}$$

By Lemma 3 we get that for each n

$$\left\|\mu_\delta - m\right\|_W \leq \left\|m - \frac{1}{n}\sum_{1\leq i\leq n}L^i\mu_\delta\right\|_W + \frac{(n-1)}{2}\left\|(L - L_\delta)\mu_\delta\right\|_W. \tag{6}$$

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 (5), we can choose δ sufficiently mall so to make $\frac{(n-1)}{2}\|(L - L_\delta)\mu_\delta\|_W$ as small as needed, hence proving the statement. □

Remark 5. The qualitative stability statements with respect to the Wasserstein distance proved in this section for circle rotations, extend directly to many other systems, for example to uniquely ergodic rotations on the multidimensional torus. In fact in the proof, aside of the general properies of the Wasserten distance and of pushforward maps, we only use that the system is uniquely ergodic, and the map is an isometry. This property could also be relaxed to a non-expansive property, ensuring that (4) is satisfied.

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 numers*. 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 $\delta \longmapsto \mu_\delta$ 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 [36]): this concept expresses quantitatively the rate of approximability of an irrational number by sequences of rationals.

In what follows, we will also use $\|\cdot\|_{\mathbb{Z}}$ to denote the distance from a real number to the nearest integer.

Definition 6. If α is irrational, the Diophantine type of α is defined by

$$\gamma(\alpha) := \sup\{\gamma \geq 0 : \liminf_{k\to\infty} \; k^\gamma \|k\alpha\|_{\mathbb{Z}} = 0\}.$$

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

Remark 7. The Diophantine type of α can be also defined by

$$\begin{aligned} \gamma(\alpha) \quad &:= \quad \inf\left\{\gamma \geq 0 : \; \exists c > 0 \text{ s.t. } \|k\alpha\|_{\mathbb{Z}} \geq c_0 |k|^{-\gamma} \; \forall \, k \in \mathbb{Z} \setminus \{0\}\right\} \\ &= \quad \inf\left\{\gamma \geq 0 : \; \exists c > 0 \text{ s.t. } \left|\alpha - \frac{p}{q}\right| \geq \frac{c}{|q|^{\gamma+1}} \quad \forall \, \frac{p}{q} \in \mathbb{Q} \setminus \{0\}\right\}. \end{aligned}$$

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

$$\left|\alpha - \frac{p}{q}\right| > \frac{c}{|q|^{1+\tau}} \qquad \forall \quad \frac{p}{q} \in \mathbb{Q} \setminus \{0\}. \tag{7}$$

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

Remark 9. Comparing with Definition 6, it follows that every $\alpha \in \mathcal{D}(\tau)$ has finite Diophantine type $\gamma(\alpha) \leq \tau$. On the other hand, if α has finite Diophantine type, then $\alpha \in \mathcal{D}(\tau)$ for every $\tau > \gamma(\alpha)$.

Remark 10. Let us point out the following properties (see [42, p. 601] for their proofs):

- if $\tau < 1$, the set $\mathcal{D}(\tau)$ is empty;
- if $\tau > 1$ the set $\mathcal{D}(\tau)$ has full Lebesgue measure;
- if $\tau = 1$, then $\mathcal{D}(\tau)$ has Lebesgue measure equal to zero, but it has Hausdorff dimension equal to 1 (hence, it has the cardinality of the continuum).

See also [33, 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

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

it can be proved (see [36, 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 11. *Let α 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:*

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

From the definition of discrepancy, Theorem 11, and the fact that the translation is an isometry, we can deduce the following corollary.

Corollary 12. *Let $x_0 \in [0,1]$, 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 11 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 $N \geq 1$*

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

Proof. It is sufficient to prove that for each x_0 it holds that $D_{N,\alpha}(x_0) \leq 2D_{N,\alpha}(0)$. Indeed, consider $\varepsilon > 0$ and an interval $I = [\alpha, \beta]$ such that

$$D_N(x_1, ..., x_N) - \varepsilon \leq \left| \frac{1}{N} \sum_{1\leq i\leq N} 1_I(x_i) - (\beta - \alpha) \right|.$$

Now consider the translation of I by $-x_0$ (mod. 1):

$$S = \{x \in [0,1] \mid x + x_0 - \lfloor x + x_0 \rfloor \in I\}$$

and the translation of the sequence x_i, which is the sequence $y_i = \alpha i - \lfloor \alpha i \rfloor$. We have that S is composed by at most two intervals $S = I_1 \cup I_2$ with lenghts $m(I_1)$ and $m(I_2)$; moreover

$$\left| \frac{1}{N} \sum_{1\leq i\leq N} 1_I(x_i) - (\beta - \alpha) \right| = \left| \frac{1}{N} \sum_{1\leq i\leq N} 1_{I_1}(y_i) - m(I_1) + \frac{1}{N} \sum_{1\leq i\leq N} 1_{I_2}(y_i) - m(I_2) \right|.$$

Then

$$D_N(x_1, ..., x_N) - \varepsilon \leq 2D_N(y_1, ..., y_N).$$

Since ε is arbitrary, we conclude that $D_{N,\alpha}(x_0) \leq 2D_{N,\alpha}(0)$. □

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 [36, Theorem 5.1, page 143 and Theorem 1.3, page 91]).

