THE RANDOM ARNOLD CONJECTURE: A NEW PROBABILISTIC CONLEY-ZEHNDER THEORY FOR SYMPLECTIC MAPS

ÁLVARO PELAYO FRAYDOUN REZAKHANLOU

Abstract. We take the first steps to develop Conley-Zehnder Theory, as conjectured by Arnold, in the world of probability. As far as we know, this paper provides the first probabilistic theorems about the density of fixed points of symplectic twist maps in dimensions greater than 2. In particular we will show that, when the analogue conditions to classical Conley-Zehnder theory hold, quasiperiodic symplectic twist maps have infinitely many fixed points almost surely. The paper contains also a number of theorems which go well beyond the quasiperidic case.

1. Introduction

Conley-Zehnder theory, as conjectured by Arnold, is one of the the great achievements at the intersection of symplectic geometry and Hamiltonian dynamics in the past few decades.

The motivation for these works can be traced back to Poincaré and later to the first developments in symplectic topology.

In the present paper we take the first steps to develop Conley-Zehnder Theory in a probabilistic setting. One of our main theorems says that a quasiperiodic twist symplectic map has infinitely many fixed points almost surely, provided the analogue conditions to those imposed by Conley and Zehnder in their famous theorem for tori, hold.

A main tool of the paper is the Ergodic Theorem, which allows us to control the behavior of random symplectic maps in analogy with how topological assumptions such as compactness are used in Conley-Zehnder theory. We also use the Ergodic Theorem to evaluate the density of fixed points.

1.1. **Poincaré's theorem on area preserving maps: from classical to random.** The work of Henri Poincaré in classical mechanics [Po93] led him to the famous Poincaré-Birkhoff Theorem [Po12] concerning fixed points of area preserving twist maps of an annulus, which he stated in 1912. It was Birkhoff [Bi13, Bi26] who finally proved the result in 1925.

The result essentially says that such a map always has at least two fixed points, and these points are genuinely different.

This result motivated us to pursue similar statements in the context of probability theory, and we took the first steps to do this in our initial paper in the subject [PR] published in 2018 (see Section 7 for a very brief account of the main results of this paper), where we proved that there is a statement, similar to the Poincaré-Birkhoff Theorem, for area preserving maps which are random.

In [PR] we rely heavily on "finite dimensional" methods, notably the theory of generating functions, which allows one to reduce infinite dimensional proofs to the finite dimensional case. More concretely, we used the ideas of Chaperon [Ch84, Ch84b, Ch89] and Viterbo [Vi11].

There is an essential mathematical difference between the classical and the random versions of this result, and which is also at the heart of the proofs: while the classical result is more topological because it is established in setting of compactness, the random result is more analytic because it only makes sense in the non-compact world. Hence, in the proof we gave, we used mostly analysis.

From the point of view of what statements to expect, in the random setting one expects to have *infinitely many fixed points*, and to come in families which are genuinely different too.

1.2. **Moving to higher dimensions: the Arnold Conjecture.** From the point of view of symplectic geometry, a result in dimension 2, while interesting is not entirely satisfactory. It was Arnold who had the great insight of formulating an analogue of the result by Poincaré-Birkhoff in higher dimensions. He saw, that one should consider "symplectic maps" instead of "volume preserving maps", and formulated the famous Arnold Conjecture. This conjecture has generated an immense amount of research in symplectic geometry in the past few decades.

Essentially the conjecture says that if (M, ω) is a compact symplectic manifold, then any time periodic Hamiltonian diffeomorphism has at least as many fixed points as a smooth function has critical points.

1.3. **A breakthrough by Conley and Zehnder, and beyond.** Conley and Zehnder [CZ83] made the first breakthrough on the Arnold conjecture, proving it for the $2d$-dimensional torus; more precisely they proved that any smooth symplectic map of the $2d$-dimensional torus that is isotopic to identity has at least $2d+1$ many fixed points.

The work by Conley and Zehnder was followed by major works by Floer where he developed the ideas of what now is known as Floer theory [Fl88, Fl89, Fl89b, Fl91], and works of many others, including Hofer, Hofer-Salamon, Liu-Tian, Ono, and Weinstein [Ho85, HS95, LT98, On95, We86].

1.4. **Goal of this paper: towards a probabilistic Arnold Conjecture and a proof of the random Conley-Zehnder Theorem.** Our goal in this paper is to take the first steps to understand the Arnold Conjecture in the world of probability, by establishing the Conley-Zehnder theorem for random symplectic twist maps.

At this time much of the technical machinery that is needed to remove the assumption "twist" is not yet available, nonetheless we believe that a much more general result will be possible in the future, and as such we state it as a meta-goal in Section 2.

The statements and proofs we present in the paper we believe are quite new, in the sense that we are not aware of results in this direction, beyond what we did in dimension 2 in our paper from five years ago [PR]. Indeed [PR] concerns area preserving maps in dimension 2, so it is not in that sense a very "symplectic" paper, while in the current paper we treat any dimension, in the spirit of the original Arnold Conjecture and Conley-Zehnder theory; so the context of our current paper is indeed, symplectic.

1.5. **Novelty of the paper: statements and proof techniques.** Our point of view in our previous paper [PR] in dimension 2 was mainly the classical theory of generating functions, more specifically Chaperon's viewpoint [Ch84, Ch84b, Ch89]. This point of view has been further advocated by Viterbo [Vi11].

We believe that the ideas of the present paper — both involved in the statements and in the proofs — are new, and are developed from combining ergodic and symplectic methods. Indeed, as far as we know, our paper provides the first probabilistic theorems about the density of fixed points of symplectic twist maps in dimensions greater than 2.

We recommend Hofer-Zehnder [HZ94] and Polterovich [Pol01] for treatments of different aspects of symplectic topology. We refer to Golé [Go01] for a treatment of symplectic twist maps and to Adler-Taylor [AT07, AT09] for treatments of certain geometric aspects of randomness.

In particular we refer to [AT07, AT09, AW09] for thorough discussions of Kac-Rice type formulas for level sets of Gaussian random fields.

1.6. **Structure and main achievements of the paper.** In Section 2 we formulate the main goal of the paper (Meta-Goal and Stochastic Analogue of Conley-Zehnder Theorem) and formulate two of our main results: Theorem 2.5 and Theorem 2.6. These two results concern quasiperiodic symplectic maps and are simpler to state, but the paper goes well beyond this case, so in this section we also announce the other main results of the paper: Theorem 3.4, Proposition 4.1, Theorem 4.2 and Theorem 5.1.

In Section 3 we study the existence of generating functions. Notably, in Theorem 3.4 we prove that the "generating function" is stationary. This allows us in Proposition 4.1 to have an almost sure candidate for the density of fixed points, meaning by "density" the number of fixed points in a box of side 2ℓ divided by the volume of the box.

This poses the problem of deciding whether the density is positive, which we achieve in Section 4 and Section 5 by deriving an explicit formula for this density in two cases: Theorem 4.2 and Theorem 5.1. One of these cases has to do with quasiperidic symplectic twist diffeomorphisms which would lead to Theorem 2.5 and Theorem 2.6. We would prefer to give a rather informal statements of these theorems in this section, and provide detailed and precise versions of these theorems later in the paper as Theorems 5.1 and 6.2.

Section 6 is devoted to the properties of time-one map of stationary Hamiltonian ODEs. These properties would allow us to deduce Theorem 2.6 (equivalently Theorem 6.2) from Theorem 2.5 (equivalently Theorem 5.1).

Finally, in the appendix (Section 7) we review the 2-dimensional case, that is, the random Poincaré-Birkhoff Theorem. This case is much simpler to discuss and relies on more standard tools, so we believe that it can serve as a warm up for the results that of the present paper.

Acknowledgments. Both authors have been generously funded by a BBVA (Bank Bilbao Vizcaya Argentaria) Foundation Grant for Scientific Research Projects with project title *From Integrability to Randomness in Symplectic and Quantum Geometry (FITRISAQG).*

During the preparation for some parts of this work, FR was hosted by Max Fathi at Université Paris Cité during June of 2023. He is thankful for the invitation and hospitality of UP Cité. During the final stages of this work (July 2023), AP was visiting the University of Cantabria and UIMP (Universidad International Menéndez Pelayo), and he is thankful to these institutions for their hospitality.

AP is also thankful to the Dean of the School of Mathematics of the Complutense University of Madrid Antonio Brú and the Chair of the Department of Algebra, Geometry and Topology Rutwig Campoamor for their support and excellent resources he is being provided with to carry out the aforementioned BBVA project.

2. Meta-Goal, main results on quasiperiodic symplectic twists, and announcements of results beyond the quasiperiodic case

We very briefly review the classical Conley-Zehnder theory and then establish its probabilistic analogue, which is our main result.

2.1. **Conley-Zehnder Theory.**

We are interested in extending to the stochastic setting the following classical result of Conley and Zehnder.

Theorem 2.1 (Conley-Zehnder [CZ83]). *Every smooth time* 1*-periodic Hamiltonian vector field on the standard torus* $\mathbb{T}^{2d}$ *has at least* $2d+1$ *contractible periodic orbits.*

Alternatively, if we write φ for the time one map of the flow of such Hamiltonian vector field, then φ is a symplectic map that has at least $2n+1$ fixed points. Writing $\Phi : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ for the lift of φ, we have a symplectic map that has at least $2d+1$ fixed points in any box of side length 1. We may state Conley–Zehnder Theorem in terms of Φ.

Theorem 2.2 (Conley-Zehnder [CZ83]). *Let* $\Phi : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ *be a symplectic diffeomorphism that is homologous to identity. Additionally assume that*

$$\omega(x) := \Phi(x) - x$$

is 1*-periodic, and*

$$\int_{[0,1]^{2d}} \omega(x)\,\mathrm{d}x = 0.$$

Then Φ *has at least* $2d+1$ *fixed points in the set* $[0,1)^{2d}$*.*

Chaperon [Ch84] carried out a proof of the Conley-Zehnder Theorem using generating functions, and the present paper pushes these ideas further.

2.2. **Random Conley-Zehnder Theory: quasiperiodic case (the simplest beyond periodic).**

For the stochastic analogue of [CZ83], we take a symplectic diffeomorphism $\Phi : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ of the form

$$\Phi(x) = x + \omega(x),$$

where ω is selected randomly, and the map

$$x \mapsto \omega(x)$$

is a stationary process with respect to the $2d$-dimensional translation

$$\theta_a \omega(x) = \omega(x+a).$$

Our typical result would assert that even if ω is not periodic, then generically the corresponding Φ would have infinitely many fixed points.

In fact we will use probabilistic means to select a generic Φ. To explain this, let us set

$$\mathcal{S} := \Big\{ \Phi\colon \mathbb{R}^{2d} \to \mathbb{R}^{2d} \,:\, \Phi \text{ symplectic diffeomorphism}, \, \omega := \Phi - \mathrm{id} \text{ bounded} \Big\},$$

and define

$$\theta'_a \Phi(x) = \Phi(x+a) - a$$

so that

$$\theta'_a \Phi(x) - x = \theta_a \omega(x).$$

We equip $\mathcal{S}$ with the topology of C^1 norm and consider the σ-algebra $\mathcal{B}$ of Borel subsets of $\mathcal{S}$.

Definition 2.3. We say that a probability measure $\mathcal{P}$ on $\mathcal{S}$ is *θ'-invariant and ergodic* if the following conditions are true:

(i) For every $A \in \mathcal{B}$ we have that $\mathcal{P}\big(\theta'_a A\big) = \mathcal{P}(A)$.
(ii) If there exists $A \in \mathcal{B}$ such that $\theta'_a A = A$ for all $a \in \mathbb{R}^{2d}$, then $\mathcal{P}(A) \in \{0, 1\}$.

In the same vein, we can talk about a probability measure $\mathbb{P}$ that is *θ-invariant and ergodic*

Example 2.4 *(Almost periodic-twists)* To ease the notation, we write n for $2d$. Given a continuous function $\bar{\Phi} : \mathbb{R}^n \to \mathbb{R}^n$, let us assume that the set $\Gamma = \{\theta'_a \bar{\Phi}\}$ is precompact with respect to the topology of uniform convergence. We write Γ' for the topological closure of Γ. By the classical theory of almost periodic functions, the set Γ' can be turned into a compact topological group and for $\mathcal{P}$, we may choose a normalized *Haar* measure on Γ'. We say $\bar{\Phi}$ is *quasiperiodic* if the group Γ' is finite dimensional (and therefore isomorphic to a torus). More concretely, let $\bar{\omega} : \mathbb{R}^N \to \mathbb{R}$ be a 1-periodic C^2 function, and let $A \in \mathbb{R}^{N \times n}$ be a matrix. Then the map

$$\bar{\Phi}(x) = x + \bar{\omega}(Ax),$$

is quasiperiodic.

Meta-Goal (Stochastic Analogue of Conley-Zehnder Theorem): *Let $\mathcal{P}$ be a θ'-invariant ergodic measure on $\mathcal{S}$ such that*

$$\int_{\mathcal{S}} \Phi(0)\ \mathcal{P}(\mathrm{d}\Phi) = 0.$$

Assume also that Φ is homologous to the identity map, $\mathcal{P}$-almost surely. Then Φ has infinitely many fixed point $\mathcal{P}$-almost surely.

We establish this Meta-Goal in an important case: We say that

$$\Phi(q,p) = \big(Q(q,p), P(q,p)\big)$$

is *twist* if for every $p \in \mathbb{R}^d$, the map $q \mapsto Q(q,p)$ is a diffeomorphism of $\mathbb{R}^d$.

Theorem 2.5. *The Meta-Goal holds when Φ is a quasiperiodic symplectic twist.*

We refer to section 5.1 for the regularity of Φ and a more detailed statement of Theorem 2.5.

One natural way of producing such a stochastic symplectic map is by using time one map of a Hamiltonian ODE for which the Hamiltonian function is a stationary process with respect to the translation.

To prepare for the statement of our results, let us write $\mathcal{H}_0$ for the set of C^2 functions

$$H : \mathbb{R}^{2d} \times \mathbb{R} \to \mathbb{R},$$

such that:

(i) $H(x, t+1) = H(x, t)$ for all $(x, t) \in \mathbb{R}^{2d} \times \mathbb{R}$,
(ii) ∇H is uniformly bounded.

We also define translation of $H \in \mathcal{H}_0$ by

$$\theta_a H(x,t) = H(x+a,t).$$

We equip $\mathcal{H}_0$ with the topology of C^2 norm and consider the σ-algebra $\mathcal{B}$ of Borel subsets of $\mathcal{H}_0$.

We write

$$X_H = J\nabla H(x,t)$$

for the Hamiltonian vector field associated with the Hamiltonian function H, and ϕ_t^H for the flow of X_H. The following is a stochastic analogue of Theorem 2.1. The regularity of H in the statement will be made precise later in the paper (Theorem 6.2).

Theorem 2.6. *Let $\mathbb{P}$ be a θ-invariant and ergodic probability measure on $\mathcal{B}$. If H is sufficiently small, then the Hamiltonian vector field X_H has infinitely many 1-periodic orbits $\mathbb{P}$-almost surely.*

2.3. **Random Conley-Zehnder Theory: well-beyond the quasiperiodic case.** It is important to note that our results go well beyond the quasiperidic case covered in the theorems presented in this section, as one can see from the theorems proven in Sections 3-5.

Indeed, the main result of Section 3, and probably the hardest and most substantial result of this paper, is Theorem 3.4 which reduces the problem of counting fixed points to counting the critical points of a stationary process.

In two cases we provide an explicit formula for the density of fixed points, which is achieved in Section 4 (Theorem 4.2) in one case (in the case that some random variable has a density) and then the second case of Section 5 is the quasiperiodic case (Theorem 5.1); this is the only section where quasiperiodic case appears at all, but we stated results earlier for this case because it is simpler to formulate.

2.4. **Examples of stationary Hamiltonian functions.** It is important to note that the quasiperiodic case is **the least random of all** and in that sense the least interesting from the point of view of stochastic processes. It is also the simplest case to deal with. We now describe some examples of stationary probability measures on $\mathcal{H}_0$. We equip $\mathcal{H}_0$ with the topology of uniform convergence.

Example 2.7 *(Periodic Hamiltonian Functions)* As the simplest example, take any $H_0(x,t)$ in $\mathcal{H}_0$, that is 1-periodic in x-variable, and set

$$O(H_0) = \left\{\theta_a H_0 \ : \ a \in \mathbb{R}^{2d}\right\}. \tag{2.1}$$

Since H_0 is a 1-periodic function, the set $O(H_0)$ is homeomorphic to $\mathbb{T}^{2d}$. Under this homeomorphism, the translation θ becomes the standard translation Θ on $\mathbb{T}^{2d}$. The only θ-invariant probability measure $\mathbb{P}$ on $O(H_0)$ is the push forward of the Lebesgue measure on $\mathbb{T}^{2d}$ under the map $a \mapsto \theta_a H_,$.

Example 2.8 *(Quasiperiodic Hamiltonian Functions)* Given $N \geqslant n$, pick a C^2 function $K_0(\omega,t)$,

$$K_0 : \mathbb{T}^N \times \mathbb{R} \to \mathbb{R},$$

that is 1-periodic in t. Pick a matrix $A \in \mathbb{R}^{N\times n}$ that satisfies the following condition:

$$mA = 0, \quad m \in \mathbb{Z}^N \quad \Rightarrow \quad m = 0. \tag{2.2}$$

Let

$$H_0(x,t) = K_0(Ax,t),$$

and define $O(H_0)$ as in (2.1). Note that if $N > 2d$, the set $O(H_0)$ is not closed. However, the condition (2.2) guarantees that its topological closure $\overline{O(H_0)}$ consists of functions of the form

$$H(x,t,\omega) := K_0(\omega + Ax, t),$$

with $\omega \in \mathbb{T}^N$. (Here we regard $\mathbb{T}^N$ as $[0,1]^N$ with $0 = 1$, and $\omega + Ax$ is a Mod 1 summation.) Assume that $\mathbb{P}$ is concentrated on the set $\overline{O(H_0)}$. Again, since $\mathbb{P}$ is θ-invariant, the pull-back of $\mathbb{P}$ with respect to the transformation $\omega \in \mathbb{T}^N \mapsto H(\cdot,\cdot,\omega)$ can only be the Lebesgue measure on $\mathbb{T}^N$. Note that our main result Theorem 2.6 only guarantees the existence of 1-periodic orbits for $H(\cdot,\cdot,\omega)$, for $\mathbb{P}$-almost all choices of ω.

Example 2.9 *(Almost periodic Hamiltonian Functions)* Given a function $H_0 \in \mathcal{H}$, let us assume that the corresponding $O(H_0)$ is precompact with respect to the topology of uniform convergence. By the classical theory of almost periodic functions, the set $\overline{O(H_0)}$ can be turned to a compact topological group and for $\mathbb{P}$, we may choose a normalized *Haar* measure on $\overline{\Omega(H_0)}$.