Theorem 13. *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 14 (Quantitative statistical stability of Diophantine rotations)**.** *Let $R_\alpha : \mathbb{S}^1 \to \mathbb{S}^1$ be an irrational rotation. Suppose α has finite Diophantine type $\gamma(\alpha)$. Let $\{T_\delta\}_{0\le\delta\le\bar{\delta}}$ be a family of Borel measurable maps of the circle such that*

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

Suppose μ_δ is an invariant measure[3] *of T_δ. 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 15. *Under the assumptions of Theorem 14, let $\{\mu_\delta\}_{0\le\delta\le\bar{\delta}}$ be a family of Borel probability measures on $\mathbb{S}^1$. Then, for every $\varepsilon > 0$*

$$\|m - \frac{1}{n}\sum_{1\le i\le n} L^i\mu_\delta\|_W = O(n^{-\frac{1}{\gamma(\alpha)}+\varepsilon}) \tag{8}$$

uniformly in δ; namely, for every $\varepsilon > 0$, there exist $C = C(\varepsilon) \ge 0$ such that for each δ and $n \ge 1$

$$\|m - \frac{1}{n}\sum_{1\le i\le n} L^i\mu_\delta\|_W \le Cn^{-\frac{1}{\gamma(\alpha)}+\varepsilon}.$$

Proof. Let us fix $\varepsilon > 0$. By Theorem 13 and Corollary 12 we have that there is $C \ge 0$ such that for each Lipschitz function f with Lipschitz constant 1, and for each $x_0 \in \mathbb{S}^1$ we have

$$\left|\frac{1}{n}\sum_{1\le i\le n} f(R_\alpha^i(x_0)) - \int_{[0,1]} f\, dx\right| \le C\, n^{-\frac{1}{\gamma(\alpha)}+\varepsilon} \qquad \forall\, n \ge 1.$$

Let δ_{x_0} be the delta-measure concentrated at a point $x_0 \in \mathbb{S}^1$. By definition of $\|\cdot\|_W$, we conclude that

$$\|m - \frac{1}{n}\sum_{1\le i\le n} L^i\delta_{x_0}\|_W \le Cn^{-\frac{1}{\gamma(\alpha)}+\varepsilon}. \tag{9}$$

Now, as in the proof of Lemma 3, any measure μ_δ can be approximated, arbitary well, in the $\|\cdot\|_W$ norm by a convex combination of delta-measures and we obtain (8) from (9), exactly in the same way as done in the proof of Lemma 3. □

Proof of Theorem 14. Let L_δ be the transfer operator of T_δ. Let us fix $\varepsilon > 0$; without loss of generality we can suppose $\varepsilon < \frac{1}{\gamma(\alpha)}$. By lemma 15 we have that

$$\|m - \frac{1}{n}\sum_{1\le i\le n} L^i\mu_\delta\|_W \le Cn^{-\frac{1}{\gamma(\alpha)}+\varepsilon}.$$

By Lemma 3 we get that for each $n \ge 1$

$$\|\mu_\delta - m\|_W \le \big\|m - \frac{1}{n}\sum_{1\le i\le n} L^i\mu_\delta\big\|_W + \frac{(n-1)}{2}\big\|(L - L_\delta)\mu_\delta\big\|_W. \tag{10}$$

[3] cf. footnote 2.

Hence

$$\begin{aligned}\|\mu_\delta - m\|_W &\le Cn^{-\frac{1}{\gamma(\alpha)}+\varepsilon} + \frac{(n-1)}{2}\|(L-L_\delta)\mu_\delta\|_W \\ &\le Cn^{-\frac{1}{\gamma(\alpha)}+\varepsilon} + \frac{(n-1)}{2}\delta,\end{aligned} \tag{11}$$

where we have used that, since $\sup_{x\in\mathbb{S}^1}|R_\alpha(x)-T_\delta(x)|\le\delta$, then

$$\|(L-L_\delta)\mu_\delta\|_W \le \delta.$$

Since the inequality is true for each $n\ge 1$, we can now consider n minimizing

$$F(n) := Cn^{-\frac{1}{\gamma(\alpha)}+\varepsilon} + \frac{n-1}{2}\delta.$$

The extension to $\mathbb{R}$ of the funcion F is convex and it goes to $+\infty$ both as $x\to 0^+$ and as $x\to+\infty$. Let us denote $a := \frac{1}{\gamma(\alpha)} - \varepsilon > 0$, then $F(x) = Cx^{-a} + \frac{x-1}{2}\delta$. This is minimized at

$$x_* := (2aC)^{\frac{1}{a+1}}\delta^{-\frac{1}{a+1}} := \tilde{c}\,\delta^{-\frac{1}{a+1}}.$$

Consider $n_* = \lfloor x_* \rfloor$ and observe that

$$\begin{aligned}F(n_*) &= \frac{C}{n_*^a} + \frac{n_*-1}{2}\delta \le \frac{C}{n_*^a} + \frac{n_*}{2}\delta = O(\delta^{\frac{a}{a+1}}) \\ F(n_*+1) &= \frac{C}{(n_*+1)^a} + \frac{n_*}{2}\delta \le \frac{C}{n_*^a} + \frac{n_*}{2}\delta = O(\delta^{\frac{a}{a+1}}).\end{aligned}$$

Substituting in (11) we conclude:

$$\begin{aligned}\|\mu_\delta - m\|_W &\le \min\{F(n_*), F(n_*+1)\} = O(\delta^{\frac{a}{a+1}}) \\ &= O\big(\delta^{\frac{1-\varepsilon\gamma(\alpha)}{1+(1-\varepsilon)\gamma(\alpha)}}\big)\end{aligned}$$

proving the statement. □

Remark 16. We remark that, as it follows from the above proof, the constants involved in $O(\delta^\ell)$ in the statement of Theorem 14 only depend on α and ℓ.

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 14 is essentially optimal. We show that for a rotation R_α with rotation number α of Diophantine type $1<\gamma(\alpha)\le+\infty$, there exist perturbations of "size δ", for which the unique physical invariant measure varies in a Hölder way.

More specifically, for any $r\ge 0$ we will construct a sequence $\delta_n\to 0$ and C^∞-maps T_n such that: $\|R_\alpha - T_n\|_{C^r}\le\delta_n$, T_n has a unique physical invariant probability measure μ_n and $\|\mu_n - m\|_W \ge C\delta_n^{\frac{1}{p}}$ for some $C\ge 0$ and $p>1$.

Proposition 17. *Let us consider the rotation $R_\alpha:\mathbb{S}^1\to\mathbb{S}^1$, where α is an irrational number with $1<\gamma(\alpha)\le+\infty$. For each $r\ge 0$ and $\gamma'<\gamma(\alpha)$ there exist a sequence of numbers $\delta_j>0$ and C^∞ diffeomorphisms $T_j:\mathbb{S}^1\to\mathbb{S}^1$ such that $\|T_j - R_\alpha\|_{C^r}\le 2\delta_j$ and*

$$\|m-\mu_j\|_W \ge \frac{1}{2}\delta_j^{\frac{1}{\gamma'+1}}$$

for every $j\in\mathbb{N}$ and for every μ_j invariant measure of T_j.