Example 2.10 *(Lorenz gas type models)* Let us write Ω_0 for the set of discrete subsets of $\mathbb{R}^d$. We also write $\mathbb{Q}_0$ for the law of a Poisson point process of intensity one on Ω_0. We set $\Omega = \Omega_0 \times \mathbb{T}^d$, and $\mathbb{Q} = \mathbb{Q}_0 \times \lambda$, where λ denotes the Lebesgue measure of $\mathbb{T}^d$. On Ω_0, we have a natural translation that is denoted by τ: For $\omega_0 = \{q_i : \ i \in I\}$, we define

$$\tau_q \omega_0 = \{q_i - q : \ i \in I\}.$$

As before, let us write Θ for the translations on the torus $\mathbb{T}^d$. We define a translation $\hat{\theta}$ on Ω by

$$\hat{\theta}_{(q,p)}(\omega_0, a) = \big(\tau_q \omega_0, \Theta_p a\big).$$

The measure $\mathbb{Q}$ is $\hat{\theta}$ invariant and ergodic. Pick a C^2 function $K^0(p,t)$ that is 1-periodic in all the coordinates of (p,t), and a C^2 function $V(q,t)$ that is of compact support in q and 1-periodic in t. Given a realization of $\omega = (\omega_0, a)$, with $\omega_0 = \{q_i : \ i \in I\}$, we define a Hamiltonian function

$$H(q,p,t,\omega) := K^0(p + a, t) + \sum_{i \in I} V(q - q_i, t).$$

The map $\omega \mapsto H(\cdot,\cdot,\cdot,\omega)$ pushes forward the probability measure $\mathbb{Q}$ to a probability measure $\mathbb{P}$ that is θ-invariant and ergodic. Note that if $x(t) = (q(t), p(t))$ solves the corresponding Hamiltonian ODE $\dot{x} = J\nabla H(x,t)$, then the speed $|\dot{q}(t)|$ is bounded by $\|K^0_p\|_{\mathrm{C}^0}$. We also have a bound on $|\dot{p}(t)|, \ t \in [0,T]$ in terms of the number of q_i in a ball $\mathrm{B}_{r(T)}(q(0))$, with a radius $r(T)$ that depends on T only. From this we deduce that the corresponding Hamiltonian ODE is well-defined even though $\mathrm{D}^2\mathrm{H}$ is not uniformly bounded. When K^0 is instead of the form $K^0(p) = |p|^2/2$, the corresponding Hamiltonian ODE is known as a *Lorenz gas* with the following interpretation: $x(t)$ is the state of a particle at time t that is interacting via a potential V with immobile particles at random locations q_i's.

Example 2.11 Let us write Ω for the set of discrete subsets of $\mathbb{R}^{2d}$. We also write $\mathbb{Q}$ for the law of a Poisson point process of intensity one on Ω. On Ω, we have a natural translation that is denoted by θ: For $\omega = \{x_i = (q_i, p_i) : \ i \in I\}$, we define

$$\theta_x \omega = \{x_i - x : \ i \in I\}.$$

The measure $\mathbb{Q}$ is θ invariant and ergodic. Pick a C^2 function $K^0(q, p, t)$ that is of compact support in $x = (q, p)$, and 1-periodic in t. Given a realization of $\omega = \{x_i : \ i \in I\}$, we define a Hamiltonian function

$$H(x, t, \omega) := \sum_{i \in I} K^0(x - x_i, t).$$

Observe that this sum is finite $\mathbb{Q}$-almost surely, because K^0 is of compact support, and ω is discrete. The map $\omega \mapsto H(\cdot, \cdot, \omega)$ pushes forward the probability measure $\mathbb{Q}$ to a probability measure $\mathbb{P}$ that is θ-invariant and ergodic. We remark that the Hamiltonian vector field

$$X(x, t, \omega) = J \nabla H(x, t, \omega)$$

is not a Lipschitz map. However we conjecture that one should be able to construct a nice flow for X, $\mathbb{P}$-almost surely.

Remark 2.12 As we mentioned earlier, our main results in Section 5, namely Theorems 5.1 and 5.2, offer an explicit expression (the formula (5.5)) for the density of 1-periodic orbits in the setting of Example 2.8. This formula is based on the classical Coarea Formula. We speculate two possible extensions of the work of this article that would allow us to study the other examples we formulated above:

- We expect an analogue of formula (5.5) to hold for the density of periodic orbits in the setting of almost periodic Hamiltonian ODEs (Example 2.8). To derive such a formula, we need an analogue of coarea formula for the Haar measure of a topological group that can be regarded as an infinite dimensional torus $\mathbb{T}^\infty$.
- We also conjecture that our Theorem 4.2 is applicable to the model we described in Example 2.11. As we mentioned before, Kac-Rice type formulas are often stated and verified for Gaussian processes. Because of this, we can build Hamiltonian functions from Gaussian processes to produce examples for which our Theorem 4.2 is applicable.

2.5. **Abstract setting and Poisson Structure.** In an equivalent formulation of our results, we start from a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and a group of measurable maps

$$\big(\theta_a : \ a \in \mathbb{R}^n\big)$$

with

$$\theta_{a+b} = \theta_a \circ \theta_b,$$

such that $\mathbb{P}$ is θ-invariant and ergodic. In our probabilistic setting, $(\Omega, \mathcal{F}, \theta, \mathbb{P})$, plays the role of a symplectic manifold. Needless to say that we have no tangent bundle to make sense of symplectic forms on Ω. However, it is possible to make sense of a Poisson structure on Ω that is inherited from the standard Poisson structure of $\mathbb{R}^{2d}$ via the translation θ.

In order to explain this, we first define an (unbounded) operator $\boldsymbol{\nabla}$ that is acting on measurable functions on Ω. For the domain of the definition of this operator we write $\mathfrak{H}^1(\mathbb{P})$ (see also Definition 3.1 below). It consists of functions

$$f : \Omega \to \mathbb{R}$$

such that

- $f \in \mathrm{L}^2(\mathbb{R})$, and that the map $x \mapsto f(\theta_x\omega)$ is differentiable at $x = 0$ for $\mathbb{P}$-almost ω. This derivative is denoted by $\boldsymbol{\nabla} f(\omega)$.
- The function $\boldsymbol{\nabla} f$ is in $\mathrm{L}^2(\mathbb{P})$.

Given a measurable $K : \Omega \times \mathbb{R} \to \mathbb{R}$, we define the corresponding Hamiltonian function H by

$$H(x, t, \omega) = K(\theta_x\omega, t).$$

When $K(\cdot, t) \in \mathfrak{H}^1(\mathbb{P})$, and is continuous in time, we can talk about the corresponding *Hamiltonian vector field*

$$X_K(\omega, t) = J\boldsymbol{\nabla} K(\omega, t).$$

Observe that when $J\nabla H(x, t, \omega)$ is C^1, then we can talk about its flow $\phi_t^{H(\cdot,\omega)}(x)$. Using this, we can define a flow $\boldsymbol{\phi}_t^K$ on ω in the following manner:

$$\boldsymbol{\phi}_t^K(\omega) = \theta_{x(t,\omega)}\omega, \quad \text{where} \quad x(t, \omega) := \phi_t^{H(\cdot,\omega)}(0).$$

In some sense, $\boldsymbol{\phi}_t^K$ is the flow of the Hamiltonian (or rather Poissonian) vector field X_K. In order to explain this, we first construct the *Poisson structure*

$$\boldsymbol{\{}\cdot, \cdot\boldsymbol{\}} : \mathfrak{H}^1(\mathbb{P}) \times \mathfrak{H}^1(\mathbb{P}) \to \mathrm{L}^1(\mathbb{P})$$

on $(\Omega, \mathcal{F}, \theta, \mathbb{P})$ given by

$$\boldsymbol{\{}f, g\boldsymbol{\}} = J(\boldsymbol{\nabla} f) \cdot (\boldsymbol{\nabla} g).$$

It is degenerate (expect when $\Omega = \mathbb{T}^{2d}$) because it is induced from the $2d$-dimensional symplectic structure of $\mathbb{R}^{2d}$ on the possibly infinite dimensional space Ω. Observe that for a function $f \in \mathfrak{H}^1(\mathbb{P})$,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{dt}} f\big(\boldsymbol{\phi}_t^K(\omega)\big) &= \frac{\mathrm{d}}{\mathrm{dt}} f\big(\theta_{x(t,\omega)}\omega\big) = (\boldsymbol{\nabla} f)\big(\theta_{x(t,\omega)}\omega\big) \cdot \dot{x}(t, \omega) \\
&= (\boldsymbol{\nabla} f)\big(\theta_{x(t,\omega)}\omega\big) \cdot J\nabla H(x(t, \omega), t, \omega) \\
&= (\boldsymbol{\nabla} f)\big(\theta_{x(t,\omega)}\omega\big) \cdot J\boldsymbol{\nabla} K(\theta_{x(t,\omega)}\omega, t) \\
&= \boldsymbol{\{}K(\cdot, t), f\boldsymbol{\}}\big(\boldsymbol{\phi}_t^K(\omega)\big).
\end{aligned}$$

Example 2.13 In the quasi periodic setting, $\Omega = \mathbb{T}^N = [0, 1]^N,\ 0 = 1$, $\mathbb{P}$ is the Lebesgue measure, and $\theta_x\omega = \omega + Ax \pmod 1$. In this case,

$$\big(\Omega, \boldsymbol{\{}, \cdot, \cdot\boldsymbol{\}}\big)$$

is a Poisson manifold, with

$$\boldsymbol{\{}f, g\boldsymbol{\}} = \big(AJA^*\big)(\nabla f) \cdot (\nabla g).$$

Given $K_0 : \mathbb{T}^N \times \mathbb{R} \to \mathbb{R}$, the corresponding Hamiltonian ODE vector field is $\big(AJA^*\big)(\nabla K)(\omega, t)$. We refer to the proof of Proposition 6.1(vii) below for more details.

3. Existence of stationary generating functions

In this section we study the existence and regularity of generating functions associated with θ' stationary twist symplectic maps. Our results require a Sobolev-type regularity of the symplectic twist diffeomorphism Φ that depends on the choice of the stationary measure $\mathcal{P}$ or $\mathbb{P}$. The corresponding Sobolev spaces will be defined in the next definition. To ease the notation, we write n for $2d$.

Definition 3.1.

(i) Given a θ-invariant probability measure $\mathbb{P}$ on a measure space Ω, we define a group of unitary operators

$$\mathcal{T}_x : \mathrm{L}^2(\mathbb{P}) \to \mathrm{L}^2(\mathbb{P}), \qquad \mathcal{T}_x f(\omega) = f(\theta_x \omega).$$

The inner product of the corresponding $\mathrm{L}^2(\mathbb{P})$ is denoted by $\langle \cdot, \cdot \rangle$. We also write $\mathbb{E}$ for the expected value with respect to $\mathbb{P}$:

$$\mathbb{E}\, f = \int_\Omega f \, \mathrm{d}\mathbb{P}.$$

The *infinitesimal generator* of the group $\mathcal{T}$ is denoted by $\boldsymbol{\nabla}$,

$$\begin{aligned} \boldsymbol{\nabla}_j f(\omega) &= \lim_{h \to 0} h^{-1}\big(f(\theta_{h\mathrm{e}_j}\omega) - f(\omega)\big), \\ \boldsymbol{\nabla} &= \big(\boldsymbol{\partial}_1, \ldots, \boldsymbol{\partial}_n\big), \end{aligned}$$

where $\{\mathrm{e}_1, \ldots, \mathrm{e}_n\}$ is the standard basis of $\mathbb{R}^n$, and the convergence is with respect to the $\mathrm{L}^2(\mathbb{P})$ norm. When

$$f = (f_1, \ldots, f_n) : \Omega \to \mathbb{R}^n,$$

is vector valued, we write $\boldsymbol{D} f(\omega)$ for a matrix whose j-th row is $\boldsymbol{\nabla} f_j$.

We write $\mathfrak{H}^1 = \mathfrak{H}^1(\mathbb{P})$ for the domain of $\boldsymbol{\nabla}$. Note that when $f \in \mathfrak{H}^1$, then the function $a \mapsto f(\theta_a \omega)$ is differentiable in $\mathrm{L}^2_{\mathrm{loc}}(\mathbb{R}^n)$. By Stone's theorem (see for example [La02]), there is a projection-valued measure $\mathbf{E}(\mathrm{d}\xi)$ such that

$$\begin{aligned} \mathcal{T}_x &= \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}x\cdot\xi} \, \mathbf{E}(\mathrm{d}\xi), \\ \boldsymbol{\nabla} &= \mathrm{i} \int_{\mathbb{R}^n} \xi \, \mathbf{E}(\mathrm{d}\xi). \end{aligned}$$

We also write $\mathfrak{H}^{-1} = \mathfrak{H}^{-1}(\mathbb{P})$ for the domain of the definition of the operator

$$\boldsymbol{\nabla}^{-1} = -\mathrm{i} \int_{\mathbb{R}^n} \xi^{-1} \, \mathbf{E}(\mathrm{d}\xi),$$

where $\xi^{-1} = (\xi_1^{-1}, \ldots, \xi_n^{-1})$. If

$$\mathrm{Z}_f(\mathrm{d}\xi) := \mathbf{E}(\mathrm{d}\xi) f, \qquad \mathrm{G}_f(\mathrm{d}\xi) := \langle \mathbf{E}(\mathrm{d}\xi) f, f \rangle,$$

then

$$f(x,\omega) := f(\theta_x \omega) = \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}x\cdot\xi} \, \mathrm{Z}_f(\mathrm{d}\xi, \omega), \tag{3.1}$$

$$\mathrm{R}_f(x) := \langle \mathcal{T}_x f, f \rangle = \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}x\cdot\xi} \, \mathrm{G}_f(\mathrm{d}\xi). \tag{3.2}$$

From

$$(-\Delta_x)^{\pm 1} f(x,\omega) = \int_{\mathbb{R}^n} |\xi|^{\pm 2}\ \mathrm{e}^{\mathrm{i}x\cdot\xi}\ \mathrm{Z}_f(\mathrm{d}\xi,\omega),$$

we learn that $f \in \mathfrak{H}^{\pm 1}$ if and only if

$$\int_{\mathbb{R}^n} |\xi|^{\pm 2}\ \mathrm{G}_f(\mathrm{d}\xi) < \infty.$$

In particular, from

$$(-\Delta_x)^{-1} R_f(x) = \int_{\mathbb{R}^n} |\xi|^{-2}\ \mathrm{e}^{\mathrm{i}x\cdot\xi}\ \mathrm{G}_f(\mathrm{d}\xi),$$

we deduce,

$$\int_{\mathbb{R}^n} |\xi|^{-2}\ \mathrm{G}_f(\mathrm{d}\xi) = (-\Delta)^{-1}\mathrm{R}_f(0) = \int_{\mathbb{R}^n} L(x)\mathrm{R}_f(x)\ \mathrm{d}x, \tag{3.3}$$

where $L : \mathbb{R}^n \to \mathbb{R}$ is given by:

$$L(x) = \begin{cases} (n\alpha_n)^{-1}|x|^{2-n} & n > 2, \\ -(2\pi)^{-1}\log|x| & n = 2, \end{cases} \tag{3.4}$$

where α_n is the $(n-1)$-dimensional surface area of the unit sphere $\mathbb{S}^{n-1}$.

(ii) We write $\widehat{\mathfrak{H}}^{-1} = \widehat{\mathfrak{H}}^{-1}(\mathbb{P})$ for the set of $f \in \mathrm{L}^2(\mathbb{P})$ such that

$$\int_{|x|\geqslant 1} |L(x)\mathrm{R}_f(x)|\ \mathrm{d}x < \infty.$$

Equivalently,

$$\int_{\mathbb{R}^{2d}} |L(x)\mathrm{R}_f(x)|\ \mathrm{d}x < \infty.$$

because $|\mathrm{R}_f| \leqslant \|f\|_{\mathrm{L}^2(\mathbb{P})}$, and the function L is integrable near 0. As a consequence,

$$\widehat{\mathfrak{H}}^{-1} \subseteq \mathfrak{H}^{-1}.$$

Remark 3.2 Assume that $f \in \mathrm{L}^2(\mathbb{P})$ and

$$\int_\Omega f\ \mathrm{d}\mathbb{P} = 0,$$

and write $\mathcal{L}(f)$ for the $\mathrm{L}^2(\mathbb{P})$-closure of the span of the set $\big\{\theta_a f :\ a \in \mathbb{R}^n\big\}$. The spectral representation (3.1) can be used to define a $\mathrm{L}^2(\mathbb{P})$-isometry between $\mathcal{L}(f)$ and $\mathrm{L}^2(\mathrm{G_f})$:

$$\mathcal{I}_f : \mathrm{L}^2(\mathrm{G_f}) \to \mathrm{L}^2(\mathbb{P}),$$

so that if $\chi_x(\xi) = \mathrm{e}^{\mathrm{ix}\cdot\xi}$, then

$$\mathcal{I}\big(\chi_x\big) = \mathcal{T}_x f,$$

(see for example [AT07, Section 5.4].) To explain this, observe that for any bounded continuous function $\zeta : \mathbb{R}^n \to \mathbb{R}$, we can use (3.1) to write

$$F_\zeta(\omega) := \int_{\mathbb{R}^n} \zeta(x) f(\theta_x\omega)\ \mathrm{d}x = \int_{\mathbb{R}^n} \hat{\zeta}(\xi)\ Z(\mathrm{d}\xi,\omega),$$

where

$$\hat{\zeta}(\xi) = \int_{\mathbb{R}^n} \zeta(x) e^{\mathrm{i}x\cdot\xi}\ \mathrm{d}x.$$

From this, one can show

$$\mathbb{E}|F_\zeta|^2 = \int_{\mathbb{R}^n} \left|\hat{\zeta}(\xi)\right|^2\ G_f(\mathrm{d}\xi).$$

Clearly, $\mathcal{I}_f(\hat{\zeta}) = F_\zeta$.

We continue with some preparatory definitions regarding stationary functions and twist maps.

Definition 3.3.

(i) Let us write $\mathcal{H}$ for the space of C^2 Hamiltonian functions $H : \mathbb{R}^{2d} \times \mathbb{R} \to \mathbb{R}$. For each $a = (b, c) \in \mathbb{R}^d \times \mathbb{R}^d$, we define

$$\begin{aligned} (\tau_b H)(q,p,t) &= H(q+b,p,t), \\ (\eta_c H)(q,p,t) &= H(q,p+c,t), \\ (\theta_a H)(q,p,t) &= H(q+b,p+c,t). \end{aligned}$$

(ii) We write $\mathcal{C}^1$ for the set of C^1 maps $\Phi : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$. We set

$$\mathcal{F}(\Phi) = \Phi - \mathrm{id},$$

where id denotes the identity map.

We write $\mathcal{S}$ for the set of symplectic diffeomorphisms $\Phi : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ such that $\mathcal{F}(\Phi)$ is uniformly bounded. We also set $\tilde{\mathcal{S}} = \mathcal{F}\big(\mathcal{S}\big)$.

For $a \in \mathbb{R}^{2d}$, the translation operators $\theta_a : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ and $\theta_a, \theta'_a : \mathcal{C}^1 \to \mathcal{C}^1$ are defined by

$$\begin{aligned} \theta_a(x) &= x + a, \\ \theta_a \omega &= \omega \circ \theta_a, \\ \theta'_a &= \mathcal{F}^{-1} \circ \theta_a \circ \mathcal{F}, \end{aligned}$$

for $x \in \mathbb{R}^{2d}$ and $\omega \in \mathcal{C}^1$. Note that for $\Phi \in \mathcal{C}^1$,

$$\begin{aligned} \big(\theta'_a \Phi\big)(x) &= (\theta_{-a} \circ \Phi \circ \theta_a)(x) \\ &= \Phi(x+a) - a. \end{aligned}$$

(iii) Let $\mathcal{P}$ be a θ'-invariant probability measure on $\mathcal{S}$. The map $\mathcal{F}$ pushes forward to a measure on $\tilde{S}$ that is denoted by $\mathcal{Q}$. This measure is θ invariant.

(iv) We define $\pi : \tilde{\mathcal{S}} \to \mathbb{R}$ to be the evaluation map $\pi(\omega) = \omega(0) = \Phi(0)$.

(v) Let Φ be a symplectic diffeomorphism with

$$\Phi(q,p) = \big(Q(q,p), P(q,p)\big).$$

We say that Φ is *twist* if for every $p \in \mathbb{R}^d$, the map $q \mapsto Q(q,p)$ is a diffeomorphism of $\mathbb{R}^d$. We write $\hat{q}(Q,p)$ for the inverse:

$$Q(q,p) = Q \quad\iff\quad q = \hat{q}(Q,p).$$

We also set $\hat{P}(Q,p)=P\big(\hat{q}(Q,p),p\big)$, and

$$\begin{aligned}\widehat{\Phi}(Q,p) &= \big(\hat{q}(Q,p),\hat{P}(Q,p)\big),\\ \widetilde{\Phi}(Q,p) &= \big(\hat{P}(Q,p),\hat{q}(Q,p)\big).\end{aligned}$$

We are now ready to state the main result of this section.