Proof. We remark that the unique invariant measure for R_α 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

$$|k_j\alpha - p_j| \le \frac{1}{k_j^{\gamma'}} \qquad \Longleftrightarrow \qquad \left|\alpha - \frac{p_j}{k_j}\right| \le \frac{1}{k_j^{\gamma'+1}}.$$

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

Consider $\hat{T}_j$ defined as $\hat{T}_j(x) = R_{\alpha+\delta_j}(x)$; for each $r \ge 0$ we have that $\|\hat{T}_j - R_\alpha\|_{C^r} = |\delta_j|$. 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

$$y_0 := 0,\ y_1 := \delta_j,\ \ldots,\ y_{k_j-1} := 1 - \delta_j,\ y_{k_j} := 0\ (\mathrm{mod.}\, \mathbb{Z}).$$

Consider the measures

$$\mu_j = \frac{1}{k_j} \sum_{0 \le i < k_j} \delta_{y_i},$$

where δ_{y_i} is the delta-measure concentrated at y_i. The measure μ_j is clearly invariant for the map $\hat{T}_j$ and it can be directly computed that

$$\|m - \mu_j\|_W \ge \frac{1}{2k_j}.$$

Observe that $|\delta_j| \le \frac{1}{k_j^{\gamma'+1}}$, hence we get $|\delta_j|^{\frac{1}{\gamma'+1}} \le \frac{1}{k_j}$; then

$$\|m - \mu_j\|_W \ge \frac{1}{2}|\delta_j|^{\frac{1}{\gamma'+1}}.$$

This example can be further improved by perturbing the map $\hat{T}_j = R_{\alpha+\delta_j}$ to a new map T_j in a way that the measure μ_j (supported on the attractor of T_j) and the measure [4] $\mu_j + \frac{k_j}{2}$ (supported on the repeller of T_j) are the only invariant measures of T_j, and μ_j is the unique physical measure for the system. This can be done by making a C^∞ perturbation on $\hat{T}_j = R_{\alpha+\delta_j}$, 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^∞ 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 $D_\delta : \mathbb{S}^1 \to \mathbb{S}^1$, 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 $\Gamma_{\mathrm{rep}} := \{y_i + \frac{1}{2k_j} :\ 0 \le i < k_j\}$ (which is a repeller), to the set $\Gamma_{\mathrm{att}} := \{y_i :\ 0 \le i < k_j\}$ (the attractor). Then, define T_j by composing $R_{\alpha+\delta_j}$ and D_δ, namely

$$T_j(x) := D_{\delta_j}(x + (\delta_j + \alpha)).$$

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

[4]The *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 $\mathbb{S}^1$, where $A - \frac{1}{2k_j}$ is the translation of the set A by $-\frac{1}{2k_j}$.

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 α.

Proposition 18. *Consider a rotation R_α with rotation angle $\alpha := \sum_1^\infty 2^{-2^{2i}}$. Let T_j be its perturbations as constructed in Proposition 17 and let μ_j denote their invariant measures; recall that $\|T_j - R_\alpha\|_{C^k} \le 2|\delta_j| = 2\sum_{n+1}^\infty 2^{-2^{2i}}$.*

Then, there is an observable $\psi : \mathbb{S}^1 \to \mathbb{R}$, with derivative in $L^2(\mathbb{S}^1)$, and $C \ge 0$ such that

$$\left| \int_{\mathbb{S}^1} \psi dm - \int_{\mathbb{S}^1} \psi d\mu_j \right| \ge C\sqrt{\delta_j}.$$

Proof. Comparing the series with a geometric one, we get that

$$\sum_{n+1}^\infty 2^{-2^{2i}} \le 2^{-2^{2(n+1)}+1}.$$

By this, it follows

$$\|2^{2^{2n}}\alpha\| \le 2^{-2^{2(n+1)}+1} = \frac{1}{2(2^{2^{2+2n}})} = \frac{1}{2(2^{2^{2n}})^4}.$$

Since it also holds that $\|2^{2^{2n}}\alpha\| \ge 2^{-2^{2(n+1)}}$, the we conclude that $\gamma(\alpha) = 4$. Following the construction in the proof of Proposition 17, we have that with a perturbation of size less than $2^{-2^{2(n+1)}+1}$ the angles $\alpha_j := \alpha - \delta_j = \sum_1^j 2^{-2^{2i}}$ generate orbits of period $2^{2^{2j}}$. Now let us construct a suitable observable which can "see" the change of the invariant measure under this perturbation. Let us consider

$$\psi(x) := \sum_{i=1}^\infty \frac{1}{(2^{2^{2i}})^2} \cos(2^{2^{2i}} 2\pi x) \tag{12}$$

and debote by $\psi_k(x) := \sum_{i=1}^k \frac{1}{(2^{2^{2i}})^2} \cos(2^{2^{2i}} 2\pi x)$ its truncations. Since for the observable ψ, the i-th Fourier coefficient decreases like i^{-2}, then ψ has derivative in $L^2(\mathbb{S}^1)$. Let $\{x_i\}_i$ be the periodic orbit of 0 for the map R_{α_j} and let $\mu_j := \frac{1}{2^{2^{2i}}} \sum_{i=0}^{\alpha_j - 1} \delta_{x_i}$ be the physical measure supported on it. Since $2^{2^{2j}}$ divides $2^{2^{2(j+1)}}$ then $\sum_{i=1}^{2^{2^{2j}}} \psi_k(x_i) = 0$ for every $k < j$, thus $\int_{\mathbb{S}^1} \psi_{j-1}\, d\mu_j = 0$. Then

$$\begin{aligned} v_j := \quad & \int_{\mathbb{S}^1} \psi \; d\mu_j \ge \frac{1}{(2^{2^{2j}})^2} - \sum_{j+1}^\infty \frac{1}{(2^{2^{2i}})^2} \\ \ge \quad & 2^{-2^{2j+1}} - 2^{-2^{2(j+1)}+1}. \end{aligned}$$