Theorem 3.4. *Let $\mathcal{Q}$ be a θ-invariant measure such that $\pi \in \widehat{\mathfrak{H}}^{-1}(\mathcal{Q})$,*

$$\int_{\mathcal{S}} \Phi(0)\ \mathcal{P}(\mathrm{d}\Phi) = \int_{\tilde{\mathcal{S}}} \omega(0)\ \mathcal{Q}(\mathrm{d}\omega) = 0, \tag{3.5}$$

and

$$\int_{\tilde{\mathcal{S}}} \|\mathrm{D}\omega\|_{\mathrm{C}^0}^d\ \mathcal{Q}(\mathrm{d}\omega) < \infty. \tag{3.6}$$

Assume that $\Phi = \mathrm{id} + \omega$ is C^2 twist diffeomorphism $\mathcal{Q}$-almost surely. Then there exists a unique function $\hat{w}:\tilde{\mathcal{S}}\to\mathbb{R}$, with $\hat{w}\in \mathrm{L}^2(\mathcal{Q})$, and

$$\int_{\tilde{\mathcal{S}}} \hat{w}(\omega)\ \mathcal{Q}(\mathrm{d}\omega) = 0, \tag{3.7}$$

such that if

$$w(x,\omega) := \hat{w}\big(\theta_x\omega\big), \qquad W(Q,p,\omega) = Q\cdot p + w(Q,p,\omega),$$

then

$$\widehat{\Phi} = (W_p, W_Q) =: \widehat{\nabla} W,$$

$\mathcal{Q}$-almost surely.

Remark 3.5 Note that $\widehat{\nabla} = \big(\partial_p,\partial_Q\big)$ represents the gradient operator with ∂_p and ∂_Q swapped. We refer to $\hat{w}$ of Theorem 3.4 as a *stationary generating function.* According to this theorem, a θ' stationary symplectic twist map always possesses a stationary generating function. A natural question is whether the converse is true. Given a function $\hat{w}$ such that the corresponding stationary process

$$w(x) = w(x,\omega) := \hat{w}\big(\theta_x\omega\big),$$

is C^2, can we use this function to produce a symplectic θ'-stationary twist map Φ? This is equivalent to the condition that

$$Q \mapsto W_p(Q,p) = Q + w_p(Q,p),$$

is a diffeomorphism for each p, so that we can solve the equation

$$W_p(Q,p) = q$$

for $Q = Q(q,p)$. This is always possible if the C^2 norm of w is small (see Proposition 3.1(iv)). Moreover, when $d=1$, we need

$$\begin{aligned} &W_{pQ} = 1 + w_{pQ} > 0, \qquad W_p(\pm\infty,p) = \pm\infty, \quad \text{or}\\ &W_{pQ} = 1 + w_{pQ} < 0, \qquad W_p(\pm\infty,p) = \mp\infty.\end{aligned}$$

The latter condition can be guaranteed by assuming that w_p is a bounded function.

With the previous definitions in mind, we state and prove three preparatory propositions.

Proposition 3.1. *The following statements hold.*

(i) *For every symplectic twist diffeomorphism* $\Phi : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$, *and* $a \in \mathbb{R}^d$, *we have*

$$\widehat{\theta'_a \Phi} = \theta'_a \widehat{\Phi}.$$

(ii) *If* $\pi \in \widehat{\mathfrak{H}}^{-1}(\mathcal{Q})$, *and*

$$\int_{\mathcal{S}} \|\mathrm{D}\Phi\|^d_{\mathrm{C}^0}\ \mathcal{P}(\mathrm{d}\Phi) = \int_{\tilde{\mathcal{S}}} \|\mathrm{D}\omega\|^{\mathrm{d}}_{\mathrm{C}^0}\ \mathcal{Q}(\mathrm{d}\omega) < \infty,$$

then

$$\hat{\pi} \in \widehat{\mathfrak{H}}^{-1}(\mathcal{Q}),$$

where $\hat{\pi}(\omega) := \widehat{\Phi}(0)$.

(iii) *Assume that* $\Phi \in \mathcal{S}$ *is symplectic twist diffeomorphism. Then there exists a* C^2 *function*

$$W\colon \mathbb{R}^{2d} \to \mathbb{R}$$

such that $\widetilde{\Phi} = \nabla W$. *Moreover, for* w *defined by*

$$w(Q,p) := W(Q,p) - Q \cdot p,$$

we have

$$\|\nabla w\|_{\mathrm{C}^0} \leqslant \|\mathcal{F}(\Phi)\|_{\mathrm{C}^0}. \tag{3.8}$$

(iv) *Let* w *be a* C^2 *function with* $\|\mathrm{D}^2 w\| < 1$, *and set*

$$W(Q,p) = Q \cdot p + w(Q,p). \tag{3.9}$$

Then there exists a symplectic twist diffeomorphism Φ *such that*

$$\Phi\big(W_p(Q,p),p\big) = \big(Q, W_Q(Q,p)\big).$$

Proof. (i) Let us write

$$\begin{aligned} \Phi'(q,p) &:= (\theta'_a\Phi)(q,p) = \big(Q'(q,p), P'(q,p)\big), \\ \widehat{\Phi}'(Q,p) &= \big(\hat{q}'(Q,p), \hat{P}'(Q,p)\big). \end{aligned}$$

This implies

$$\begin{aligned} Q'(q,p) = Q(q+b,p+c) - b = Q \quad &\Longleftrightarrow \quad \hat{q}'(Q,p) = q, \\ Q(q+b,p+c) = Q + b \quad &\Longleftrightarrow \quad \hat{q}(Q+b,p+c) = q + b. \end{aligned}$$

Hence

$$\hat{q}'(Q,p) = \hat{q}(Q+b,p+c) - b.$$

On the other hand

$$\begin{aligned} \hat{P}'(Q,p) &= P'\big(\hat{q}'(Q,p),p\big) \\ &= P\big(\hat{q}'(Q,p) + b, p + c\big) - c \\ &= P\big(\hat{q}(Q+b,p+c), p + c\big) - c \\ &= \hat{P}\big(Q+b,p+c\big) - c, \end{aligned}$$

as desired.

(ii) To ease the notation, we write $\mathbb{E}$ for the $\mathrm{d}\mathcal{Q}$ integration. Fix some $x^0 = (q^0, p^0) \in \mathbb{R}^{2d}$. Our claim reads as follows: If

$$\mathbb{E}\int_{|x-x^0|\geqslant 1} L(x-x^0)\big|(\Phi(x)-x)\cdot(\Phi(x^0)-x^0)\big|\ \mathrm{d}x < \infty, \tag{3.10}$$

then

$$\mathbb{E}\int_{|x-x^0|\geqslant 1} L(x-x^0)\big|(\widehat{\Phi}(x)-x)\cdot(\widehat{\Phi}(x^0)-x^0)\big|\ \mathrm{d}x < \infty. \tag{3.11}$$

By stationarity, neither the statement (3.10) nor the statement (3.11) depend on the choice of the point x^0. To ease the notation, we write

$$\begin{aligned} X &= (Q,P) = \Phi(x) = \Phi(q,p), \\ X' &= (Q,p). \end{aligned}$$

Write $y^0 = (Q^0, p^0)$, where $Q^0 = Q(x^0)$. Since

$$(\Phi(x)-x)\cdot(\Phi(x^0)-x^0) = \left(\widehat{\Phi}(X')-X'\right)\cdot\left(\widehat{\Phi}(y^0)-y^0\right),$$

the statement (3.10) can be rewritten as

$$\mathbb{E}\int_{|x-x^0|\geqslant 1} L(x-x^0)\big|(\widehat{\Phi}(X')-X')\cdot(\widehat{\Phi}(y^0)-y^0)\big|\ \mathrm{d}q\ \mathrm{d}p < \infty. \tag{3.12}$$

Assuming this, we wish to show

$$\mathbb{E}\int_{|X'-y^0|\geqslant 1} L(X'-y^0)\big|(\widehat{\Phi}(X')-X')\cdot(\widehat{\Phi}(y^0)-y^0)\big|\ \mathrm{d}Q\ \mathrm{d}p < \infty. \tag{3.13}$$

Since the law of both

$$x \mapsto \Phi(x) - x,$$

and

$$X' \mapsto \widehat{\Phi}(X') - X'$$

are θ-invariant, we can choose $x^0 = 0$ in (3.12), and $y^0 = 0$ in (3.13).

We first assume that $d > 1$. Observe that if $c_0 \geqslant \sup|\mathcal{F}(\Phi)|$, then

$$|X'-x| \leqslant c_0,$$

and

$$\begin{aligned} |x| \geqslant 2c_0 &\implies \frac{1}{2}|x| \leqslant |X'| \leqslant \frac{3}{2}|x|, \\ |X'| \geqslant 3c_0 &\implies \frac{2}{3}|X'| \leqslant |x| \leqslant \frac{4}{3}|X'|. \end{aligned} \tag{3.14}$$

This in turn implies under the assumption $|x| \geqslant 2c_0$,

$$\begin{aligned} \big||X'|^{-r} - |x|^{-r}\big| &= |X'|^{-r}|x|^{-r}\big||X'|^r - |x|^r\big| \\ &= (|X'||x|)^{-r}\big||X'| - |x|\big|\sum_{j=0}^{r-1}|X'|^j|x|^{r-1-j} \\ &\leqslant c_1 c_0 |x|^{-r-1}, \end{aligned}$$

for a constant $c_1 = c_1(r)$ that depends on $r = 2d - 2$ only. We additionally require $c_0 \geqslant 1/2$ so that $2c_0 \geqslant 1$. Hence, for

$$\Lambda := \int_{|x|\geqslant 2c_0} \left|(\widehat{\Phi}(X') - X') \cdot (\widehat{\Phi}(y^0) - y^0)\right| \, |X'|^{-r} \, \mathrm{d}x,$$

we have

$$\begin{aligned}\Lambda &= \int_{|x|\geqslant 2c_0} \left|(\widehat{\Phi}(X') - X') \cdot (\widehat{\Phi}(y^0) - y^0)\right| \, |X'|^{-r} \, \mathrm{d}x \\ &\leqslant \int_{|x|\geqslant 2c_0} \left|(\widehat{\Phi}(X') - X') \cdot (\widehat{\Phi}(y^0) - y^0)\right| \, |x|^{-r} \, \mathrm{d}x \\ &\quad + c_1 c_0 \int_{|x|\geqslant 2c_0} \left|(\widehat{\Phi}(X') - X') \cdot (\widehat{\Phi}(y^0) - y^0)\right| \, |x|^{-r-1} \, \mathrm{d}x \\ &=: \Lambda_0 + \Lambda_1,\end{aligned}$$

where $X^0 = \Phi(x^0)$. The assumption (3.12) implies then

$$\mathbb{E}(\Lambda_0 + \Lambda_1) < \infty.$$

Hence

$$\mathbb{E}\Lambda < \infty. \tag{3.15}$$

We now make a change of variables to replace q with $\hat{q}(Q,p)$ in Λ. Note

$$\begin{aligned} \mathrm{d}Q \ \mathrm{d}p &= \left|\det Q_q(q,p)\right| \ \mathrm{d}q \ \mathrm{d}p, \\ \left|\det Q_q(q,p)\right| &\leqslant d! |\mathrm{D}\Phi|^d =: c_2. \end{aligned} \tag{3.16}$$

From this, (3.14), and (3.15) we learn

$$\begin{aligned} &\mathbb{E} \int_{|X'|\geqslant 3c_0} \left|(\widehat{\Phi}(X') - X') \cdot (\widehat{\Phi}(y^0) - y^0)\right| \, |X'|^{-r} \, \mathrm{d}Q \ \mathrm{d}p \\ \leqslant\ & \mathbb{E} \int_{|x|\geqslant 2c_0} \left|(\widehat{\Phi}(X') - X') \cdot (\widehat{\Phi}(y^0) - y^0)\right| \, |X'|^{-r} \, \mathrm{d}Q \ \mathrm{d}p \\ \leqslant\ & c_2 \int_{|x|\geqslant 2c_0} \left|(\widehat{\Phi}(X') - X') \cdot (\widehat{\Phi}(y^0) - y^0)\right| \, |X'|^{-r} \, \mathrm{d}q \ \mathrm{d}p \\ =\ & c_3 \mathbb{E}\Lambda < \infty. \end{aligned}$$

Because of this, the claim (3.13) (in the case $y_0 = 0$) would follow if we can show

$$\mathbb{E} \int_{1\leq |X'|\leqslant 3c_0} \left|\left(\widehat{\Phi}(X') - X'\right) \cdot \left(\widehat{\Phi}(y^0) - y^0\right)\right| \ |X'|^{-r} \, \mathrm{d}Q \ \mathrm{d}p < \infty,$$

Since the law of $\widehat{\Phi}(X') - X'$ is θ-stationary, we can bound the left-hand side by

$$\mathbb{E}\left|\widehat{\Phi}(0)\right|^2 \int_{1\leq |X'|\leqslant 3c_0} |X'|^{-r} \, \mathrm{d}Q \ \mathrm{d}p =: c_3 \mathbb{E}\left|\widehat{\Phi}(0)\right|^2.$$

It remains to verify

$$\mathbb{E}\left|\widehat{\Phi}(0)\right|^2 < \infty. \tag{3.17}$$

Indeed from $|X'-x|\leqslant c_0$, (3.15), and the stationarity, we deduce

$$\begin{aligned}
\mathbb{E}\big|\widehat{\Phi}(0)\big|^2 =& \frac{1}{|\mathrm{B}_1(0)|}\ \mathbb{E}\int_{\mathrm{B}_1(0)} \big|\widehat{\Phi}(X')-X'\big|^2\ \mathrm{d}Q\ \mathrm{d}p\\
\leqslant& c_2\frac{1}{|\mathrm{B}_1(0)|}\ \mathbb{E}\int_{\mathrm{B}_{1+c_0}(0)} \big|\widehat{\Phi}(X')-X'\big|^2\ \mathrm{d}q\ \mathrm{d}p\\
=& c_2\frac{1}{|\mathrm{B}_1(0)|}\ \mathbb{E}\int_{\mathrm{B}_{1+c_0}(0)} \big|\Phi(x)-x\big|^2\ \mathrm{d}q\ \mathrm{d}p\\
=& c_2\frac{|\mathrm{B}_{1+c_0}(0)|}{|\mathrm{B}_1(0)|}\ \mathbb{E}\big|\Phi(0)\big|^2<\infty,
\end{aligned}$$

where we used

$$\big|\widehat{\Phi}(X')-X'\big|^2=\big|\Phi(x)-x\big|^2.$$

for the second equality. The proof is complete when $d>1$.

The proof in the case $d=1$ is similar. Observe that from

$$\begin{aligned}
\big|\log|X'|-\log|x|\big| &= \left|\int_{|x|}^{|X'|}\frac{\mathrm{d}r}{r}\right|\\
&\leqslant \frac{|X'-x|}{|X'|\wedge|x|},
\end{aligned}$$

and $|X'-x|\leqslant c_0$, we deduce

$$\begin{aligned}
\max\{|x|,|X'|\}\geqslant 2c_0 \quad &\Longrightarrow \quad \min\{|x|,|X'|\}\geqslant c_0\\
&\Longrightarrow \quad \big|\log|X'|-\log|x|\big|\leqslant 1.
\end{aligned}$$

This would allow us to repeat our proof for the case $d>1$ and finish the proof.

(iii) Since Φ is symplectic, we have

$$\begin{aligned}
0 &= \mathrm{d}\big(P\cdot\mathrm{d}Q-p\cdot\mathrm{d}q\big)\\
&= \mathrm{d}\big(\hat{P}\cdot\mathrm{d}Q-p\cdot\mathrm{d}\hat{q}\big)\\
&= \mathrm{d}\big(\hat{P}\cdot\mathrm{d}Q+\hat{q}\cdot\mathrm{d}p\big).
\end{aligned}$$

Hence, there exists a function $W=W(Q,p)$ such that

$$\mathrm{d}W=\hat{P}\cdot\mathrm{d}Q+\hat{q}\cdot\mathrm{d}p.$$

As a result,

$$\nabla W=\widetilde{\Phi}.$$

The inequality (3.8) is an immediate consequence of

$$\begin{aligned}
\mathcal{F}(\Phi)(q,p) &= (Q-q,P-p)\\
&= (Q-W_p(Q,p),W_Q(Q,p)-p)\\
&= (-w_p(Q,p),w_Q(Q,p)).
\end{aligned}$$

(iv) If we define

$$\begin{aligned}\widehat{\Phi}(Q,p) &= \big(W_p(Q,p), W_Q(Q,p)\big)\\ &= (Q,p) + \big(w_p(Q,p), w_Q(Q,p)\big),\end{aligned}$$

then $\widehat{\Phi}$ is a C^1 diffeomorphism by our assumption on w. In particular, the equation $W_p(Q,p) = q$, can be solved implicitly for $Q = Q(q,p)$. We may define

$$P(q,p) = W_Q\big(Q(q,p),p\big),$$

and

$$\Phi(q,p) = \big(Q(q,p), P(q,p)\big),$$

which concludes the proof. □

As we have learned from Proposition 3.1 (iii), a symplectic twist diffeomorphism always has a generating function. What Theorem 3.4 claims is the existence of a *stationary* generating function. Note that if we set

$$B(Q,p) = B(Q,q,\omega) := \widehat{\Phi}(Q,p) - (Q,p), \tag{3.18}$$

then B is a θ-stationary by Proposition 3.1(i). By Proposition 3.1(iii), we can express B as $\widehat{\nabla} w$ for a function $w(Q,p) = w(Q,p,\omega)$. We wish to show that this w can be chosen to be a stationary process with respect to θ. In the next Proposition, we state a sufficient condition (which is also necessary) for the existence of such stationary generating function.

Proposition 3.2. *Let $\mathcal{Q}$ be a θ-invariant probability measure on Ω, and let*

$$\hat{B} : \Omega \to \mathbb{R}^{2d}$$

be a function with the following properties:

(i) *$\hat{B} \in \mathfrak{H}^{-1}(\mathcal{Q})$, and*

$$\int_\Omega \hat{B}(\omega)\ \mathcal{Q}(\mathrm{d}\omega) = 0. \tag{3.19}$$

(ii) *There exists a C^2 function $v(x,\omega)$ such that*

$$B(x,\omega) := \hat{B}(\theta_x\omega) = \widehat{\nabla} v(x,\omega).$$

Then there exists a unique $\hat{w} \in \mathrm{L}^2(\mathcal{Q})$ such that

$$\int_\Omega \hat{w}\ \mathrm{d}\mathcal{Q} = 0,$$

and if $w(x,\omega) := \hat{w}(\theta_x\omega)$, then $\widehat{\nabla} w = B$, $\mathcal{Q}$-almost surely.

Proof. Since the process B is stationary, by the Spectral Theorem (3.1), we can find a vector measure

$$Z(\mathrm{d}\xi,\omega) = (Z_j(\mathrm{d}\xi,\omega) :\ j = 1,\dots,2d)$$

such that

$$\begin{aligned} B(x,\omega) &= \int_{\mathbb{R}^{2d}} \mathrm{e}^{\mathrm{i}x\cdot\xi}\ Z(\mathrm{d}\xi,\omega),\\ \mathrm{e}^{\mathrm{i}a\cdot\xi} Z(\mathrm{d}\xi,\omega) &= Z(\mathrm{d}\xi,\theta_a\omega).\end{aligned}$$

Let us write $\eta : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ for the function that swaps Q with p:

$$\eta(Q,p) = (p,Q).$$

If we write B' for $\eta(B)$, and Z' for $\eta(Z)$, then

$$B'(x,\omega) = \int_{\mathbb{R}^{2d}} \mathrm{e}^{\mathrm{i}x\cdot\xi}\ Z'(\mathrm{d}\xi,\omega).$$

Since $B = \widehat{\nabla} v$, for some function v, we have $B' = \nabla v$ is an exact derivative. This means

$$\mathrm{D}B'(x,\omega) = \mathrm{i}\left[\int_{\mathbb{R}^{2d}} \mathrm{e}^{\mathrm{i}x\cdot\xi}\ \xi_j Z'_k(\mathrm{d}\xi,\omega)\right]_{j,k=1}^{2d},$$

is a symmetric matrix. As a result,

$$\xi_j Z'_k(\mathrm{d}\xi,\omega) = \xi_k Z'_j(\mathrm{d}\xi,\omega),$$

which in turn implies that the scalar measure

$$z(\mathrm{d}\xi,\omega) = \xi_j^{-1} Z'_j(\mathrm{d}\xi,\omega),$$

is independent of j. In summary,

$$\begin{aligned} Z'(\mathrm{d}\xi,\omega) &= \xi\ z(\mathrm{d}\xi,\omega), \\ \mathrm{e}^{\mathrm{i}a\cdot\xi} z(\mathrm{d}\xi,\omega) &= z(\mathrm{d}\xi,\theta_a\omega). \end{aligned} \tag{3.20}$$

Hence

$$Z(\mathrm{d}\xi,\omega) = \eta(\xi)\ z(\mathrm{d}\xi,\omega). \tag{3.21}$$

Our candidate for $\hat{w}$ is simply

$$\hat{w}(\omega) := -\mathrm{i}z\big(\mathbb{R}^{2d},\omega\big).$$

We claim that our assumption $\hat{B} \in \mathfrak{H}^{-1}(\mathcal{Q})$ guarantees that $\hat{w}$ is well-defined and $\hat{w} \in \mathrm{L}^2(\mathcal{Q})$. Once this is done, we can then use (3.20) to deduce

$$w(x,\omega) := w(\theta_x\omega) = -\mathrm{i}\int_{\mathbb{R}^{2d}} \mathrm{e}^{\mathrm{i}x\cdot\xi}\ z(\mathrm{d}\xi,\omega). \tag{3.22}$$

Recall that if

$$R(a) := \mathbb{E}\ B(0,\omega) \otimes \bar{B}(0,\theta_a\omega)$$

represents the correlation of B, then by (3.2),

$$R(a) = \int_{\mathbb{R}^{2d}} \mathrm{e}^{\mathrm{i}\xi\cdot a}\ G_B(\mathrm{d}\xi), \quad \text{where} \quad G_B(\mathrm{d}\xi) = \mathbb{E}\ Z(\mathrm{d}\xi,\omega) \otimes \bar{Z}(\mathrm{d}\xi,\omega).$$

From this, and (3.21) we learn,

$$\begin{aligned} G_B(\mathrm{d}\xi) &= \eta(\xi) \otimes \overline{\eta(\xi)}\ \mathbb{E}\ |z|^2(\mathrm{d}\xi,\omega) \\ &=: \eta(\xi) \otimes \overline{\eta(\xi)}\ g(\mathrm{d}\xi) \\ &=: \left[G_B^{jk}(\mathrm{d}\xi)\right]_{\mathrm{j,k=1}}^{\mathrm{2d}}. \end{aligned}$$

Since $\hat{B} \in \mathfrak{H}^{-1}(\mathcal{Q})$, we have

$$\int_{\mathbb{R}^{2d}} |\xi|^{-2}\ G_B(d\xi) < \infty.$$

This means that the map $\xi \mapsto \zeta_j(\xi) := \xi_j^{-1}$ is in $\mathrm{L}^2(\mathrm{G})$. We then use the isometry $\mathcal{I}_B$ of Remark 3.2 to assert that $\hat{w} = \mathcal{I}_B(\zeta_j)$ is in $\mathrm{L}^2(\mathcal{Q})$. Moreover, if

$$r(a) := \mathbb{E}\ w(0,\omega)\bar{w}(0,\theta_a\omega),$$

represents the correlation of w, then

$$r(a) = \int_{\mathbb{R}^{2d}} \mathrm{e}^{\mathrm{i}\xi\cdot a}\ g(\mathrm{d}\xi).$$

In particular,

$$\mathbb{E}|\hat{w}|^2 = r(0) = g(\mathbb{R}^d) = \int_{\mathbb{R}^{2d}} |\xi_j|^{-2}\ G_B^{jj}(\mathrm{d}\xi) < \infty,$$

for any $j \in \{1,\dots,2d\}$. From differentiating (3.22), we can readily deduce that $\widehat{\nabla} w = B$ weakly. Since B is C^1, we conclude that $\hat{w} \in \mathrm{C}^2$.

It remains to verify the uniqueness of $\hat{w}$. Note that if $\hat{w}' \in \mathrm{L}^2(\mathcal{Q})$ such that

$$\int \hat{w}'\ \mathrm{d}\mathcal{Q} = 0,$$

and the corresponding w' is C^1 function satisfying $\widehat{\nabla} w' = B$, then $\hat{\zeta} = \hat{w} - \hat{w}'$ satisfies $\widehat{\nabla}\zeta = 0$, for $\zeta(x,\omega) = \hat{\zeta}(\theta_x\omega)$. This means that $\hat{\zeta}(\theta_x\omega)$ does not depend on x. Since the measure $\mathcal{Q}$ is ergodic with respect to θ, we deduce that $\hat{\zeta}$ is constant $\mathcal{Q}$-almost surely. Since $\mathcal{Q}$-integral of $\hat{z}$ is zero, we deduce that $\hat{z} = 0$. Hence $\hat{w} = \hat{w}'$, proving the uniqueness of w. □

Example 3.6 In this example, we examine the set $\mathfrak{H}^{-1}$ in the setting of quasiperiodic functions (see Example 2.8). On the torus $\Omega = [0,1]^N,\ 0 = 1$, we define the flow

$$\theta_x\omega = \omega + Ax \mod 1,$$

where A is a $N \times (2d)$ matrix, and $x \in \mathbb{R}^{2d}$. Recall that if $\mathbb{P}$ denotes the Lebesgue measure on Ω, then $\mathbb{P}$ is ergodic with respect to θ if and only if

$$m \in \mathbb{Z}^N \setminus \{0\} \quad \Longrightarrow \quad mA \neq 0.$$

Consider the function

$$\begin{aligned} u(\omega) &= \sum_{m\in\mathbb{Z}^N} a_m \mathrm{e}^{\mathrm{i}m\cdot\omega}, \\ u(\theta_x\omega) &= \sum_{m\in\mathbb{Z}^N} a_m \mathrm{e}^{\mathrm{i}(mA)\cdot x}\mathrm{e}^{\mathrm{i}m\cdot\omega}. \end{aligned}$$

From this and

$$\mathbb{E}\ u(\theta_x\omega)\overline{u(\omega)} = \sum_{m\in\mathbb{Z}^N} |a_m|^2 \mathrm{e}^{\mathrm{i}(mA)\cdot x},$$

we deduce,

$$\begin{aligned} Z(\mathrm{d}\xi,\omega) &= \sum_{m\in\mathbb{Z}^N} a_m \mathrm{e}^{\mathrm{i}m\cdot\omega}\ \delta_{mA}(\mathrm{d}\xi), \\ G(\mathrm{d}\xi) &= \sum_{m\in\mathbb{Z}^N} |a_m|^2 \delta_{mA}(\mathrm{d}\xi), \\ g(\mathrm{d}\xi) &= \sum_{m\in\mathbb{Z}^N} |a_m|^2 |mA|^{-2} \delta_{mA}(\mathrm{d}\xi). \end{aligned}$$

Hence

$$g(\mathbb{R}^{2d}) = \sum_{m\in\mathbb{Z}^N} |a_m|^2 |mA|^{-2}.$$

The function $u \in \mathfrak{H}^{-1}(\mathbb{P})$ if $g(\mathbb{R}^{2d}) < \infty$. For example, a *Diophantine condition* of the form

$$m \in \mathbb{Z}^N \setminus \{0\} \quad \Longrightarrow \quad |mA| \geqslant |m|^{-k},$$

yields

$$\sum_{m\in\mathbb{Z}^N} |m|^{2k} |a_m|^2 < \infty \quad \Longrightarrow \quad g(\mathbb{R}^{2d}) < \infty.$$

Hence if u possesses k many derivatives in L^2, then $u \in \mathfrak{H}^{-1}(\mathbb{P})$.

Our next ingredient for the proof of Theorem 3.4 is an application of Ergodic Theorem.

Proposition 3.3. *Let $v(x) = v(x,\omega) = \hat{v}(\tau_x\omega)$ be a stationary process with*

$$c_0 := \mathbb{E}|\hat{v}| < \infty.$$

Given $\ell = (\ell_1, \ldots, \ell_{2d})$, write

$$\mathrm{I}(\ell) = \prod_{i=1}^{2d} [-\ell_i, \ell_i].$$

Then almost surely, we can finds a sequence

$$\ell^r = (\ell_1^r, \ldots, \ell_{2d}^r)$$

such that $\ell_i^r \to \infty$ in large r limit, and

$$\sup_r \sigma(\partial\mathrm{I}(\ell^r))^{-1} \int_{\partial\mathrm{I}(\ell^r)} |v(x,\omega)|\ \sigma(\mathrm{d}x) < \infty,$$

where σ denotes the $2d-1$-dimensional surface measure.

Proof. To ease the notation, we write $h = |v|$. Given $r > 0$, write I_r for $[-r,r]^{2d}$, and define

$$\mathrm{M}(\omega) = \sup_{r\geqslant 1} |\mathrm{I}_r|^{-1} \int_{\mathrm{I}_r} h(x,\omega)\ \mathrm{d}x.$$

By the Maximal Ergodic Theorem (see for example Theorem 1.4 of [R]),

$$\mathbb{P}(A_s) := \mathbb{P}\left(\{\omega :\ \mathrm{M}(\omega) > s\}\right) \leqslant s^{-1} c_0.$$

Fix a large $s > 0$, and $\omega \in A_s^c$, so that

$$\mathrm{M}(\omega) \leqslant s.$$

Set $\hat{x} = (x_2, \ldots, x_{2d})$. Since $\omega \in A_s^c$, we can write

$$\begin{aligned}
\int_{\mathrm{I}_r} w(\theta_x \omega)\ \mathrm{d}x &= \int_0^r J_1^r(x_1)\ \mathrm{d}x_1 \\
&:= \int_0^r \left[\int_{\hat{\mathrm{I}}_r} (w(x_1,\hat{x}) + w(-x_1,\hat{x}))\ \mathrm{d}\hat{x}\right]\ \mathrm{d}x_1 \\
&\leqslant s|\mathrm{I}_r| = s(2r)^{2d},
\end{aligned}$$

where $\hat{\mathrm{I}}_r = [-r, r]^{2d-1}$. From this and Chebyshev's inequality,

$$\left|\left\{x_1 \in [0,r] :\ J_1^r(x_1) > 4s(2r)^{2d-1}\right\}\right| \leqslant \frac{r}{2}.$$

As a result, there exists $\ell_1^r \in [r/3, r]$ such that

$$J_1^r(\ell_1^r) \leqslant 4s(2r)^{2d-1}.$$

In the same fashion, we can write

$$\int_{\mathrm{I_r}} h(x,\omega)\ \mathrm{d}x = \int_0^r J_i^r(x_i)\ \mathrm{d}x_i,$$

for $i \in \{1, \ldots, 2d-1\}$, and find $\ell_i^r \in [r/3, r]$ such that

$$J_i^r(\ell_i^r) \leqslant 4s(2r)^{2d-1}.$$

For $\omega \in A_s^c$, and $\ell^r = \ell_r(\omega) = (\ell_1^r, \ldots, \ell_{2d}^r)$ as above, observe that for each j,

$$\begin{aligned}
\int_{\partial \mathrm{I}(\ell^r)} h(x,\omega)\ \sigma(\mathrm{d}x) &= \sum_{i=1}^{2d} J_i^r(\ell_i^r) \\
&\leqslant 8ds(2r)^{2d-1} \\
&\leqslant 8ds3^{2d-1} \prod_{i \neq j}^{2d} (2\ell_i^r),
\end{aligned}$$

because $\ell_i \in [r/3, r]$ for every i. From this we learn

$$\begin{aligned}
\int_{\partial \mathrm{I}(\ell^r)} h(x,\omega)\ \sigma(\mathrm{d}x) &\leqslant 4s3^{2d-1} \sum_{j=1}^{2d} \prod_{i \neq j}^{2d} (2\ell_i^r) \\
&= 2s3^{2d-1}\sigma(\partial \mathrm{I}(\ell^r)).
\end{aligned}$$

This completes the proof for $\omega \in A_s^c$. Since $\mathbb{P}(A_s) \to 0$ as $s \to \infty$, we are done. $\square$

With the aid of Propositions 3.1-3.3, we are now ready to tackle Theorem 3.4.

Proof of Theorem 3.4. *(Step 1)* Recall the process B that was defined in (3.18). By Proposition 3.1(i), the process B is stationary. We are done if we can apply Proposition 3.2 to B. For this, we need to verify the properties (i) and (ii) of this Proposition. Proposition 3.1(iii) verifies property (ii). The property (i) consists of two condition. The first condition of this property requires $\hat{A}$ to be in $\mathfrak{H}^{-1}(\mathcal{Q})$. For this, it suffices to show $\hat{A} \in \widehat{\mathfrak{H}}^{-1}(\mathcal{Q})$, which is an immediate consequence of Proposition 3.1(ii), and our assumptions $\pi \in \widehat{\mathfrak{H}}^{-1}(\mathcal{Q})$ and (3.6). It remains to verify (3.19):

$$a = (b, c) := \int_{\mathcal{S}} B(0)\ \mathcal{P}(\mathrm{d}\Phi) = \int_{\mathcal{S}} \widehat{\Phi}(0)\ \mathcal{P}(\mathrm{d}\Phi) = 0. \tag{3.23}$$

Observe that by Proposition 3.2 is applicable to $B(Q,p) - a$. In other words, there exists a C^2 stationary function $w(Q,p) = \hat{w}(\theta_{(Q,p)}\omega)$ such that

$$B(Q,p) - a = \widehat{\nabla} w,$$

or

$$\begin{aligned} \hat{P}(Q,p) &= \hat{P}(Q,p,\omega) = c + p + w_Q(Q,p,\omega), \\ \hat{q}(Q,p) &= \hat{q}(Q,p,\omega) = b + Q + w_p(Q,p,\omega). \end{aligned}$$

Recall that $\omega = \mathcal{F}(\Phi) = \Phi - \mathrm{id}$.

(Step 2) Using our assumption (3.5) and the stationarity of $\mathcal{Q}$,

$$\begin{aligned} 0 &= \iint_{\mathrm{I}(\ell)} \big(\Phi(q,p) - (q,p)\big)\, \mathrm{d}q\ \mathrm{d}p\ \mathcal{Q}(\mathrm{d}\omega) \\ &= \iint_{\mathrm{I}(\ell)} \big(Q(q,p) - q, P(q,p) - p\big)\ \mathrm{d}q\ \mathrm{d}p\ \mathcal{Q}(\mathrm{d}\omega) \\ &= \iint_{\mathrm{I}'(\ell)} \big(Q - \hat{q}(Q,p,\omega), \hat{P}(Q,p,\omega) - p\big)\ \mathrm{d}q\ \mathrm{d}p\ \mathcal{Q}(\mathrm{d}\omega) \\ &= \iint_{\mathrm{I}'(\ell)} \big(Q - \hat{q}(Q,p,\omega), \hat{P}(Q,p,\omega) - p\big)\ \det\big(\hat{q}_Q(Q,p,\omega)\big)\ \mathrm{d}Q\ \mathrm{d}p\ \mathcal{Q}(\mathrm{d}\omega) \\ &= \iint_{\mathrm{I}'(\ell)} \big(-b - w_p(Q,p,\omega), c + w_Q(Q,p,\omega)\big)\ \det\big(\mathrm{I}_d + w_{Qp}(Q,p,\omega)\big)\ \mathrm{d}Q\ \mathrm{d}p\ \mathcal{Q}(\mathrm{d}\omega) \\ &= |\mathrm{I}(\ell)|(-b,c) - \iint_{\mathrm{I}'(\ell)} J\nabla w(Q,p,\omega)\ \det\big(\mathrm{I}_d + w_{Qp}(Q,p,\omega)\big)\ \mathrm{d}Q\ \mathrm{d}p\ \mathcal{Q}(\mathrm{d}\omega), \end{aligned}$$

where I_d is the $d \times d$ identity matrix, $\ell = (\ell_1, \dots, \ell_{2d})$, $\mathrm{I}(\ell) = \prod_{i=1}^{2d} [-\ell_i, \ell_i]^{2d}$, and

$$\mathrm{I}'(\ell) = \big\{(Q,p):\ \ (\hat{q}(Q,p,\omega), p) \in \mathrm{I}(\ell)\big\}.$$

In summary,

$$(-b, c) = \frac{1}{|\mathrm{I}(\ell)|} \iint_{\mathrm{I}'(\ell)} J\nabla w(Q,p,\omega)\ \det\big(\mathrm{I_d} + w_{Qp}(Q,p,\omega)\big)\ \mathrm{d}Q\ \mathrm{d}p\ \mathcal{Q}(\mathrm{d}\omega). \tag{3.24}$$

Let us write

$$\begin{aligned} & c_0 = c_0(\omega) := \|\omega\|_{\mathrm{C}^0}, \qquad c_1 = c_1(\omega) := \|\mathrm{D}\omega\|_{\mathrm{C}^0}, \\ & Z_\ell = \int_{\mathrm{I}(\ell)} \nabla w(Q,p,\omega)\ \det\big(\mathrm{I}_d + w_{Qp}(Q,p,\omega)\big)\ \mathrm{d}Q\ \mathrm{d}p, \\ & Z'_\ell = \int_{\mathrm{I}'(\ell)} \nabla w(Q,p,\omega)\ \det\big(\mathrm{I}_d + w_{Qp}(Q,p,\omega)\big)\ \mathrm{d}Q\ \mathrm{d}p. \end{aligned} \tag{3.25}$$

Observe

$$\mathrm{I}(\ell - \mathrm{c}_0(\omega)) \subset \mathrm{I}'(\ell) \subset \mathrm{I}(\ell + \mathrm{c}_0(\omega)).$$

Since

$$\left|\mathrm{I}(\ell+\mathrm{c}_0(\omega))\setminus \mathrm{I}(\ell-\mathrm{c}_0(\omega))\right| \leqslant \mathrm{c}_2\mathrm{c}_0(\omega)\ell^{2\mathrm{d}-1},$$

for a constant c_2, we learn

$$\left|Z_\ell - Z'_\ell\right| \leqslant c_2 c_0(\omega)^2 c_1(\omega)^d d!\ \ell^{2d-1}. \tag{3.26}$$

Recall that we wish to show $b = c = 0$. On account of (3.26), (3.25), and $\mathcal{Q}$-almost sure finiteness of $c_0(\omega)+c_1(\omega)$, it suffices to show

$$\lim_{r\to\infty} |\mathrm{I}(\ell^r)|^{-1} \int_{\mathrm{I}(\ell^r)} \nabla w(Q,p,\omega)\ \det\big(\mathrm{I}_d + w_{Qp}(Q,p,\omega)\big)\ \mathrm{d}Q\ \mathrm{d}p = 0, \tag{3.27}$$

for ℓ^r as in Proposition 3.3.