For j big enough

$$2^{-2^{2j+1}} - 2^{-2^{2(j+1)}+1} \ge \frac{1}{2}(2^{-2^{2j}})^2.$$

Summarizing, with a perturbation of size

$$\delta_j = \sum_{j+1}^\infty 2^{-2^{2i}} \le 2 \cdot 2^{-2^{2(j+1)}} = 2^{-2^{2(j+1)}} = 2(2^{-2^{2j}})^4$$

we get a change of average for the observable ψ from $\int_{\mathbb{S}^1} \psi dm = 0$ to $v_n \ge \frac{1}{2}(2^{-2^{2j}})^2$. Therefore, there is $C \ge 0$ such that with a perturbation of size δ_j, we get a change of average for the observable ψ of size bigger than $C\sqrt{\delta_j}$. □

Remark 19. Using in (12) $\frac{1}{(2^{2^{2i}})^{\sigma}}$, for some $\sigma > 2$, instead of $\frac{1}{(2^{2^{2i}})^{2}}$, 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[5]. Suppose hence T_0 has a physical measure μ_0 and let μ_δ be a family of physical measures of T_δ.

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

$$\dot{\mu} := \lim_{\delta\to 0} \frac{\mu_\delta - \mu_0}{\delta} \tag{13}$$

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:

$$\mu_\delta = \mu_0 + \dot{\mu}\delta + o(\delta) \tag{14}$$

which holds in some weaker or stronger sense, depending on which topology the convergence in (13) holds.

We remark that given an observable function $c : X \to \mathbb{R}$, if the convergence in (13) is strong enough with respect to the regularity [6] of c, we get

$$\lim_{t\to 0} \frac{\int_{\mathbb{S}^1} \; c \; d\mu_t - \int_{\mathbb{S}^1} \; c \; d\mu_0}{t} = \int_{\mathbb{S}^1} \; c \; d\dot{\mu} \tag{15}$$

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

[5] An invariant measure μ is said to be *physical* if there is a positive Lesbegue measure set B such that for each continuous observable f

$$\int_{\mathbb{S}^1} f \; d\mu = \lim_{n\to\infty} \frac{f(x) + f(T(x)) + \ldots + f(T^n(x))}{n+1}$$

for each $x \in B$ (see [48]).

[6] For example, L^1 convergence in (13) allows to control the behavior of L^∞ observables in (15), while a weaker convergence in (13), for example in the Wasserstein norm (see definition 1) allows to get information on the behavior of Lipschitz obsevable.

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

Let $\mathrm{Diff}^r_+(\mathbb{S}^1)$ denote the set of orientation preserving homeomorphism of the circle of class C^r with $r \in \mathbb{N} \cup \{+\infty, \omega\}$. Let $\mathrm{rot}(f) \in \mathbb{S}^1$ denote the rotation number of f (see, for example, [33, Section II.2] for more properties on the rotation number).

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 homeomorphim $h: \mathbb{S}^1 \longrightarrow \mathbb{S}^1$ such that the following diagram commutes:

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

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

Remark 20. Observe that if h exists, then it is essentially unique, in the sense that if $h_i : \mathbb{S}^1 \longrightarrow \mathbb{S}^1$, $i = 1, 2$, are homeomorphisms conjugating f to $R_{\mathrm{rot}(f)}$, 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 [33, 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 [18] 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 21 (Denjoy)**.** *Let T be an orientation preserving diffeomorphism of the circle with an irrational rotation number α and such that* $\log(T')$ *has bounded variation. Then there exists a homeomorphism $h : \mathbb{S}^1 \to \mathbb{S}^1$ such that*

$$T \circ h = h \circ R_\alpha.$$

Remark 22. 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 μ 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 [5] and J. Moser [38]. These results are in the perturbative setting and are generally referred to as *KAM theory*. Namely, they consider perturbations of *Diophantine* rotations

$$f_\varepsilon(x) = R_\alpha + \varepsilon u(x, \varepsilon) \tag{16}$$

and prove that, under suitable regularity assumptions on u, there exist $\varepsilon_0 > 0$ (depending on the properties of α and u) and a Cantor set $\mathcal{C} \subset (-\varepsilon_0, \varepsilon_0)$ such that f_ε is conjugated to a $R_{\mathrm{rot}(f_\varepsilon)}$ for every $\varepsilon \in \mathcal{C}$. Observe that the conjugacy does not exist in general for an interval of ε, but only for those values of ε for which the rotation number of f_ε satisfies suitable arithmetic properties (*e.g.*, it is Diophantine). See below for a more precise statement.

Remark 23. Observe that f_ε has not necessarily rotation number α, even if one asks that $u(\cdot,\varepsilon)$ has zero average.

Remark 24. 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 [38] under the assumption that u is sufficiently smooth (the minimal regularity needed was later improved by Rüssmann [41]). 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 mention a few recent articles and surveys, like [16, 21], which contain a more exhaustive list of references therein.

Later, Herman [33] and Yoccoz [46, 47] provided a thorough analysis of the situation in the general (non-perturbative) context. Let us briefly summarize their results (see also [22] for a more complete account).

Theorem 25 (Herman [33], Yoccoz [46, 47])**.**

- *Let* $f \in \mathrm{Diff}^r_+(\mathbb{S}^1)$ *and* $\mathrm{rot}(f) \in \mathcal{D}(\tau)$*. If* $r > \max\{3, 2\tau-1\}$*, then there exists* $h \in \mathrm{Diff}^{r-\tau-\varepsilon}_+(\mathbb{S}^1)$*, for every* $\varepsilon > 0$*, conjugating* f *to* $R_{\mathrm{rot(f)}}$*.*
- *Let* $f \in \mathrm{Diff}^\infty_+(\mathbb{S}^1)$ *and* $\mathrm{rot}(f) \in \mathcal{D}(\tau)$*. Then, there exists* $h \in \mathrm{Diff}^\infty_+(\mathbb{S}^1)$ *conjugating* f *to* $R_{\mathrm{rot(f)}}$*.*
- *Let* $f \in \mathrm{Diff}^\omega_+(\mathbb{S}^1)$ *and* $\mathrm{rot}(f) \in \mathcal{D}(\tau)$*. Then, there exists* $h \in \mathrm{Diff}^\omega_+(\mathbb{S}^1)$ *conjugating* f *to* $R_{\mathrm{rot(f)}}$*.*

Remark 26. 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 $\mathcal{H}$-condition (named in honour of Herman); we refer to [46, 47] 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 [43, Theorem 9.0.4] (cf. also [16, Theorem 2] and [17]).