(Step 3) Note that a interchanging Q with p, or performing a permutation among the variables $(Q_1,\dots,Q_d)$, or $(p_1,\dots,p_d)$ does not alter the integral in (3.27). Because of this, (3.27) would follow if we can show

$$\lim_{r\to\infty} |\mathrm{I}(\ell^r)|^{-1} \iint_{\mathrm{I}(\ell^r)} w_{Q_1}(Q,p,\omega)\ \det\big(\mathrm{I}_d + w_{Qp}(Q,p,\omega)\big)\ \mathrm{d}Q\ \mathrm{d}p\ \mathcal{Q}(\mathrm{d}\omega) = 0. \tag{3.28}$$

To simplify the notation, set

$$w_{Qp} =: A = [a_{ij}]_{i,j=1}^d.$$

Expanding the determinant in (3.28) yields

$$\det(\mathrm{I}+A) = 1 + \sum_{k=1}^{d} \sum_{1\leqslant i_1<\dots<i_k\leqslant d} \det[a_{i_j i_l}]_{j,l=1}^k.$$

This expansion yields an analogous expansion for the left-hand of (3.28). As we examine this expansion, we encounter two types of terms: Given k, either $1\in\{i_1,\dots,i_k\}$, or $1\notin\{i_1,\dots,i_k\}$. If the former occurs, we perform a permutation to rewrite the corresponding integral as

$$Z_\ell(\omega) := \iint_{\mathrm{I}(\ell)} w_{Q_1}(Q,p,\omega)\ \det\big[w_{Q_i p_j}(Q,p,\omega)\big]_{i,j=1}^k\ \mathrm{d}Q\ \mathrm{d}p.$$

If the latter occurs, then we must have $k<d$, and after a permutation, we rewrite the corresponding integral as

$$Z'_\ell(\omega) := \int_{\mathrm{I}(\ell)} w_{Q_1}(Q,p,\omega)\ \det\big[w_{Q_{i+1} p_j}(Q,p,\omega)\big]_{i,j=1}^k\ \mathrm{d}Q\ \mathrm{d}p.$$

We verify(3.28) by showing

$$\lim_{\ell\to\infty} |\mathrm{I}(\ell)|^{-1} Z_\ell(\omega) = 0, \tag{3.29}$$

$$\lim_{r\to\infty} |\mathrm{I}(\ell^r)|^{-1} Z'_{\ell^r}(\omega) = 0, \tag{3.30}$$

$\mathcal{Q}$-almost surely. (Recall $\ell^r = \ell^r(\omega)$ of Proposition 3.3 can depend on ω.)

(Step 4) We first focus on Z_ℓ. Let us write $\bar p = (p_1,\dots,p_k)$. If we fix Q and

$$p' = (p_{k+1},\dots,p_d),$$

and write

$$F(\bar p; Q, p') = F(\bar p) := \big(w_{Q_1}(Q,p),\dots,w_{Q_k}(Q,p)\big),$$

then for each (Q, p'), the $\mathrm{d}\bar{p}$ integration in Z_ℓ takes the form

$$\begin{aligned}
&\int_{\bar{\mathrm{I}}(\ell)} w_{Q_1}(Q,\bar{p},p')\ \det \mathrm{D}_{\bar{p}} F(\bar{p})\ \mathrm{d}\bar{p} \\
=& \int_{\bar{\mathrm{I}}(\ell)} \det \mathrm{D}_{\bar{p}} F'(\bar{p})\ \mathrm{d}\bar{p} \\
=& \int_{F'(\bar{\mathrm{I}}(\ell))} \mathrm{d}p_1 \wedge \cdots \wedge \mathrm{d}p_k,
\end{aligned}$$

where $D_{\bar{p}}$ denotes the differentiation with respect of $\bar{p}$, and

$$\bar{\mathrm{I}}(\ell) = \prod_{i=1}^{k} [-\ell_i, \ell_i], \qquad F' = \big(w_{Q_1}^2/2, w_{Q_2}, \dots, w_{Q_k}\big).$$

Here we are using the fact that $\mathrm{D}_{\bar{p}}F'$ is obtained from $\mathrm{D}_{\bar{p}}F$ by multiplying its first row by w_{Q_1}. Since $\|\nabla w\|_{\mathrm{C}^0} \leqslant c_0$, for $c_0 = c_0(\omega)$ as in (3.25), we can write

$$F'\big(\bar{\mathrm{I}}(\ell)\big) \subseteq \big[-c_0^2/2, c_0^2/2\big] \times [-c_0, c_0]^{k-1}.$$

Hence

$$\left| \int_{\bar{\mathrm{I}}(\ell)} \det \mathrm{D}_{\bar{p}} F'(\bar{p})\ \mathrm{d}\bar{p} \right| \leqslant 2^{k-1} c_0(\omega)^{k+1},$$

which yields the bound

$$|Z_\ell(\omega)| \leqslant 2^{k-1} c_0(\omega)^{k+1} \prod_{i=k+1}^{2d} (2\ell_i).$$

This certainly yields (3.29) because $k \geqslant 1$, and $c_0(\omega) < \infty$, $\mathcal{Q}$-almost surely.

(Step 5) We now turn our attention to Z'_ℓ. To ease the notation, let us set

$$f = w_{\hat{p}} = \big(w_{p_2}, \dots, w_{k+1}\big),$$

with

$$\hat{p} = (p_2, \dots, p_{k+1}),$$

and regard f as a column vector. With this interpretation, we can write

$$Z'_\ell = \int_{I(\ell)} w_{Q_1}\ \det\big[f_{Q_2}, \dots, f_{Q_{k+1}}\big]\ \mathrm{d}Q\ \mathrm{d}p.$$

We wish to integrate by parts with respect to Q_1. This can be performed with a boundary contribution that involves the functions w, and the first derivatives of f. More precisely,

$$Z'_\ell = \sum_{j=2}^{k+1} A_\ell^j + E_\ell^1, \tag{3.31}$$

where E_ℓ^1 satisfies a bound of the form

$$\big|E_\ell^1\big| \leqslant k!\ c_1(\omega)^k \int_{\partial \mathrm{I}(\ell)} |w|\ \mathrm{d}x, \tag{3.32}$$

and

$$A_\ell^j = -\int_{\mathrm{I}(\ell)} w\ \det\big[f_{Q_2}, \dots, f_{Q_{j-1}}, f_{Q_j Q_1}, f_{Q_{j+1}}, \dots, f_{Q_{k+1}}\big]\ \mathrm{d}Q\ \mathrm{d}p,$$

when $3 < j < k$ (and a similar expression when $j \in \{2, 3, k, k+1\}$). Here we are using the fact that $f \in \mathrm{C}^2$ (because $\Phi \in \mathrm{C}^2$), and

$$\|\mathrm{D}f\|_{\mathrm{C}^0} \leqslant \|\mathrm{D}^2 w\|_{\mathrm{C}^0} = \|\mathrm{D}\omega\|_{\mathrm{C}^0} = c_1(\omega).$$

We now perform an integration by parts with respect to the variable Q_j. This involves a boundary contribution that again involves w and first derivatives of f. Hence

$$A_\ell^j = B_\ell^j + \sum_{i=2, i\neq j}^{k+1} C_\ell^{ij} + E_\ell^{2,j}, \tag{3.33}$$

where $E_\ell^{2,j}$ satisfies a bound of the form

$$\left|E_\ell^{2,j}\right| \leqslant k!\ c_1(\omega)^k \int_{\partial \mathrm{I}(\ell)} |w|\ \mathrm{d}x, \tag{3.34}$$

and

$$B_\ell^j = \int_{\mathrm{I}(\ell)} w_{Q_j}\ \det\left[f_{Q_2}, \dots, f_{Q_{j-1}}, f_{Q_1}, f_{Q_{j+1}}, \dots, f_{Q_{k+1}}\right] \mathrm{d}Q\ \mathrm{d}p,$$

$$C_\ell^{ij} = \int_{\mathrm{I}(\ell)} w\ \det\left[g_2^{ij}, \dots, g_{k+1}^{ij}\right] \mathrm{d}Q\ \mathrm{d}p.$$

Here $g_r^{ij} = f_{Q_r}$ when $r \neq i, j$, $g_j^{ij} = f_{Q_1}$, and $g_i^{ij} = f_{Q_i Q_j}$. From anti-symmetry of the determinant, we can readily see that

$$C_\ell^{ij} = -C_\ell^{ji}.$$

From this, and (3.31)–(3.34) we deduce,

$$Z'_\ell = \sum_{j=2}^{k+1} B_\ell^j + E_\ell^3, \tag{3.35}$$

with E_ℓ^3 satisfying a bound of the form

$$\left|E_\ell^3\right| \leqslant [1 + k!k]\ c_1(\omega)^k \int_{\partial I(\ell)} |w|\ \mathrm{d}x. \tag{3.36}$$

(Step 6) We now study B_ℓ^j. Let us fix Q and

$$p'' = (p_1, p_{k+2}, \dots, p_d),$$

and focus on the $\hat{p}$-integration. We also set $h_r = w_{Q_r}$, and

$$\hat{\mathrm{I}}(\ell) = \prod_{i=2}^{k+1} [-\ell_i, \ell_i].$$

Note that the expression

$$B_\ell^j(Q, p'') := \int_{\hat{\mathrm{I}}(\ell)} w_{Q_j}\ \det\left[f_{Q_2}, \dots, f_{Q_{j-1}}, f_{Q_1}, f_{Q_{j+1}}, \dots, f_{Q_{k+1}}\right] \mathrm{d}\hat{p},$$

can be written as

$$\int_{\hat{\mathrm{I}}(\ell)} \mathrm{d}h_2 \wedge \dots \wedge \mathrm{d}h_{j-1} \wedge (h_j\ \mathrm{d}h_1) \wedge \mathrm{d}h_{j+1} \wedge \dots \wedge \mathrm{d}h_{k+1}.$$

If we set

$$\hat{B}^j_\ell(Q,p'') := \int_{\hat{\mathrm{I}}(\ell)} w_{Q_1}\ \det\left[f_{Q_2},\dots,f_{Q_{j-1}},f_{Q_j},f_{Q_{j+1}},\dots,f_{Q_k}\right]\,\mathrm{d}\hat{p}$$
$$= \int_{\hat{\mathrm{I}}(\ell)} \mathrm{d}h_2\wedge\cdots\wedge\mathrm{d}h_{j-1}\wedge(h_1\ \mathrm{d}h_j)\wedge\mathrm{d}h_{j+1}\wedge\cdots\wedge\mathrm{d}h_{k+1},$$

then

$$B^j_\ell(Q,p'')+\hat{B}^j_\ell(Q,p'') = \int_{\hat{\mathrm{I}}(\ell)} \mathrm{d}h_2\wedge\cdots\wedge\mathrm{d}h_{j-1}\wedge\mathrm{d}(h_jh_1)\wedge\mathrm{d}h_{j+1}\wedge\cdots\wedge\mathrm{d}h_{k+1}.$$

But if

$$G(\hat{p}) = (h_2,\dots,h_{j-1},h_jh_1,h_{j+1},\dots,h_{k+1}),$$

then

$$B^j_\ell(Q,p'')+\hat{B}^j_\ell(Q,p'') = \int_{G(\hat{\mathrm{I}}(\ell))} \mathrm{d}p_2\wedge\cdots\wedge\mathrm{d}p_{k+1}. \tag{3.37}$$

Since the function $|h_j|$ is bounded by c_0, we deduce that

$$G(\hat{\mathrm{I}}(\ell)) \subset [-c_0,c_0]^{j-2}\times[-c_0^2,c_0^2]\times[-c_0,c_0]^{k-j+1}.$$

This and (3.37) imply

$$\left|B^j_\ell(Q,p'')+\hat{B}^j_\ell(Q,p'')\right| \leqslant 2^k c_0(\omega)^{k+1}.$$

From this we learn

$$B^j_\ell = -Z'_\ell + E^{4,j}_\ell,$$

where

$$\left|E^{4,j}_\ell\right| \leqslant 2^k c_0(\omega)^{k+1}(2\ell)^{2d-k}.$$

This and (3.35) imply,

$$Z'_\ell = -kZ'_\ell + E^3_\ell + E^4_\ell,$$

with

$$\left|E^4_\ell\right| \leqslant k2^k c_0(\omega)^{k+1}(2\ell)^{2d-k}. \tag{3.38}$$

In other words,

$$(k+1)Z'_\ell = E^3_\ell + E^4_\ell. \tag{3.39}$$

We now divide both sides of (3.39) by $|\mathrm{I}(\ell)|$, and choose

$$\ell = \ell^r$$

with ℓ^r as in Proposition 3.3, where v is chosen to be w. Finally we send $r\to\infty$ and use (3.36), (3.38), and Proposition 3.3 to deduce (3.30). This completes the proof.

4. The density of fixed points

From Theorem 3.4, we learn that a stationary symplectic twist diffeomorphism can be represented as

$$\Phi\big(Q + w_p(Q,p), p\big) = \big(Q, p + w_Q(Q,p)\big),$$

for a stationary process $w(Q,p,\omega)$. From this representation it is clear that

$$\Phi(Q,p) = (Q,p)$$

if and only if

$$f(Q,p,\omega) := \widehat{\nabla} w(Q,p,\omega) = 0,$$

(or equivalently $\nabla w(Q,p,\omega) = 0$). In words, there exists a one-to-one correspondence between the fixed points of Φ and the critical points of w. We now have the question of existence of critical points of w before us. To ease the notation, we write x for (Q,p).

We wish to use Ergodic Theorem to count the number of points in the zero set of the stationary process f, restricted to a large box.

As a preparation for the statement of the main results of this section, we make some definitions.

Definition 4.1. Recall that a function $f : \mathbb{R}^n \times \Omega \to \mathbb{R}^n$ is *θ-stationary* if

$$f(x + a, \omega) = f(x, \theta_a \omega),$$

or equivalently, $f(x,\omega) = \hat{f}(\theta_x \omega)$ for $\hat{f}(\omega) = f(0,\omega)$. Given a stationary function $f(x,\omega)$, and an open set U of $\mathbb{R}^{n\times n}$, we define

$$\mathcal{N}_U(a,\Lambda,\omega) = \sharp Z_U(a,\Lambda,\omega),$$

where $Z_U(a,\Lambda,\omega) = Z_U(a,\omega) \cap \Lambda$, where

$$Z_U(a,\omega) = \big\{x \in \mathbb{R}^n : \ f(x,\omega) = a, \ \mathrm{D}f(x,\omega) \in U\big\}.$$

We simply write

$$\mathcal{N}_U(a,\omega) := \mathcal{N}_U(a,[0,1]^n,\omega).$$

As an example of U, we may consider the set of matrices $\Gamma \in \mathbb{R}^{n\times n}$ with exactly k many negative eigenvalues.

Proposition 4.1. *Assume f is θ-stationary with respect to the probability measure $\mathcal{Q}$. If*

$$\mathbb{E}\mathcal{N}_U(a,\omega) = \int_\Omega \mathcal{N}_U(a,\omega)\ \mathcal{Q}(\mathrm{d}\omega) < \infty, \tag{4.1}$$

then

$$\lim_{\ell\to\infty} (2\ell)^{-n} \mathcal{N}_U\big(a, [-\ell,\ell)^n, \omega\big) = \mathbb{E}\mathcal{N}_U(a,\omega), \tag{4.2}$$

$\mathcal{Q}$-almost surely, and in $\mathrm{L}^1(\mathcal{Q})$ sense.

Proof. Observe that if k and k' are two positive integers with $k < k'$, then

$$\mathcal{N}_U\big(a,[k,k')^n,\omega\big) = \sum_{i_1,\dots,i_n=k}^{k'-1} \mathcal{N}_U\big(a,[i_1,i_1+1)\times\ldots\times[i_n,i_n+1),\omega\big).$$

This and the stationarity imply

$$\mathbb{E}\ \mathcal{N}_U\big(a,[k,k')^n,\omega\big) = (k'-k)^n \mathbb{E}\mathcal{N}_U(a,\omega). \tag{4.3}$$

If we assume (4.1), then the right-hand side of (4.3) is finite. This in particular implies that the set $Z_U(a,\omega)$ is discrete, and that $\mathcal{N}_U\big(a,[-\ell,\ell]^n,\omega\big)$ is finite, $\mathcal{Q}$-almost surely.

By the Ergodic Theorem,

$$\lim_{\ell\to\infty}(2\ell)^{-n}\int_{[-\ell,\ell]^n}\mathcal{N}_U(a,\theta_x\omega)\;\mathrm{d}x=\mathbb{E}\mathcal{N}_U(a,\omega),\tag{4.4}$$

almost surely, and in $\mathrm{L}^1(\mathcal{Q})$ sense. We wish to use (4.4) to deduce (4.2).

Observe that by the stationarity,

$$Z_U(a,\theta_x\omega)=Z_U(a,\omega)-x:=\big\{y-x:\;y\in Z_U(a,\omega)\big\}.\tag{4.5}$$

From this we can readily deduce

$$\mathcal{N}_U(a,\Lambda,\theta_x\omega)=\mathcal{N}_U(a,\Lambda+x,\omega).$$

This in turn implies

$$\begin{aligned}\mathcal{N}_U(a,\theta_x\omega)=\mathcal{N}_U\big(a,[0,1]^n+x,\omega\big)&=\sum_{z\in Z(a,\omega)}\mathbb{1}\left(z\in[0,1]^n+x\right)\\&=\sum_{z\in Z(a,\omega)}\mathbb{1}\left(x\in[-1,0]^n+z\right).\end{aligned}$$

(By convention, the right-hand side is 0, when $Z_U(a,\omega)=\varnothing$.) As a consequence,

$$\begin{aligned}\int_{[-\ell-1,\ell+1]^n}\mathcal{N}_U(a,\theta_x\omega)\;\mathrm{d}x&=\sum_{z\in Z(a,\omega)}\int_{[-\ell-1,\ell+1]^n}\mathbb{1}\left(x\in[-1,0]^n+z\right)\;\mathrm{d}x\\&\geqslant\mathcal{N}_U\big(a,[-\ell,\ell]^n,\omega\big).\end{aligned}\tag{4.6}$$

In the same manner we can show

$$\int_{[-\ell+1,\ell-1]^n}\mathcal{N}_U(a,\theta_x\omega)\;\mathrm{d}x\le\mathcal{N}_U\big(a,[-\ell,\ell]^n,\omega\big)\;\mathrm{d}x.$$

From this, (4.6), and (4.4), we can readily deduce (4.2). □

On account of (4.3), we wish to find an explicit formula $\mathbb{E}\mathcal{N}(0,\omega)$. In particular we would like to find conditions that would guarantee

$$\mathbb{E}\mathcal{N}(0,\omega)>0,$$

so that we can deduce the existence of critical points of w. Formally speaking, we expect

$$\mathbb{E}\;\mathcal{N}_U(0,\omega)=\mathbb{E}\big[\mathbb{1}\big(\boldsymbol{D}\hat{f}(\omega)\in U\big)\;|\det\boldsymbol{D}\hat{f}(\omega)|\;\delta_0\big(\hat{f}(\omega)\big)\big],\tag{4.7}$$

provided that the right-hand side is well-defined. (See Definition 3.1(i), for the definition of the operator $\boldsymbol{D}$.) In this section, we give one set of sufficient conditions that would allow us to make sense of (4.7) (see (4.9) below). In Section 5, we will be able to use the classical coarea formula to rewrite the right-hand of (4.7) in a more tractable form when f is quasiperiodic.

As it turns out, a multi-dimensional generalization of the classical Kac-Rice formula, would allow us to express $\mathbb{E}\mathcal{N}(0,\omega)$ in terms of the probability density of the random variable

$$(\hat{f}(\omega),\boldsymbol{D}\hat{f}(\omega)).$$

We refer to references [AT07, AT09, AW09] for thorough discussions of Kac-Rice type formulas, and their applications for Gaussian processes. For our purposes, we need the following variant of Kac-Rice formula.

Theorem 4.2. *Suppose that* $f : \mathbb{R}^n \to \mathbb{R}^n$, $f(x,\omega) = \hat{f}(\theta_x\omega)$ *is* C^2 θ*-stationary process. Assume*

$$\mathbb{E}\ \mathcal{N}(a,\omega) < \infty, \tag{4.8}$$

for a *near* 0*, and the random variable* $(\hat{f}(\omega), \boldsymbol{D}\hat{f}(\omega))$ *has a density* $p(a,\Gamma)$ *such that the following conditions are true:*

- *The function*
$$Q(a) := \int_U |\det\Gamma|\ p(a,\Gamma)\ \mathrm{d}\Gamma,$$
is continuous near 0.
- *The function*
$$p(a) := \int_U\ p(a,\Gamma)\ \mathrm{d}\Gamma,$$
is bounded near 0.

Then

$$\mathbb{E}\ \mathcal{N}_U(0,\omega) = \int_U |\det\Gamma|\ p(0,\Gamma)\ \mathrm{d}\Gamma. \tag{4.9}$$

Proof. *(Step 1)* According to Area (or Coarea) Formula,

$$\int_{\mathbb{R}^n} \zeta(a)\mathcal{N}_U(a,\omega)\ \mathrm{d}a = \int_{[0,1]^n} \zeta(f(x))\ \mathbb{1}\big(\mathrm{D}f(x,\omega)\in U\big)\ |\det \mathrm{D}f(x,\omega)|\ \mathrm{d}x, \tag{4.10}$$

for every continuous function ζ (see for example [AW09]). From taking $\mathcal{Q}$-expectation of both sides we deduce

$$\int_{\mathbb{R}^n} \zeta(a)\big[\mathbb{E}\ \mathcal{N}_U(a,\omega)\big]\ \mathrm{d}a = \int_{\mathbb{R}^n}\int_U \zeta(a)|\det\Gamma|\ p(a,\Gamma)\ \mathrm{d}\Gamma\ \mathrm{d}a = \int_{\mathbb{R}^n} \zeta(a)Q(a)\ \mathrm{d}a.$$

Hence,

$$\mathbb{E}\ \mathcal{N}_U(a,\omega) = \int_U |\det\Gamma|\ p(a,\Gamma)\ \mathrm{d}\Gamma, \tag{4.11}$$

for Lebesgue almost all $a \in \mathbb{R}^n$. We wish to show that (4.11) holds for *all* a near 0. To achieve this we will show that for each $\delta > 0$, there exists a measurable set $\Omega(\delta) \subset \Omega$ such that the following statements hold:

(i) $\lim_{\delta\to 0} \mathcal{Q}\big(\Omega(\delta)\big) = 1$.

(ii) If $\omega \in \Omega(\delta)$, and $|u| \leqslant \delta$, then
$$\mathcal{N}_U(u,\omega) = \mathcal{N}_U(0,\omega).$$

(iii) If $\omega \in \Omega(\delta)$, and $\varepsilon \in (0,\delta)$, then

$$\mathcal{N}_U(0,\omega) = \frac{1}{|\mathrm{B}_\varepsilon(0)|}\int_{[0,1]^n} \mathbb{1}(|f(x,\omega)| \leqslant \varepsilon)\ \mathbb{1}\big(\mathrm{D}f(x,\omega)\in U\big)\ |\det \mathrm{D}f(x,\omega)|\ \mathrm{d}x. \tag{4.12}$$

Let us demonstrate how (i)-(iii), and (4.11) imply (4.9). Indeed from (iii) we deduce that when $\varepsilon \in (0,\delta)$, the expression

$$\mathbb{E}\ \mathcal{N}_U(0,\omega)\mathbb{1}\big(\omega \in \Omega(\delta)\big),$$

equals

$$\begin{aligned}
&\mathbb{E}\ \frac{1}{|\mathrm{B}_\varepsilon(0)|}\int_{[0,1]^n} \mathbb{1}\big(|f(x,\omega)| \leqslant \varepsilon,\ \mathrm{D}f(x,\omega)\in U\big)\ |\det \mathrm{D}f(x,\omega)|\ \mathrm{d}x\ \mathbb{1}\big(\omega \in \Omega(\delta)\big)\\
&\leqslant \mathbb{E}\ \frac{1}{|\mathrm{B}_\varepsilon(0)|}\int_{[0,1]^n} \mathbb{1}\big(|f(x,\omega)| \leqslant \varepsilon,\ \mathrm{D}f(x,\omega)\in U\big)\ |\det \mathrm{D}f(x,\omega)|\ \mathrm{d}x\\
&= \frac{1}{|\mathrm{B}_\varepsilon(0)|}\mathbb{E}\ \mathbb{1}\big(|f(0,\omega)| \leqslant \varepsilon,\ \mathrm{D}f(0,\omega)\in U\big)\ |\det \mathrm{D}f(0,\omega)|\\
&= \frac{1}{|\mathrm{B}_\varepsilon(0)|}\int_U\int_{\mathrm{B}_\varepsilon(0)} |\det\Gamma|\ p(a,\Gamma)\ \mathrm{d}a\ \mathrm{d}\Gamma = \frac{1}{|\mathrm{B}_\varepsilon(0)|}\int_{\mathrm{B}_\varepsilon(0)} Q(a)\ \mathrm{d}a,
\end{aligned}$$

where we use the stationarity for the first equality.

We then send $\varepsilon \to 0$, and $\delta \to 0$ (in this order), and use the continuity of Q at 0 to deduce

$$\mathbb{E}\ \mathcal{N}_U(0,\omega) \leqslant \int_U |\det\Gamma| p(0,\Gamma)\ \mathrm{d}a\ \mathrm{d}\Gamma. \tag{4.13}$$

On the other-hand, by (4.11), we can find a sequence $a_k \to 0$ such that

$$\mathbb{E}\ \mathcal{N}_U(a_k,\omega) = \int_U |\det\Gamma| p(a_k,\Gamma)\ \mathrm{d}a\ \mathrm{d}\Gamma. \tag{4.14}$$

We use such a sequence to argue

$$\begin{aligned}
\mathbb{E}\ \mathcal{N}_U(0,\omega) &\leqslant \int_U |\det\Gamma| p(0,\Gamma)\ \mathrm{d}a\ \mathrm{d}\Gamma\\
&= \lim_{k\to\infty}\int_U |\det\Gamma| p(a_k,\Gamma)\ \mathrm{d}a\ \mathrm{d}\Gamma\\
&= \lim_{k\to\infty}\mathbb{E}\ \mathcal{N}_U(a_k,\omega)\\
&= \lim_{k\to\infty}\lim_{\delta\to 0}\mathbb{E}\ \mathcal{N}_U(a_k,\omega)\ \mathbb{1}\big(\omega\in\Omega(\delta)\big)\\
&= \lim_{k\to\infty}\lim_{\delta\to 0}\mathbb{E}\ \mathcal{N}_U(0,\omega)\ \mathbb{1}\big(\omega\in\Omega(\delta)\big)\\
&= \mathbb{E}\ \mathcal{N}_U(0,\omega),
\end{aligned}$$

where we used

- (4.13) for the beginning inequality,
- the continuity of Q at 0, for the first equality,
- (4.14) for the second equality,
- (i) for the third equality,
- (ii) for the fourth equality
- (i) for the last equality.

Since we must have equality for the inequalities in the above display, we arrive at (4.9). It remains to construct the sets $\Omega(\delta),\ \delta > 0$, satisfying (i)-(iii).

(Step 2) Let us write $Z = Z(\omega)$ for the level set $Z_U(a,\omega)$, when $a = 0$. We also set

$$\Omega_0 = \big\{\omega : \text{ there exists } x \in [0,1]^n \text{ such that } f(x,\omega) = 0, \ \det \mathrm{D} f(x,\omega) = 0\big\},$$
$$\Omega_1 = \big\{\omega : \text{ there exists } x \in \partial\big([0,1]^n\big) \text{ such that } f(x,\omega) = 0\big\}.$$

We assert,

$$\mathcal{Q}(\Omega_0) = 0, \tag{4.15}$$
$$\mathcal{Q}(\Omega_1) = 0. \tag{4.16}$$

Our assumptions on f would allow us to use Proposition 6.5 of [AW09] to deduce (4.15).

Recall that by our assumption $\mathbb{E}\mathcal{N}(0,\omega) < \infty$, the set

$$Z(\omega) \cap [0,1]^n$$

is finite almost surely. This and stationarity imply that

$$Z(\omega) \cap [-\ell,\ell]^n$$

is finite for every positive ℓ. Hence the set $Z(\omega)$ is discrete almost surely.

We now argue that (4.16) is a consequence of the discreteness and the stationarity of the set $Z(\omega)$. To explain this, let us write $\lambda(\mathrm{d}x)$ for the Lebesgue measure on $\mathbb{R}$, and let us write

$$\pi_i(x_1,\dots,x_n) = x_i$$

for the i-th coordinate projection. Evidently, the discreteness of Z implies that $\lambda(Z(\omega)) = 0$, which in turn implies

$$(\lambda \times \mathcal{Q})\big(\big\{(a,\omega) : \ a \in \pi_i Z(\omega)\big\}\big) =: (\lambda \times \mathcal{Q})(\mathcal{Z}_i) = 0,$$

by Fubini's theorem. From this and Fubini's theorem again we learn

$$\lambda\left(\big\{a : \ \mathcal{Q}\big(\big\{\omega : \ (a,\omega) \in \mathcal{Z}_i\big\}\big) > 0\big\}\right) = 0. \tag{4.17}$$

By stationarity of $Z(\omega)$, the probability

$$\mathcal{Q}\big(\big\{\omega : \ (a,\omega) \in \mathcal{Z}_i\big\}\big),$$

is independent of a. From this and (4.17) we deduce

$$\mathcal{Q}\big(\big\{\omega : \ (a,\omega) \in \mathcal{Z}_i\big\}\big) = 0,$$

for every $a \in \mathbb{R}$. In particular,

$$\mathcal{Q}\big(\big\{\omega : \ \pi_i Z(\omega) \cap \{0,1\} \neq \varnothing\big\}\big) = 0,$$

As a consequence,

$$\mathcal{Q}\big(\big\{\omega : \ \pi_i Z(\omega) \cap \{0,1\} \neq \varnothing \ \text{ for some } \ i \in \{1,\dots,n\}\big\}\big) = 0.$$

This is exactly the claim (4.16).

(Step 3) We are now ready to construct our sets $\Omega(\delta)$, $\delta > 0$. Take $\omega \in \Omega \setminus \big(\Omega_0 \cup \Omega_1\big)$, and assume that $N = N(\omega) = \sharp Z(\omega) \neq 0$. If

$$Z(\omega) = \big\{a^1,\dots,a^N\big\},$$

then we can use the (local) inverse mapping theorem to find $\varepsilon(\omega) > 0$, and disjoint open sets

$$V_1(\omega),\dots,V_N(\omega) \subset [0,1]^n \cap U,$$

such that $a^i \in V_i(\omega)$, $f(V_i(\omega)) = \mathrm{B}_{\delta(\omega)}(0)$, and the restriction of f to each V_i is a diffeomorphism for each i. To have $\varepsilon(\omega)$ a measurable function, we choose $\varepsilon = \varepsilon(\omega)$ to be the largest positive number for which such sets $V_1, \dots, V_N$ exist. We set

$$\Omega(\delta) = \big\{\omega \in \Omega \setminus \big(\Omega_0 \cup \Omega_1\big) : \ \varepsilon(\omega) \geqslant \delta\big\}.$$

We now verify (i)-(iii) of *(Step 1)*. The property (i) is an immediate consequence of (4.14) and (4.15). The property (ii) follows from the fact that $\mathcal{N}_U(a,\omega) = N$ for $u \in \mathrm{B}_{\delta(\omega)}(0)$. To verify the third property, set

$$W_\varepsilon(\omega) = \big\{x : \ |f(x,\omega)| \leqslant \varepsilon\big\}.$$

When $\varepsilon < \delta(\omega)$, the right-hand side of (4.12) equals

$$\frac{1}{|\mathrm{B}_\varepsilon(0)|} \sum_{i=1}^{N} \int_{V_i(\omega)\cap W_\varepsilon(\omega)} |\det \mathrm{D}f(x,\omega)| \ \mathrm{d}x$$
$$= \frac{1}{|\mathrm{B}_\varepsilon(0)|} \sum_{i=1}^{N} \big|f(V_i(\omega) \cap W_\varepsilon(\omega))\big| = N,$$

because $f(V_i(\omega) \cap W_\varepsilon(\omega)) = \mathrm{B}_\varepsilon(0)$. This completes the verification of the third property. $\square$

Remark 4.3 In Theorem 4.2 we assumed that the law of the random variable $(\hat{f}, \boldsymbol{D}\hat{f})$ has a density $p(a,\Gamma)$. This requirement can be replaced with the following two conditions:

- The law of the random variable $\hat{f}$ has a density $p(a)$ that is continuous near 0.
- If $q(a, d\Gamma)$ is the the conditional probability distribution of $\boldsymbol{D}\hat{f}$, given $\hat{f} = a$, then the function

  $$a \mapsto \int_{\mathbb{R}^{n\times n}} |\det \Gamma| \ q(a, \mathrm{d}\Gamma),$$

  is continuous at $a = 0$. We remark that if $(\hat{f}, \boldsymbol{D}\hat{f})$ has a density $p(a,\Gamma)$, then q in terms of p is given by $q(a,\Gamma)/p(a)\mathrm{d}G$.

Theorem 4.2 and Proposition 4.1 give us a way of counting the fixed points of a stationary symplectic twist diffeomorphism provided that the conditions of Theorem 4.2 are met. The conditions of this theorem are stated for the density of the pair

$$(f, \mathrm{Df}) = \big(\widehat{\Phi} - \mathrm{id} \ , \ \mathrm{D}\widehat{\Phi} - \mathrm{I}\big).$$

In practice, we need conditions that are formulated for the original symplectic map Φ, not $\widehat{\Phi}$. The following result will remedy this.

Proposition 4.2. *Let Φ be as in Theorem 3.4. Assume that the pair*

$$(\Phi(x), \mathrm{D}\Phi(x))$$

has a density $\rho(x, X, \Gamma)$ with respect to the Lebesgue measure $\mathrm{d}X \ \mathrm{d}\Gamma$ of $\mathbb{R}^{2d} \times \mathbb{R}^{2d\times 2d}$. Then the pair $(\widehat{\Phi}(x), \mathrm{D}\widehat{\Phi}(x))$ has a density $\hat{\rho}(\hat{x}, \hat{X}, \hat{\Gamma})$ with respect to the Lebesgue measure $\mathrm{d}\hat{X} \ \mathrm{d}\hat{\Gamma}$, where

$$\hat{\rho}(\hat{x}, \hat{X}, \hat{\Gamma}) = |\det \hat{A}|^{1-4d} \rho(x, X, \mathcal{A}(\hat{\Gamma})), \tag{4.18}$$

with $x = (q,p)$*,* $X = (Q,P)$*,* $\hat{x} = (Q,p)$*,* $\hat{X} = (q,P)$*, and*

$$\mathcal{A}(\hat{\Gamma}) := \begin{bmatrix} \hat{A}^{-1} & -\hat{A}^{-1}\hat{B} \\ \hat{C}\hat{A}^{-1} & \hat{D} - \hat{C}\hat{A}^{-1}\hat{B} \end{bmatrix}, \qquad \text{for} \qquad \hat{\Gamma} = \begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix}. \tag{4.19}$$

Proof. (Step 1) We first find an expression for $\mathrm{D}\widehat{\Phi}$ in terms of $\mathrm{D}\Phi$. From the definition of $\hat{q}(Q,p)$, and

$$\hat{P}(Q,p) = P(\hat{q}(Q,p),p),$$

we learn

$$\begin{aligned} \big(Q_q, Q_p\big) &= \big(\hat{q}_Q^{-1}, -\hat{q}_Q^{-1}\ \hat{q}_p\big), \\ \big(P_q, P_p\big) &= \big(\hat{P}_Q\ \hat{q}_Q^{-1}, \hat{P}_p - \hat{P}_Q\ \hat{q}_Q^{-1}\ \hat{q}_p\big). \end{aligned}$$

From

$$\widehat{\Phi}(Q,p) = \big(\hat{q}, \hat{P}\big)(Q,p)$$

we learn that if

$$\begin{aligned} \Gamma : = \mathrm{D}\Phi &= \begin{bmatrix} A & B \\ C & D \end{bmatrix} := \begin{bmatrix} Q_q & Q_p \\ P_q & P_p \end{bmatrix}, \\ \hat{\Gamma} : = \mathrm{D}\hat{\Phi} &= \begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix} := \begin{bmatrix} \hat{q}_Q & \hat{q}_p \\ \hat{P}_Q & \hat{P}_p \end{bmatrix}, \end{aligned}$$

then $\Gamma = \mathcal{A}(\hat{\Gamma})$, with $\mathcal{A}(\hat{\Gamma})$ given by (4.19). Clearly,

$$\rho(x, X, \Gamma)\ \mathrm{d}X\ \mathrm{d}\Gamma = m(\hat{x}, \hat{X}, \hat{\Gamma})\ \mathrm{d}\hat{X}\ \mathrm{d}\hat{\Gamma}, \tag{4.20}$$

where

$$m(\hat{x}, \hat{X}, \hat{\Gamma}) = m(Q,p,q,P,\hat{\Gamma}) = \rho\big(q,p,Q,P,\mathcal{A}(\hat{\Gamma})\big).$$

From

$$\mathrm{d}X = \mathrm{d}q\ \mathrm{d}p = \big|\det(\hat{q}_Q)\big|\ \mathrm{d}Q\ \mathrm{d}p,$$

we deduce

$$\mathrm{d}X = \big|\det \hat{A}\big|\ \mathrm{d}\hat{X}. \tag{4.21}$$

It remains to express $\mathrm{d}\Gamma$ in terms of $\mathrm{d}\hat{\Gamma}$. We write $\mathrm{d}A,\ \mathrm{d}B,\ \mathrm{d}C$, and $\mathrm{d}D$ for the volume forms in $\mathbb{R}^{d\times d}$ associated with $A,\ B,\ C$ and D. Here by $\mathrm{d}A$ we really mean

$$\mathrm{d}A = \big(\mathrm{d}a_{11} \wedge \mathrm{d}a_{12} \wedge \cdots \wedge \mathrm{d}a_{1d}\big) \wedge \cdots \wedge \big(\mathrm{d}a_{d1} \wedge \mathrm{d}a_{12} \wedge \cdots \wedge \mathrm{d}a_{dd}\big),$$

where $a_{ij},\ i,j = 1,\dots d$, denote the entries of A. In the same manner we define $\mathrm{d}B,\ \mathrm{d}C$, and $\mathrm{d}D$. Analogously $\mathrm{d}\hat{A},\ \mathrm{d}\hat{B},\ \mathrm{d}\hat{C}$, and $\mathrm{d}\hat{D}$ are defined. On account of (4.20), (4.21), the proof is complete if we show

$$(\mathrm{d}A) \wedge (\mathrm{d}B) \wedge (\mathrm{d}C) \wedge (\mathrm{d}D) = \pm(\det \hat{A})^{-4d}\ (\mathrm{d}\hat{A}) \wedge (\mathrm{d}\hat{B}) \wedge (\mathrm{d}\hat{C}) \wedge (\mathrm{d}\hat{D}). \tag{4.22}$$

Note the equality in (4.22) is for volume forms, where as in (4.20) the expressions $dX\ d\Gamma$ and $d\hat{X}\ d\hat{\Gamma}$ refer to the measures in $\mathbb{R}^{2d} \times \mathbb{R}^{2d\times 2d}$. Since we are interested in the probability densities of our random variables, we do not keep track of signs of the volume forms that will appear in the subsequent calculations.

(Step 2) We first try to express

$$\mathrm{d}A = \mathrm{d}(\hat{A}^{-1})$$

in terms of $\mathrm{d}\hat{A}$. By our Lemma 4.1 below, we have

$$\mathrm{d}A = \pm(\det \hat{A})^{-2d}\, \mathrm{d}\hat{A}. \tag{4.23}$$

We next study

$$(\mathrm{d}A) \wedge (\mathrm{d}B) = \big(\mathrm{d}\hat{A}^{-1}\big) \wedge \mathrm{d}\big(-\hat{A}^{-1}\hat{B}\big).$$

As we apply the exterior derivative on $\hat{A}^{-1}\hat{B} = A\hat{B}$, we can treat A^{-1} as a constant because of the wedge product with dÂ (we are using $\mathrm{d}\hat{a}^{ij} \wedge \mathrm{d}\hat{a}^{ij} = 0$).