Theorem 27 (KAM Theorem for circle diffeomorphisms)**.** *Let* $\alpha \in \mathcal{D}(\tau)$*, with* $\tau > 1$ *and let us consider a smooth family of circle diffeomorphisms*

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

with

(i) $u(x,\varepsilon) \in C^\infty(\mathbb{S}^1)$ *for every* $|\varepsilon| < 1$*;*
(ii) the map $\varepsilon \longmapsto u(\cdot,\varepsilon)$ *is* C^∞*;*
(iii) $\int_{\mathbb{S}^1} u(x,\varepsilon)dx = A\varepsilon^m + o(\varepsilon^m)$*, where* $A \neq 0$ *and* $m \geq 0$*.*

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

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

such that

$$\begin{array}{ccc} \mathbb{S}^1 & \xrightarrow{f_\varepsilon} & \mathbb{S}^1 \\ \uparrow h_\varepsilon & & \uparrow h_\varepsilon \\ \mathbb{S}^1 & \xrightarrow{R_{\alpha_\varepsilon}} & \mathbb{S}^1 \end{array} \qquad \Longleftrightarrow \qquad f_\varepsilon \circ h_\varepsilon = h_\varepsilon \circ R_{\alpha_\varepsilon}. \tag{17}$$

Moreover:

- *the maps* $\varepsilon \longmapsto h_\varepsilon$ *and* $\varepsilon \longmapsto \alpha_\varepsilon$ *are* C^∞ *on the Cantor set* $\mathcal{C}$*, in the sense of Whitney;*
- $\alpha_\varepsilon = \alpha + A\varepsilon^{m+1} + o(\varepsilon^{m+1})$.

Remark 28. Observe that f_ε does not have necessarily rotation number α. In particular, the map $rot : \mathrm{Diff}^0_+(\mathbb{S}^1) \longrightarrow \mathbb{S}^1$ is continuous with respect to the C^0-topology (see for example [33, Ch. II, Proposition 2.7])

Remark 29.

(i) Theorem 27 is proved in [43] 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 differentiablity, cf. Theorem 25). In particular, the proof of the asymptotic expansion of α_ε appears on [43, 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 $(-r,r)\cap\mathcal{C}$ has lebesgue measure $\geq Mr^{\frac{1}{m+1}}$ (see [43, 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 [43]).

Let us discuss how to deduce from this result the existence of linear response for the circle diffeomorphisms f_ε.

Theorem 30. *Let* $\alpha \in \mathcal{D}(\tau)$*, with* $\tau > 1$ *and let us consider a family of circle diffeomorphisms obtained by perturbing the rotation* R_α *in the following way:*

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

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

Then, the circle rotation R_α *admits linear response, in the limit as* ε *goes to* 0*, by effect of this family of perturbations.*

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

$$\lim_{\varepsilon\in\mathcal{C},\varepsilon\to 0} \frac{\mu_\varepsilon - m}{\varepsilon} = 2\pi i \sum_{n\in\mathbb{Z}\setminus\{0\}} \left(\frac{n\,\hat{u}(n)}{1-e^{2\pi i n\alpha}}\right) e^{2\pi i n x} \qquad \textit{(in the } L^1\textit{-sense)} \qquad (18)$$

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

Remark 31. 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 [43]).

As we have already observed in Remark 28, the rotation number of f_ε varies continuously with respect to the perturbation, from here the need of taking the limit in (18) 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 ε, hence $\mathcal{C}$ coincides with the whole set of parameters; therefore the limit in (18) can be taken in the classical sense.

Corollary 32. *Under the same hypotheses and notation of Theorem 30, if in addition we have that* $\mathrm{rot}(f_\varepsilon) = \alpha$ *for every* $|\varepsilon| < 1$*, then there exists linear response without any need of restricting to a Cantor set and it is given by*

$$\lim_{\varepsilon\to 0} \frac{\mu_\varepsilon - m}{\varepsilon} = 2\pi i \sum_{n\in\mathbb{Z}\setminus\{0\}} \left(\frac{n\,\hat{u}(n)}{1-e^{2\pi i n\alpha}}\right) e^{2\pi i n x} \qquad \text{(in the } L^1\text{-sense).} \tag{19}$$

Proof. **(Corollary 32).** As we have remarked above, this corollary easily follows from Theorem 30 by observing that $\mathrm{rot}(f_\varepsilon) = \alpha \in \mathcal{D}(\tau)$ for every $|\varepsilon| < 1$, hence $\mathcal{C} \equiv (-1, 1)$. In fact, this follows from [43, Section 9.2, pp. 147-148]: in their notation our parameter ε corresponds to μ and their $a(\mu)$ corresponds to our $\mathrm{rot}(f_\varepsilon)$. In particular, they define the Cantor set as $\mathcal{C}_F = v^{-1}(D_\Upsilon)$ (see [43, p.148]): in our notation this corresponds to the values of $\varepsilon \in (-1, 1)$ for which $\mathrm{rot}(f_\varepsilon)$ belongs to the a certain set of Diophantine numbers that includes α. Since, by hypothesis, $\mathrm{rot}(f_\varepsilon) \equiv \alpha$, it follows that $\mathcal{C} \equiv (-1, 1)$ and, in particular, the limit in (18) is meant in the classical sense. □

Let us now prove Theorem 30.