Lemma 4.1 below allows us write

$$\mathrm{d}(E\hat{B}) = \pm(\det E)^{d}\, \mathrm{d}\hat{B}, \tag{4.24}$$

for a constant matrix E. Hence

$$\begin{aligned}(\mathrm{d}A) \wedge \mathrm{d}B) &= \pm(\det \hat{A})^{-2d}\, (\mathrm{d}\hat{A}) \wedge \big(\mathrm{d}\big(\hat{A}^{-1}\hat{B}\big)\big) \\ &= \pm(\det \hat{A})^{-2d}\, \Big(\det\big(\hat{A}^{-1}\big)\Big)^{d} (\mathrm{d}\hat{A}) \wedge (\mathrm{d}\hat{B}) \\ &= \pm(\det \hat{A})^{-3d}\, (\mathrm{d}\hat{A}) \wedge (\mathrm{d}\hat{B}).\end{aligned}$$

In the same fashion,

$$(\mathrm{d}A) \wedge (\mathrm{d}B) \wedge (\mathrm{d}C) = \pm(\det \hat{A})^{-4d}\, (\mathrm{d}\hat{A}) \wedge (\mathrm{d}\hat{B}) \wedge (\mathrm{d}\hat{C}).$$

From this and

$$(\mathrm{d}A) \wedge (\mathrm{d}B) \wedge (\mathrm{d}C) \wedge (\mathrm{d}D) = (\mathrm{d}A) \wedge (\mathrm{d}B) \wedge (\mathrm{d}C) \wedge (\mathrm{d}\hat{D}),$$

we can readily derive (4.22). □

It remains to verify (4.23) and (4.24).

Lemma 4.1. *The following statements hold.*

(i) *Given $E \in \mathbb{R}^{d\times d}$, consider the map*

$$\zeta\colon \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d},$$

such that $\zeta(Z) = EZ$. If we write $\mathrm{d}Z$ for the volume form of $\mathbb{R}^{d\times d}$, then

$$\zeta^*(\mathrm{d}Z) = \pm(\det E)^{d}\, \mathrm{d}Z. \tag{4.25}$$

A similar formula is valid if $\zeta(Z) = ZE$.

(ii) *If $\eta(Z) = Z^{-1}$, then*

$$\eta^*(\mathrm{d}Z) = \pm(\det Z)^{-2d}\, \mathrm{d}Z. \tag{4.26}$$

Proof. (i) Given

$$Z = [z_{ij}]_{i,j=1}^{d}, \quad EZ = [\zeta_{ij}]_{i,j=1}^{d},$$

we can write

$$
\begin{aligned}
\bigwedge_{j=1}^{d}\bigwedge_{i=1}^{d} \mathrm{d}\zeta_{ij} &= \bigwedge_{j=1}^{d}\bigwedge_{i=1}^{d}\left(\sum_{k=1}^{d} e_{ik}\ \mathrm{d}z_{kj}\right)\\
&= \bigwedge_{j=1}^{d}\sum_{\sigma\in S_d}\prod_{i=1}^{d} e_{i\sigma(i)}\ \bigwedge_{j=1}^{d}\mathrm{d}z_{\sigma(i)j}\\
&= \bigwedge_{j=1}^{d}\sum_{\sigma\in S_d}\varepsilon(\sigma)\prod_{k=1}^{d} e_{i\sigma(i)}\ \bigwedge_{i=1}^{d}\mathrm{d}z_{ij}\\
&= (\det E)^d\bigwedge_{j=1}^{d}\bigwedge_{i=1}^{d}\mathrm{d}z_{ij},
\end{aligned}
$$

where S_d denotes the set of permutations of $\{1,\dots,d\}$. This completes the proof of (4.25).

(ii) Let us write $\eta^*(\mathrm{d}Z)=\alpha(Z)\ \mathrm{d}Z$. Fix E as in part (i). Observe

$$
\begin{aligned}
(\eta\circ\zeta)^*(\mathrm{d}Z) &= \zeta^*\eta^*(\mathrm{d}Z)\\
&= \zeta^*\big(\alpha(Z)\mathrm{d}Z\big)\\
&= \pm(\det E)^d\ \alpha(EZ)\ \mathrm{d}Z. \qquad (4.27)
\end{aligned}
$$

On the other hand, if we set $\xi(Z)=ZE^{-1}$, then $\eta\circ\zeta=\xi\circ\eta$, and

$$
\begin{aligned}
(\xi\circ\eta)^*(dZ) &= \eta^*\xi^*(\mathrm{d}Z)\\
&= \pm(\det E)^{-d}\ \eta^*(\mathrm{d}Z)\\
&= (\det E)^{-d}\ \alpha(Z)\ \mathrm{d}Z.
\end{aligned}
$$

From this and (4.27) we deduce

$$\alpha(EZ)=\pm(\det E)^{-2d}\ \alpha(Z),$$

which in turn implies that

$$a(E)=(\det E)^{-2d}\alpha(\mathrm{I}), \qquad (4.28)$$

where $\mathrm{I}=\mathrm{I_d}$ denotes the $d\times d$ identity matrix. Furthermore, since $\eta\circ\eta=\mathrm{id}$, we know

$$
\begin{aligned}
\mathrm{d}Z &= (\eta\circ\eta)^*(\mathrm{d}Z)\\
&= \eta^*\big(\alpha(Z)\ \mathrm{d}Z\big)\\
&= \alpha(Z^{-1})\alpha^*(\mathrm{d}Z)\\
&= \alpha(Z^{-1})\alpha(Z)\ \mathrm{d}Z,
\end{aligned}
$$

which in particular implies that $\alpha(\mathrm{I})^2=1$. From this and (4.28) we can readily deduce (4.26). □

Remark 4.4 We express our formula (4.18) in terms of the density of $\big(\Phi(x),\mathrm{D}\Phi(\mathrm{x})\big)$ with no reference to the stationarity of the process $\omega=\mathcal{F}(\Phi)$. In fact, if the law of pair

$(\omega(x), \mathrm{D}\omega(\mathrm{x}))$ with respect to the measure $\mathcal{Q}$ has a density $\gamma(X,\Gamma)$, then it does not depend on x by stationarity, and ρ can be expressed in terms of γ by the following formula:

$$\rho(x,X,\Gamma) = \gamma(X-x,\Gamma-\mathrm{I}),$$

where I is $(2d)\times(2d)$ identity matrix.

5. The density of fixed points in the case of quasiperiodic maps

Theorem 3.4 and Proposition 4.1 reduce the counting of the fixed points to the evaluation of $\mathbb{E}\mathcal{N}_U(0,\omega)$.

If we take the expected value of both sides of (4.10), and use the stationarity we always have

$$\int \zeta(a)\big[\mathbb{E}\mathcal{N}_U(a,\omega)\big]\ da = \int \zeta(\hat{f}(\omega))\ \mathbb{1}\big(\boldsymbol{D}\hat{f}(\omega)\in U\big)\ |\det \boldsymbol{D}\hat{f}(\omega)|\ \mathcal{Q}(\mathrm{d}\omega), \tag{5.1}$$

for every bounded continuous function ζ. (See Definition 3.1(i) for the definition of $\boldsymbol{D}$.)

If we can choose ζ to be the delta function at 0, then we have the informal expression (4.7) for $\mathbb{E}\mathcal{N}_U(0,\omega)$. Theorem 4.2 offers an explicit formula for the right-hand side of (4.7) in terms of the density of the pair $(\hat{f},\boldsymbol{D}\hat{f}))$. The existence of a density is rather a restrictive requirement and not valid for many examples of interest. In this section, we offer a new explicit formula for $\mathbb{E}\mathcal{N}_U(0,\omega)$ when Φ is quasi periodic.

Recall that $\hat{f} = \widehat{\boldsymbol{\nabla}}\hat{w}$. To simplify our notation, we may instead consider the zero set of the function $\hat{g} = \boldsymbol{\nabla}\hat{w}$. Before stating our first main result, let us review the setting we will be working with.

Setting 5.1 Given a C^2 function

$$\hat{w}:\mathbb{T}^N\to\mathbb{R},$$

and $N\times n$ matrix A, define

$$w(x) = w(x,\omega) = \hat{w}(\Theta_{Ax}\omega),$$

$\hat{g} = \boldsymbol{\nabla}\hat{w}$, and $g = \nabla w$. We assume that the $n\times n$ matrix

$$E := A^*A$$

is of full rank. Let us write

$$\mathcal{N}_U(a,\omega) = \sharp\left\{x\in[0,1]^n:\ g(x,\omega)=a,\ \mathrm{D}g(x,\omega))\in U\right\}.$$

Observe

$$\hat{g}(\omega) = \nabla\hat{w}(\omega)A, \qquad \boldsymbol{D}\hat{g}(\omega) = A^*\mathrm{D}^2\hat{w}(\omega)A, \tag{5.2}$$

where $\nabla\hat{w}$ and $\mathrm{D}\hat{g}$ represent the standard derivatives of $\hat{w}$ and $\hat{g}$ (as opposed to the $\boldsymbol{\nabla}$ and $\boldsymbol{D}$ which denote the differentiation in the sense of Definition 3.1(i)). We write $\mathcal{Q}$ for the Lebesgue measure on $\mathbb{T}^N$. □

We are now ready to present our formula for $\mathbb{E}\mathcal{N}_U(0,\omega)$, namely (5.3) below. The proof of (5.3) is similar to the proof of (4.9). One of the main tool we use is the celebrated Coarea Formula that we now recall; given a measurable set V, a C^1 function

$$S:\mathbb{R}^N\to\mathbb{R}^n,$$

with $N > n$, and a non-negative measurable function $T : \mathbb{R}^N \to \mathbb{R}$, we have

$$\int_V T(x)\ (\mathcal{J}S)(x)\ \mathrm{d}x = \int_{\mathbb{R}^n} \left[\int_{V\cap S^{-1}(a)} T\ \mathrm{d}\sigma_{N-n}\right]\ \mathrm{d}a, \tag{5.3}$$

where

$$(\mathcal{J}S)(x) = \Big(\det\big((\mathrm{D}S)(x)(\mathrm{D}S)^*(x)\big)\Big)^{1/2},$$

(here A^* denotes the transpose of A), and σ_{N-n} denotes the $N-n$ dimensional (Hausdorff) measure. For our purposes, we wish to choose

$$T(x) = W(x)(\mathcal{J}S)(x)^{-1},$$

in (5.3). This function is well-defined so long as

$$V \subset \Sigma := \big\{X \in \mathbb{R}^N :\ (\mathcal{J}S)(x) \neq 0\big\}.$$

For such a choice of T, (5.3) reads as

$$\int_V W(x)\ \mathrm{d}x = \int_{\mathbb{R}^n} \left[\int_{V\cap S^{-1}(a)} W(\mathcal{J}S)^{-1}\ \mathrm{d}\sigma_{N-n}.\right]\ \mathrm{d}a, \tag{5.4}$$

We are now ready to state and prove the main result of this section.

Theorem 5.1. *Let $\hat{w}$ and $\hat{g}$ be as in Setting 5.1. Then*

$$\mathbb{E}\mathcal{N}_U(0,\omega) = \int_{\mathbb{T}^N} \mathcal{N}_U(0,\omega)\ \mathrm{d}\omega = \int_{\Lambda(0)} \frac{\big|\det(A^*\mathrm{D}^2\hat{w}(\omega)A)\big|}{\det(A^*(\mathrm{D}^2\hat{w}(\omega))^2A)^{1/2}}\ \sigma_{N-n}(\mathrm{d}\omega), \tag{5.5}$$

where,

$$\Lambda(a) = \big\{\omega \in \mathbb{T}^N :\ \nabla\hat{w}(\omega)A = a,\ A^*\mathrm{D}^2\hat{w}(\omega)A \in U,\ \det\big(A^*\mathrm{D}^2\hat{w}(\bar{\omega})A\big) \neq 0\big\}.$$

Moreover, if there exists $\bar{\omega}$ such that

$$\nabla\hat{w}(\bar{\omega})A = 0, \qquad A^*\mathrm{D}^2\hat{w}(\bar{\omega})A \in U,$$

and

$$A^*\mathrm{D}^2\hat{w}(\bar{\omega})A$$

is invertible, then the right-hand side of (5.4) *is nonzero.*

Proof. (Step 1) We may apply the area formula to the function g to derive the analogue of (5.1),

$$\int_{\mathbb{R}^n} \zeta(a)\big[\mathbb{E}\mathcal{N}_U(a,\omega)\big]\ \mathrm{d}a = \int_{\mathbb{T}^N} \zeta(\hat{g}(\omega))\ \mathbb{1}\big(\boldsymbol{D}\hat{g}(\omega) \in U\big)\ |\det \boldsymbol{D}\hat{g}(\omega)|\ \mathrm{d}\omega, \tag{5.6}$$

for every continuous function ζ. We wish apply Coarea Formula (5.4) to the right-hand side of (5.6), for the choices of

$$S(\omega) = \hat{g}(\omega), \qquad W(\omega) = \zeta(\hat{g}(\omega))\ ,$$

$$V = \big\{\omega :\ \boldsymbol{D}\hat{g}(\omega) \in U\big\} \cap \Sigma, \tag{5.7}$$

where

$$\Sigma = \big\{\omega \in \mathbb{T}^N :\ (\mathcal{J}\hat{g})(\omega) \neq 0\big\}.$$

Observe ,

$$\mathrm{D}\hat{g}(\omega) = A^*\mathrm{D}^2\hat{w}(\omega), \qquad (\mathcal{J}\hat{g})(\omega) = \det(A^*(\mathrm{D}^2\hat{w}(\omega))^2A)^{1/2}.$$

On the other hand, if $\omega \notin \Sigma$, then there exists a nonzero vector b such that $A^*C^2Ab = 0$, for $C = \mathrm{D}^2\hat{w}(\omega)$. As a consequence, $A^*CAb = 0$ because

$$0 = A^*C^2Ab \cdot b = |CAb|^2.$$

From this we learn

$$\big\{x \in \mathbb{T}^N : \ \det(A^*\mathrm{D}^2\hat{w}(\omega)A) \neq 0\big\} \subset \Sigma.$$

Because of this, the right-hand side of (5.6) equals to

$$\int_\Sigma \zeta(\hat{g}(\omega)) \ \mathbb{1}\big(\boldsymbol{D}\hat{g}(\omega) \in U\big) \ |\det \boldsymbol{D}\hat{g}(\omega)| \ \mathrm{d}\omega. \tag{5.8}$$

We now apply Coarea Formula (5.4) to (5.8), for the choices of (5.7), to assert

$$\int_{\mathbb{R}^n} \zeta(a)\big[\mathbb{E}\mathcal{N}_U(a,\omega)\big] \ \mathrm{d}a = \int_{\mathbb{R}^n} \zeta(a) \ \left[\int_{\Lambda(a)} \frac{\big|\det(A^*\mathrm{D}^2\hat{w}(\omega)A)\big|}{\det(A^*(\mathrm{D}^2\hat{w}(\omega))^2A)^{1/2}} \ \sigma_{N-n}(\mathrm{d}\omega)\right] \ \mathrm{d}a. \tag{5.9}$$

From this we deduce

$$\mathbb{E}\mathcal{N}_U(a,\omega) = \int_{\Lambda(a)} \frac{\big|\det(A^*\mathrm{D}^2\hat{w}(\omega)A)\big|}{\det(A^*(\mathrm{D}^2\hat{f}(\omega))^2A)^{1/2}} \ \sigma_{N-n}(\mathrm{d}\omega) =: G(a), \tag{5.10}$$

for Lebesgue almost all $a \in \mathbb{R}^n$. We wish to show that (5.7) holds for *all* a. We achieve this by verifying the continuity of the function G, and a repetition of some of the steps of the proof of Theorem 4.2.

(Step 2) With a verbatim argument as in the proof of Theorem 4.2, we can show that there exists a collection of measurable sets $\big\{\Omega(\delta) : \ \delta > 0\big\}$, with $\Omega(\delta) \subset \Omega = \mathbb{T}^N$, such that the following statements hold:

(i) $\lim_{\delta\to 0} \mathcal{Q}\big(\Omega(\delta)\big) = 1$.
(ii) If $\omega \in \Omega(\delta)$, and $|u| \leqslant \delta$, then $\mathcal{N}_U(u,\omega) = \mathcal{N}_U(0,\omega)$.
(iii) If $\omega \in \Omega(\delta)$, and $\varepsilon \in (0,\delta)$, then

$$\mathcal{N}_U(0,\omega) = \frac{1}{|\mathrm{B}_\varepsilon(0)|} \int_{[0,1]^n} \mathbb{1}(|g(x,\omega)| \leqslant \varepsilon) \ \mathbb{1}\big(\mathrm{D}g(x,\omega) \in U\big) \ |\det \mathrm{D}g(x,\omega)| \ \mathrm{d}x. \tag{5.11}$$

Let us demonstrate how the continuity of G, (i)-(iii), and (5.10) for almost all a, imply that (5.10) holds for $a = 0$ (continuity at any other a can be shown in exactly the same way).

As in the proof of Theorem 4.2 we use (5.11) to assert that when $\varepsilon \in (0,\delta)$, the expression

$$\mathbb{E} \ \mathcal{N}_U(0,\omega)\mathbb{1}\big(\omega \in \Omega(\delta)\big),$$

equals

$$
\begin{aligned}
&\mathbb{E}\ \frac{1}{|\mathrm{B}_\varepsilon(0)|}\int_{[0,1]^n} \mathbb{1}\big(|g(x,\omega)|\leqslant\varepsilon,\ \mathrm{D}g(x,\omega)\in U\big)\ |\det\mathrm{D}g(x,\omega)|\ \mathrm{d}x\ \mathbb{1}\big(\omega\in\Omega(\delta)\big)\\
&\quad\leqslant \mathbb{E}\ \frac{1}{|\mathrm{B}_\varepsilon(0)|}\int_{[0,1]^n} \mathbb{1}\big(|g(x,\omega)|\leqslant\varepsilon,\ \mathrm{D}g(x,\omega)\in U\big)\ |\det\mathrm{D}g(x,\omega)|\ \mathrm{d}x\\
&\quad= \frac{1}{|\mathrm{B}_\varepsilon(0)|}\mathbb{E}\ \mathbb{1}\big(|g(0,\omega)|\leqslant\varepsilon,\ \mathrm{D}g(0,\omega)\in U\big)\ |\det\mathrm{D}g(0,\omega)|\\
&\quad= \frac{1}{|\mathrm{B}_\varepsilon(0)|}\int_{\mathbb{T}^N} \mathbb{1}\big(|\hat g(\omega)|\leqslant\varepsilon,\ \boldsymbol{D}\hat g(\omega)\in U\big)\ |\det\boldsymbol{D}\hat g(\omega)|\ \mathrm{d}\omega\\
&\quad= \frac{1}{|\mathrm{B}_\varepsilon(0)|}\int_{\mathrm{B}_\varepsilon(0)} G(b)\ \mathrm{db},
\end{aligned}
$$

where we used the stationarity for the first equality, and Coarea Formula for the last equality.