Proof. **(Theorem 30).** First of all, applying Theorem 27, it follows that for every $\varepsilon \in \mathcal{C}$, the map $f_\varepsilon := R_\alpha + \varepsilon u(x, \varepsilon)$ possesses a unique invariant probability measure given by

$$\mu_\varepsilon = h_{\varepsilon *} m$$

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

$$\frac{d\mu_\varepsilon}{dx}(x) = \frac{1}{\partial_x h_\varepsilon(h_\varepsilon^{-1}(x))}. \tag{20}$$

In fact, if A is a Borel set in $\mathbb{S}^1$, 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 (20) that

$$\begin{aligned} \frac{d\mu_\varepsilon}{dx}(x) &= \frac{1}{\partial_x h_\varepsilon(h_\varepsilon^{-1}(x))} = \frac{1}{1+\varepsilon\partial_x v(h_\varepsilon^{-1}(x),0)+o(\varepsilon)} \\ &= \frac{1}{1+\varepsilon\partial_x v(x,0)+o_{\mathcal{C}}(\varepsilon)} = 1-\varepsilon\partial_x v(x,0)+o_{\mathcal{C}}(\varepsilon), \end{aligned} \tag{21}$$

where $o_{\mathcal{C}}(\varepsilon)$ denotes a term that goes to zero faster than $\varepsilon \in \mathcal{C}$, uniformly in x.

Then the linear response is given by

$$\dot{\mu} = \lim_{\varepsilon\in\mathcal{C},\varepsilon\to 0} \frac{\mu_\varepsilon - \mu_0}{\varepsilon} = \lim_{\varepsilon\in\mathcal{C},\varepsilon\to 0} \frac{\mu_\varepsilon - m}{\varepsilon}$$

which, passing to densities and using (21), corespond to

$$\lim_{\varepsilon\in\mathcal{C},\varepsilon\to 0} \frac{1}{\varepsilon}(1-\varepsilon\partial_x v(x,0)+o_0(\varepsilon)-1) = -\partial_x v(x,0).$$

Giving a formula for the response

$$\frac{d\dot{\mu}}{dx}(x) = -\partial_x v(x,0). \tag{22}$$

Moreover, we can find a more explicit representation formula (the above formula, in fact, is somehow implicit, since v depends on h_ε). Observe that it follows from (17) that $f_\varepsilon \circ h_\varepsilon = h_\varepsilon \circ R_{\alpha_\varepsilon}$:

$$x + \varepsilon v(x,\varepsilon) + \alpha + \varepsilon u(x + \varepsilon v(x,\varepsilon), \varepsilon) = x + \alpha_\varepsilon + \varepsilon v(x + \alpha_\varepsilon, \varepsilon). \tag{23}$$

Recall, from the statement of Theorem 27 that

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

where m and A are defined by (see item (ii) in Theorem 27)

$$< u(\cdot,\varepsilon) >:= \int_{\mathbb{S}^1} u(x,\varepsilon)dx = A\varepsilon^m + o(\varepsilon^m).$$

Hence, expanding equation (23) in terms of ε and equating the terms of order 1, we obtain the following (observe that α_ε will contribute to the first order in ε only if $m = 0$ and, therefore, $A =< u(\cdot,0) >:= \int_{\mathbb{S}^1} u(x,0)dx \neq 0$):

$$v(x+\alpha,0) - v(x,0) = u(x,0) - < u(\cdot,0) > \qquad \forall\, x \in \mathbb{S}^1, \tag{24}$$

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 (24) 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_α:

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

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

$$v(x,0) := \sum_{n\in\mathbb{Z}} \hat{v}(n) e^{2\pi i n x} \qquad \text{and} \qquad u(x,0) := \sum_{n\in\mathbb{Z}} \hat{u}(n) e^{2\pi i n x}.$$

In Fourier terms, (24) becomes:

$$\sum_{n\in\mathbb{Z}} \hat{v}(n) \left(e^{2\pi i n\alpha} - 1\right) e^{2\pi i n x} = \sum_{n\in\mathbb{Z}\setminus\{0\}} \hat{u}(n) e^{2\pi i n x}$$

and therefore for $n \neq 0$

$$\hat{v}(n) = \frac{\hat{u}(n)}{e^{2\pi i n\alpha} - 1};$$

we do not determine $\hat{v}(0)$, as it should be expected, since v is determined by (24) only up to constants.

Substituting in (22), we conclude:

$$\begin{aligned} \frac{d\dot{\mu}}{dx}(x) &= -\partial_x v(x,0) = -2\pi i \sum_{n\in\mathbb{Z}} n\, \hat{v}(n) e^{2\pi i n x} \\ &= 2\pi i \sum_{n\in\mathbb{Z}\setminus\{0\}} \left(\frac{n\, \hat{u}(n)}{1 - e^{2\pi i n\alpha}}\right) e^{2\pi i n x}. \end{aligned}$$

□

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 33. *Let T_0 be an orientation preserving diffeomorphism of the circle with an irrational rotation number α and such that $\log(T')$ has bounded variation (for example f is of class C^2). Let μ_0 be its unique invariant (absolutely continuous) probability measure (see Theorem 21). Let $\{T_\delta\}_{0\leq\delta\leq\overline{\delta}}$ 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 $0 \leq \delta \leq \overline{\delta}$, μ_δ is an invariant measure[7] *of T_δ. 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 21.

Proof of Theorem 33. By Theorem 21 we can conjugate T_0 with the rotation R_α. We apply the same coniugation to T_δ 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

$$\begin{array}{ccccccc} \mathbb{S}^1 & \stackrel{T_0}{\longrightarrow} & \mathbb{S}^1 & \qquad & \mathbb{S}^1 & \stackrel{T_\delta}{\longrightarrow} & \mathbb{S}^1 \\ \downarrow h & & \downarrow h & & \downarrow h & & \downarrow h \\ \mathbb{S}^1 & \stackrel{R_\alpha}{\longrightarrow} & \mathbb{S}^1 & & \mathbb{S}^1 & \stackrel{U_\delta}{\longrightarrow} & \mathbb{S}^1 \end{array} \tag{25}$$

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 $\overline{\mu}_\delta := h_*\mu_\delta$. These measures are invariant for U_δ. Then, by Theorem 2 we get

$$\lim_{\delta\to 0} ||\overline{\mu}_\delta - m||_W = 0.$$

This implies (uniformly approximating any continuous fuction with a sequence of Lipschitz ones) that for each $g \in C^0(\mathbb{S}^1)$

$$\lim_{\delta\to 0} \int_{\mathbb{S}^1} g\ d\overline{\mu}_\delta = \int_{\mathbb{S}^1} g\ dm. \tag{26}$$

Now consider $f \in C^0(\mathbb{S}^1)$ and remark that (using the definition of push-forward of a measure)

$$\begin{aligned} \int_{\mathbb{S}^1} f\ d\mu_\delta &= \int_{\mathbb{S}^1} f \circ h^{-1} \circ h\ d\mu_\delta = \int_{\mathbb{S}^1} f \circ h^{-1}\ d\overline{\mu}_\delta, \\ \int_{\mathbb{S}^1} f\ d\mu_0 &= \int_{\mathbb{S}^1} f \circ h^{-1}\ d\overline{\mu}_0. \end{aligned}$$

[7] cf. footnote 2.

By 26, considering $g = f \circ h^{-1}$ this shows

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

□

Similarly, one can extend the quantitative stability results proved in Theorem 14 to smooth diffeomorphisms of the circle.

Remark 34. 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 $T_0 \in C^r(\mathbb{S}^1)$ with r sufficiently big so that the cojugation h is bi-Lipschitz; compare with Theorem 25.

Theorem 35. *Let T_0 be a C^∞ diffeomorphism of the circle with Diophantine rotation number $\alpha \in \mathcal{D}(\tau)$, for some $\tau > 1$. Let $\{T_\delta\}_{0 \le \delta \le \overline{\delta}}$ be a family of Borel measurable maps of the circle such that*

$$\sup_{x \in \mathbb{S}^1} |T_0(x) - T_\delta(x)| \le \delta.$$

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

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

Proof. By Theorem 25 , there exists $h \in \mathrm{Diff}^\infty_+(\mathbb{S}^1)$ conjugating T_0 with the rotation R_α. We apply the same coniugation to T_δ for each $\delta > 0$ obtaining a family of maps U_δ. The situation is still summarized by (25). Since h is a bilipschitz map we have

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

and there is a $C \ge 1$ such that for any pair of probability measures μ_1, μ_2

$$C^{-1}||\mu_1 - \mu_2||_W \le ||h_*^{-1}\mu_1 - h_*^{-1}\mu_2||_W \le C||\mu_1 - \mu_2||_W$$

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

By Theorem 14 we then get that for each $\ell < \frac{1}{\gamma(\alpha)+1}$ we have:

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

This imply

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

□

Finally, one can also extend the existence of linear response, along the same lines of Theorem 30 and Corollary 32. In fact, as observe in Remark 29 (*iii*), KAM theorem can be extended to sufficiently regular diffeomorphisms of the circle (one can prove it either directly (*e.g.*, [5, 16, 38, 42, 43]), or by combining the result for rotations of the circle, with Theorem 25). Since the proof can be adapted *mutatis mutandis* (of course, leading to a different expression for the linear response), 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 $\mathbb{S}^1$ defined by

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

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

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

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

Given a map $T : \mathbb{S}^1 \to \mathbb{S}^1$ and let $N \in \mathbb{N}$; we define its N-*discretization* $T_N : \mathbb{S}^1 \to \mathbb{S}^1$ 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 pievewise expanding maps (see [28]) and for topologically generic diffeomorphisms of the torus (see [30], [31], [37]).

Since the discretization is a small perturbation in the uniform convergence topology, a direct application of Theorem 33 gives

Corollary 36. *Let T_0 be an orientation preserving diffeomorphism of the circle with an irrational rotation number α and such that $\log(T_0')$ has bounded variation and let $N \ge 1$. Let $T_N = P_N \circ T_0$ be the family of maps given by its $N-$ discretizations. Suppose μ_N is an invariant measure of T_N. Then*

$$\lim_{N\to\infty} \int_{\mathbb{S}^1} f \; d\mu_N = \int_{\mathbb{S}^1} f \; d\mu_0$$

for all $f \in C^0(\mathbb{S}^1)$.

Proof. The statement follows by Theorem 33 noticing that

$$\sup_{x\in\mathbb{S}^1} |T_0(x) - T_N(x)| \le \frac{1}{N}.$$

□

We think this result is very similar to the one shown in Proposition 8.1 of [37]. Comparing this kind of results with the ones in [30], 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 [30] the convergence is proved for a certain sequence of finer and finer discretizations.

As an application of our quantitative stability result (Theorem 14 and 35), 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 37. *Let T_0 be a C^∞ diffeomorphism of the circle with Diophantine rotation number $\alpha \in \mathcal{D}(\tau)$. Let $T_N = P_N \circ T_0$ be the family of its N-discretizations. Suppose μ_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 37 is similar to the one of Corollary 36.

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. Celletti and R. de la Llave for sharing with them their results in [17]. 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). A.S. also acknowledges the support of the MIUR Department of Excellence grant CUP E83C18000100006.

REFERENCES

[1] J. F. Alves, Strong statistical stability of non-uniformly expanding maps, *Nonlinearity*, **17** (2004), 1193–1215.

[2] J. F. Alves and M. Soufi, Statistical stability in chaotic dynamics, *Progress and Challenges in Dyn. Sys. Springer Proc. in Math. & Statistics*, **54** (2013), 7–24.

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

[4] L. Ambrosio, N. Gigli and 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.

[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 and X. Niu, A rigorous computational approach to linear response, *Nonlinearity*, **31** (2018), 1073–1109.

[7] W. Bahsoun, M. Ruziboev and B. Saussol, Linear response for random dynamical systems, *Adv. Math.*, **364** (2020), 107011, 44 pp.

[8] W. Bahsoun and B. Saussol, Linear response in the intermittent family: Differentiation in a weighted C^0-norm, *Discrete Contin. Dyn. Syst.*, **36** (2016), 6657–6668.

[9] W. Bahsoun and S. Vaienti, Metastability of certain intermittent maps, *Nonlinearity*, **25** (2012), 107–124.

[10] V. Baladi, Linear response, or else, *Proceedings of the International Congress of Mathematicians–Seoul 2014*, Kyung Moon Sa, Seoul, **III** (2014), 525–545.

[11] V. Baladi , M. Benedicks and D. Schnellmann, Whitney-Hölder continuity of the SRB measure for transversal families of smooth unimodal maps, *Invent. Math.*, **201** (2015), 773–844.

[12] V. Baladi, T. Kuna and V. Lucarini, Linear and fractional response for the SRB measure of smooth hyperbolic attractors and discontinuous observables, *Nonlinearity*, **30** (2017), 1204-1220.

[13] V. Baladi and D. Smania, Linear response formula for piecewise expanding unimodal maps, *Nonlinearity*, **21** (2008), 677–711. (*Corrigendum, Nonlinearity* **25** (2012), 2203– 2205.)

[14] V. Baladi and D. Smania, Linear response for smooth deformations of generic nonuniformly hyperbolic unimodal maps, *Ann. Sci. Éc. Norm. Sup.*, **45** (2012), 861–926.

[15] V. Baladi and M. Todd, Linear response for intermittent maps, *Comm. Math. Phys.*, **347** (2016), 857–874.

[16] H. W. Broer and M. B. Sevryuk, *KAM Theory: Quasi-periodicity in Dynamical Systems*, Handbook of Dynamical Systems, Vol. **3** (2010), Elsevier/North-Holland, Amsterdam.

[17] R. Calleja, A. Celletti and 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.

[18] A. Denjoy, Sur les courbes définies par les équations différentielles a la surface du tore, *J. Math. Pures et Appl.*, **11** (1932), 333–375.

[19] D. Dolgopyat, On differentiability of SRB states for partially hyperbolic systems, *Invent. Math.*, **155** (2004), 389–449.

[20] D. Dolgopyat, Prelude to a kiss, *Modern Dynamical Systems (ed. M. Brin , B.Hasselblatt and Ya. Pesin)*, (2004), 313–324.

[21] 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, 2014.

[22] H. Eliasson, B. Fayad and R. Krikorian, Jean-Christophe Yoccoz and the theory of circle diffeomorphisms, *La Gazette des Mathématiciens*, Société mathé matiques de France (Jean-Christophe Yoccoz - numéro spécial Gazette), (2018), 55–66.

[23] S. Galatolo, Quantitative statistical stability and speed of convergence to equilibrium for partially hyperbolic skew products, *J. Éc. Pol. Math.*, **5** (2018), 377–405.

[24] S. Galatolo, Quantitative statistical stability and convergence to equilibrium. An application to maps with indifferent fixed points, *Chaos Solitons Fractals*, **103** (2017), 596–601.

[25] S. Galatolo and P. Giulietti, A linear response for dynamical systems with additive noise, *Nonlinearity*, **32** (2019), 2269–2301.

[26] S. Galatolo and J. Sedro, Quadratic response of random and deterministic dynamical systems, *Chaos*, **30** (2020), 023113, 15 pp.

[27] M. Ghil and V. Lucarini, The Physics of Climate Variability and Climate Change, *Rev. Modern Phys.*, **92** (2020), 035002, 77 pp.

[28] P. Góra and A. Boyarsky, Why computers like Lebesgue measure, *Comput. Math. Appl.*, **16** (1988), 321–329.

[29] S. Gouëzel and C. Liverani, Banach spaces adapted to Anosov systems, *Ergodic Theory and Dynamical Systems*, 26 (2006), 189–217.

[30] P.-A. Guihéneuf, Physical measures of discretizations of generic diffeomorphisms, *Ergodic Theory Dynam. Systems*, **38** (2018), 1422–1458.

[31] P.-A. Guihéneuf, *Discrétisations Spatiales de Systémes Dynamiques Génériques,* PhD Thesis, Université Paris-Sud, 2015.

[32] M. Hairer and A. J. Majda, A simple framework to justify linear response theory, *Nonlinearity*, **23** (2010), 909–922.

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

[34] G. Keller and C. Liverani, Stability of the spectrum for transfer operators, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, **28** (1999), 141–152.

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

[36] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Pure and Applied Mathematics.Wiley-Interscience, 1974.

[37] T. Miernowski, Discrétisations des homéomorphismes du cercle, *Erg. Th. Dyn. Sys.*, **26** (2006), 1867–1903.

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

[39] M. Pollicott and P. Vytnova, Linear response and periodic points, *Nonlinearity*, 29 (2016), 3047–3066.

[40] D. Ruelle, Differentiation of SRB states, *Comm. Math. Phys.*, **187** (1997), 227–241.

[41] H. Rüssmann, Kleine Nenner. I. Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes, (German) *Nachr Akad Wiss., Göttingen Math-Phys KI II*, **1970** (1970), 67–105.

[42] H. Rüssmann, On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus, In: Moser J. (eds) *Dynamical Systems, Theory and Applications*. Lecture Notes in Physics, vol **38**. Springer, Berlin, Heidelberg, (1975), 598–624.

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

[44] 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).

[45] C. L. Wormell and G. A. Gottwald, On the validity of linear response theory in high-dimensional deterministic dynamical systems, *J. Stat. Phys.*, **172** (2018), 1479–1498.

[46] J.-C. Yoccoz, Conjugaison différentiable des difféomorphismes du cercle dont le nombre de rotation vérifie une condition diophantienne, *Ann. Sci. Ecole Norm. Sup. (4)*, **17** (1984), 333–359.

[47] J.-C. Yoccoz, Analytic linearization of circle diffeomorphisms, *Dynamical Systems and Small Divisors* (Cetraro, 1998), 125–173, Lecture Notes in Math., **1784**, Fond. CIME/CIME Found. Subser., Springer, Berlin, (2002).

[48] L.-S. Young, What are SRB measures, and which dynamical systems have them?, *J. Stat. Phys.*, **108** (2002), 733–754.

[49] Z. Zhang, On the smooth dependence of SRB measures for partially hyperbolic systems, *Comm. Math. Phys.*, **358** (2018), 45–79.

Received March 2021; 1st revision July 2021; 2nd revision July 2021; early access September 2021.

E-mail address: stefano.galatolo@unipi.it
E-mail address: sorrentino@mat.uniroma2.it