We then send $\varepsilon\to 0$, and $\delta\to 0$ (in this order), and use the continuity of G to deduce

$$
\mathbb{E}\ \mathcal{N}_U(0,\omega)\leqslant G(0). \tag{5.12}
$$

On the other-hand, by the validity of (5.10) for almost all points, we can find a sequence $a_k\to 0$ such that

$$
\mathbb{E}\ \mathcal{N}_U(a_k,\omega)=G(a_k). \tag{5.13}
$$

As in the proof of Theorem 4.2, we use (5.12), the continuity of G, (5.13), and (i)-(ii), to argue

$$
\begin{aligned}
\mathbb{E}\ \mathcal{N}_U(0,\omega) &\leqslant G(0)=\lim_{k\to\infty} G(a_k)\\
&= \lim_{k\to\infty}\mathbb{E}\ \mathcal{N}_U(a_k,\omega)\\
&= \lim_{k\to\infty}\lim_{\delta\to 0}\mathbb{E}\ \mathcal{N}_U(a_k,\omega)\ \mathbb{1}\big(\omega\in\Omega(\delta)\big)\\
&= \lim_{k\to\infty}\lim_{\delta\to 0}\mathbb{E}\ \mathcal{N}_U(0,\omega)\ \mathbb{1}\big(\omega\in\Omega(\delta)\big)\\
&= \mathbb{E}\ \mathcal{N}_U(0,\omega).
\end{aligned}
$$

Since we must have equality for the inequalities in the above display, we arrive at

$$
\mathbb{E}\ \mathcal{N}_U(0,\omega)=G(0),
$$

which is (5.5). It remains to verify the continuity of the function G.

(Step 3) As a preparation for the proof of the continuity, we first study the level set $\Lambda(a)$, which is a subset of Σ. Observe that if $\omega\in\Sigma$, then

$$
\det\big(A^*C^2A\big)=\det(M^*M)\neq 0,
$$

where $C=C(\omega)=D^2\hat w$ as before, and $M=CA$. Note that $M\in\mathbb{R}^{N\times n}$, with $N>n$. Hence, we may apply Cauchy-Binet Formula, to write

$$
\det(M^*M)=\sum_{I\in\mathcal{I}}(\det M_I)^2, \tag{5.14}
$$

where $\mathcal{I}$ denotes the collection of sets $I\subset\{1,2,\dots,N\}=:[N]$ such that $\sharp I=n$, and for

$$
M=[m_{ij}]_{i\in[N],j\in[n]},
$$

by M_I we mean the $n \times n$ submatrix of M, given by

$$M_I = [m_{ij}]_{i \in I, j \in [n]}.$$

Let us write

$$\omega = (\omega_1, \dots, \omega_N)$$

for the coordinates of $\omega \in \mathbb{T}^N$ (regarding $\mathbb{T}^N = [0,1]^N$, with $0 = 1$). We also write

$$\nabla_I = \left(\frac{\partial}{\partial \omega_j} : \ j \in I \right),$$

and Σ_I for the set of $\omega \in \mathbb{T}^N$ such that

$$\nabla_I \hat{g}(\omega) = \nabla_I \nabla \hat{w} A = \left[\sum_{k=1}^{N} \hat{w}_{\omega_i \omega_k} a_{kj} \right]_{i \in I, j \in [n]} = [M_{ij}]_{i \in I, j \in [n]} = M_I,$$

is invertible. From (5.14) we learn

$$\Sigma = \cup_{I \in \mathcal{I}} \Sigma_I. \tag{5.15}$$

We now examine the set $\Lambda(a) \cap \Sigma_I$, for each $I \in \mathcal{I}$.

Without loss of generality, we may assume that $I = [n]$. Let us examine the set Σ_I, when $I = [n]$. Regarding $\omega \in \mathbb{T}^N$, as a point in $[0,1]^N$, we may write $\omega = (\omega^1, \omega^2) \in \mathbb{R}^n \times \mathbb{R}^{N-n}$. For $\omega \in \Sigma_{[n]}$, we know

$$\nabla_{\omega^1} \hat{g}(\omega) = \partial^2_{\omega^1 \omega} \hat{w}(\omega) A$$

is invertible. Fix

$$\bar{\omega} = (\bar{\omega}^1, \bar{\omega}^2) \in \Sigma_{[n]} \cap \Lambda(a).$$

If we define

$$F(\omega^1, \omega^2) = (\hat{g}(\omega), \omega^2),$$

then $\mathrm{D}F(\bar{\omega})$ is invertible, and for b in a neighborhood of

$$\bar{b} = (\hat{g}(\bar{\omega}), \bar{\omega}^2) = (a, \bar{\omega}^2),$$

we can define $F^{-1}(b)$. This, and the compactness of $\Lambda(a)$ allow us to find an open covering

$$\Sigma_I \cap \Lambda(a) \subset \bigcup_{i=1}^{\alpha(I)} \Sigma_I^i,$$

such that for every $i \in \{1, \dots, \alpha(I)\}$,

$$F(\Sigma_I^i) = \mathrm{B}_{\delta_i}(a) \times U_i^I,$$

for some $\delta_i > 0$, and some open set $U_i^I \subset \mathbb{R}^{N-n}$, and we can make sense of

$$F^{-1}(a, \omega^2) = (R_I^i(a, \omega^2), \omega^2),$$

with

$$R_I^i : \mathrm{B}_{\delta_i}(a) \times U_I^i \to \mathbb{R}^n,$$

a C^1-function. In other words,

$$F\big(R_I^i(a, \omega^2), \omega^2\big) = a,$$

and the graph of $R_I^i(a, \cdot)$ yields a parametrization of $\Lambda(a) \cap \Sigma_I^i$.

To ease the notation, let us write $X : \mathbb{T}^N \to \mathbb{R}$, for the integrand of the integral that appeared in the definition of G in (5.10). We may use a partition of unity

$$\{\varphi_I^i : \ i = 1, \dots, \alpha(I), \ I \in \mathcal{I}\},$$

associated with the covering $\{\Sigma_I^i : \ i = 1, \dots, \alpha(I), \ I \in \mathcal{I}\}$, to write

$$\begin{aligned} G(a) &= \int_{\Lambda(a)} X(\omega) \ \sigma_{N-n}(\mathrm{d}\omega) \\ &= \sum_{I\in\mathcal{I}} \sum_{i=1}^{\alpha(I)} \int_{\Lambda(a)\cap\Sigma_I^i} \big(\varphi_I^i X\big)(\omega) \ \sigma_{N-n}(\mathrm{d}\omega) \\ &= \sum_{I\in\mathcal{I}} \sum_{i=1}^{\alpha(I)} \int_{U_I^i} \big(\varphi_I^i X\big)\big(R_I^i(a,\omega^2),\omega^2\big) \ J_I^i(a,\omega^2) \ \mathrm{d}\omega^2, \end{aligned}$$

where J_I^i is the corresponding Jacobian factor:

$$J_I^i(a,\omega^2) = \det\big(\mathrm{I_{N-n}} + \mathrm{E_I^i(a,\omega_2)}\big)\big)^{1/2},$$

where $\mathrm{I_{N-n}}$ is the identity matrix of $\mathbb{R}^{N-n}$, and $E_I^i \in \mathbb{R}^{(N-n)\times(N-n)}$ is a matrix with (r,s) entry

$$\big(E_I^i\big)_{rs} = \frac{\partial R_I^i}{\partial \omega_2^r} \cdot \frac{\partial R_I^i}{\partial \omega_2^s}.$$

From this representation, it is not hard to deduce the continuity of the map $a \mapsto G(a)$.

Finally observe that when $a = 0$, and $\mathrm{D}F(\bar{\omega})$ is invertible, then $\Lambda(0)$ contains an $N-n$ dimensional surface (namely the graph of $R_I^i(0,\cdot)$). This implies that the right-hand side of (5.5) is not zero. □

Formula (5.5) offers an explicit expression for the fix points density when $\mathcal{F}(\widehat{\Phi}) = \widehat{\Phi} - \mathrm{id}$ is quasiperiodic. We still need to show that indeed a quasiperiodic $\mathcal{F}(\Phi)$ yields a quaiperiodic $\mathcal{F}(\widehat{\Phi})$. For this we need a refinement of Proposition 3.1(i).

Proposition 5.1. *Assume that $x \mapsto \Phi(x,\omega)$ is a symplectic twist diffeomorphism such that $\Phi(x,\omega) = x + K(\Theta_{Ax}\omega)$, for a continuous function $K : \mathbb{T}^N \to \mathbb{R}^{2d}$. Then $\mathcal{F}(\widehat{\Phi})$ is quasiperiodic.*

Proof. Recall $\theta_{(q,p)} = \eta_p \circ \tau_q = \tau_q \circ \eta_p$. We also write

$$\Phi(q,p,\omega) = \big(q + \alpha(\tau_q\eta_p\omega), p + \beta(\tau_q\eta_p\omega)\big).$$

The twist condition means that the map

$$q \mapsto \gamma(q,\omega) := q + \alpha(\tau_q\omega),$$

is a diffeomorphism. If we write $\gamma^{-1}(Q,\omega)$ for its inverse, and set

$$\hat{\alpha}(\omega) = \gamma^{-1}(0,\omega), \tag{5.16}$$

then we can then write

$$\hat{\alpha}(\omega) + \alpha\big(\tau_{\hat{\alpha}(\omega)} \ \omega\big) = 0.$$

From

$$q + \alpha\big(\tau_q\omega\big) = Q \quad \Leftrightarrow \quad q - Q + \alpha\big(\tau_{q-Q}\tau_Q\omega\big) = 0,$$

we deduce

$$\gamma^{-1}(Q,\omega) = q = Q + \hat\alpha\big(\tau_Q\omega\big).$$

From this and the definition of $(\hat q, \hat P)$ we learn

$$\begin{aligned}\hat q(Q,p) &= Q + \hat\alpha\big(\tau_Q\eta_p\omega\big) = Q + \hat\alpha\big(\theta_{\hat x}\omega\big),\\ \hat P(Q,p) &= P\big(\hat q(Q,p),p\big) = p + \beta\big(\tau_Q\eta_p\tau_{\hat\alpha(\theta_{\hat x}\omega)}\omega\big)\\ &= p + \hat\beta\big(\theta_{\hat x}\ \omega\big),\end{aligned}$$

where $\hat x = (Q,p)$, and

$$\hat\beta(\omega) = \beta\big(\tau_{\hat\alpha(\omega)}\omega\big). \tag{5.17}$$

In summary,

$$\mathcal{F}(\widehat\Phi)(\hat x,\omega) = \gamma(\theta_{\hat x}\omega), \quad \text{where} \quad \gamma = (\hat\alpha,\hat\beta), \tag{5.18}$$

with $\hat\alpha$ and $\hat\beta$ as in (5.16) and (5.17).

We now apply our general formula (5.18) to the case of a quasiperiodic $\mathcal{F}(\Phi)$. In this case, $\omega \in \mathbb{T}^N$, $\alpha,\beta : \mathbb{T}^N \to \mathbb{R}^d$ are two continuous functions, and we have two $N\times d$ matrices A^1 and A^2, such that $A = [A^1, A^2]$, and

$$\begin{aligned}\tau_q\omega &= \Theta_{A^1q}\ \omega = \omega + A^1q \mod 1,\\ \eta_p\omega &= \Theta_{A^2p}\ \omega = \omega + A^2p \mod 1.\end{aligned}$$

Analogously $\gamma = (\hat\alpha,\hat\beta) : \mathbb{T}^N \to \mathbb{R}^{2d}$ is a continuous function such that (5.18) holds. This certainly implies the quasiperiodicity of $\mathcal{F}(\widehat\Phi)$. □

5.1. **Proof of Theorem 2.5.** We are now ready to offer a more precise statement of Theorem 2.5 and give a proof.

Theorem 5.2. *Assume that $x \mapsto \Phi(x,\omega)$ is a C^2 symplectic twist diffeomorphism such that*

$$\Phi(x,\omega) = x + K(\Theta_{Ax}\omega),$$

for a C^1 function $K : \mathbb{T}^N \to \mathbb{R}^{2d}$. Let $\mathcal{Q}$ denotes the Lebesgue measure on $\mathbb{T}^N$, and assume that $K \in \widehat{\mathfrak{H}}^{-1}(\mathcal{Q})$, and

$$\int_{\mathbb{T}^N} K\ \mathrm{d}\mathcal{Q} = 0. \tag{5.19}$$

Then the set

$$\big\{x \in \mathbb{R}^{2d} :\ \Phi(x,\omega) = x\big\},$$

is of positive (possibly infinite) density, $\mathcal{Q}$-almost surely.

Proof. Our regularity assumption $K \in \widehat{\mathfrak{H}}^{-1}(\mathcal{Q}) \cap \mathrm{C}^1$, and (5.19) allow us to apply Theorem 3.4 to deduce the existence of a stationary generating function $w(x,\omega) = \hat w(\theta_x\omega)$. By Proposition 5.1, The map

$$\mathcal{F}(\widehat\Phi) = \widehat\nabla w,$$

is quasiperiodic. As we illustrated in Example 3.6, the quasiperiodicity of $\widehat\nabla w$ implies the quasiperiodicity of $w(x,\omega)$. Proposition 4.2 guarantees the $\mathcal{Q}$-almost sure existence of a density (4.2) for the set $Z(\omega)$. When the right-hand side of (4.2) is infinite, there is nothing to prove. When the right-hand side of (4.2) is finite, we apply Theorem 5.1 to find an explicit expression given by (5.5) for the density. By choosing U to be the set of all symmetric matrices (or even the set of positive or negative matrices), we can guarantee the positivity of the density of the set $Z(\omega)$, $\mathcal{Q}$-almost surely. $\square$

Remark 5.3 We refer to Example 3.6 for sufficient conditions that would guarantee $K \in \widehat{\mathfrak{H}}^{-1}(\mathcal{Q})$.

6. Stationary Hamiltonian ODEs

In this section we study the time one map ϕ^H for a Hamiltonian function that is selected randomly according to a θ-invariant probability measure $\mathbb{P}$ on $\mathcal{H}$. As it is well-known, there is a one-to-one correspondence between 1-periodic orbits of the Hamiltonian vector field $X_H(x,t) = J\nabla H(x,t)$ and the fixed points of ϕ^H. The map $H \mapsto \phi^H$ pushes forward $\mathbb{P}$ to a probability measure $\mathcal{P}$ on $\mathcal{S}$. To count the fixed points of ϕ^H, we wish to apply Theorem 3.4. For this, we need to make some regularity assumptions on H, and verify the applicability of Theorem 3.4. Let us first make a useful definition concerning the regularity of Hamiltonian functions.

Definition 6.1. Let us write $\mathcal{C}^2(\ell)$ for the set of continuous maps

$$H : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$$

such that H is twice differentiable in x, and

$$\|\nabla_x H\|_{\mathrm{C}^0},\ \|\mathrm{D}_x^2 H\|_{\mathrm{C}^0} \leqslant \ell. \tag{6.1}$$

To ease the notation, we will write ∇ and D for ∇_x and $\mathrm{D_x}$, respectively.

In the next Proposition, we verify various properties of ϕ^H in terms of the properties of H. This will prepare us to apply Theorem 3.4 to ϕ^H, where H is selected according to the θ-invariant measure $\mathbb{P}$.

Proposition 6.1. *The following statements hold:*

(i) *We have the following equalities*

$$\phi^{\theta_a H} = \theta_{-a} \circ \phi^H \circ \theta_a = \theta'_a \phi^H.$$

In particular, if $\mathcal{G} : \mathcal{H} \to \mathcal{S}$ is defined by

$$\mathcal{G}(H) = \phi_1^H = \phi^H,$$

then $\mathcal{G}$ pushes forward any θ-invariant ergodic measure $\mathbb{P}$ on $\mathcal{H}$, to a θ'-invariant ergodic measure $\mathcal{P}$ on $\mathcal{S}$.

(ii) *Let $\mathbb{P}$ be a θ-invariant probability measure such such that*

$$\int_{\mathcal{H}} \|H\|_{\mathrm{C}^1}^{2d+1}\ \mathbb{P}(\mathrm{d}H) < \infty.$$

Then,

$$\int_{\mathcal{H}} \mathcal{G}(H)(0)\ \mathbb{P}(\mathrm{d}H) = 0.$$

(iii) *For* $H \in \mathcal{C}^2(\ell)$*, we have* $\mathcal{F}(\phi^H) \leqslant \ell$*, and*

$$\big\|\mathrm{D}\phi^H - I\|_{\mathrm{C}^0} \leqslant (\mathrm{e}^\ell - 1). \tag{6.2}$$

In particular, ϕ^H *is a twist map if* $\mathrm{e}^\ell < 2$.

(iv) *For* $H, H' \in \mathcal{C}^2(\ell)$*, we have*

$$\|\phi^{H'} - \phi^H\|_{\mathrm{C}^0} \leqslant \mathrm{e}^\ell \|\nabla H' - \nabla H\|_{\mathrm{C}^0}. \tag{6.3}$$

(v) *Assume that* $\mathbb{P}$ *is concentrated on* $\mathcal{C}^2(\ell)$ *for some* $\ell > 0$*. Assume*

$$\int_{\mathcal{H}} \int_{\mathbb{R}^{2d}} \int_0^1 \int_0^1 |\nabla H(x,t) \cdot \nabla H(0,s)| \; |L(x)| \; \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}s \; \mathbb{P}(\mathrm{d}H) < \infty. \tag{6.4}$$

Then the map $H \mapsto \phi^H(0)$ *is in* $\widehat{\mathfrak{H}}^{-1}(\mathbb{P})$*:*

$$\int_{\mathcal{H}} \int_{\mathbb{R}^{2d}} \big|\big(\phi^H(x) - x\big) \cdot \phi^H(0)\big| \; |L(x)| \; \mathrm{d}x \; \mathbb{P}(\mathrm{d}H) < \infty. \tag{6.5}$$

(vi) *If* $H \in \mathcal{C}^2(\ell)$ *and* ∇H *is almost periodic, then* $\mathcal{F}\big(\phi^H\big)$ *is almost periodic.*

(vii) *Assume that* $H \in \mathcal{C}^2(\ell)$ *and* H *is quasiperiodic i.e. we can find an integer* $N \geqslant n$*, a 1-periodic function*

$$K : \mathbb{R}^N \to \mathbb{R},$$

a matrix $A \in \mathbb{R}^{N \times n}$*, and* $\omega \in \mathbb{R}^N$*, such that*

$$H(x,t) = H(x,t,\omega) = K(\Theta_{Ax}\omega, t).$$

(Here Θ_a *denotes the translation of* $\mathbb{R}^N$*.) Assume that the null set of* A *is trivial:*

$$Ax = 0 \quad \Longrightarrow \quad x = 0.$$

Then $\mathcal{F}\big(\phi^H\big)$ *is quasiperiodic.*

Proof. (i) This is an immediate consequence of the fact that if $y(\cdot)$ is an orbit of $X_{\theta_a H}$, then

$$x(\cdot) = \theta_a y(\cdot) = y(\cdot) + a$$

is an orbit of X_H.

(ii) Let us write $\mathrm{B}_\ell = \mathrm{B}_\ell(0)$ for the ball of radius ℓ that is centered at the origin. Since

$$\phi_1^H(x) - x = \int_0^1 J\nabla H\big(\phi_t^H(x), t\big) \; \mathrm{d}t,$$

we have

$$\begin{aligned} \int_{\mathrm{B}_\ell} \big(\phi_1^H(x) - x\big) \; \mathrm{d}x &= \int_0^1 J \int_{\mathrm{B}_\ell} \nabla H\big(\phi_t^H(x), t\big) \; \mathrm{d}x \; \mathrm{d}t \\ &= \int_0^1 J \int_{\phi_t^H(\mathrm{B}_\ell)} \nabla H(x,t) \; \mathrm{d}x \; \mathrm{d}t. \end{aligned} \tag{6.6}$$

Note that since

$$\begin{aligned} \big|\phi_t^H(x) - x\big| &\leqslant t \sup |\nabla H| \\ &\leqslant \|H\|_{\mathrm{C}^1} \\ &=: c_0, \end{aligned}$$

for $t \in [0,1]$, we have

$$\mathrm{B}_{\ell-c_0} \subset \phi_t^H(\mathrm{B}_\ell) \subset \mathrm{B}_{\ell+c_0}.$$

From this and (6.6) we learn

$$\begin{aligned}
&\left|\int_{\mathrm{B}_\ell} \big(\phi_1^H(x) - x\big)\ \mathrm{d}x\right| \\
\leqslant\ &\left|\int_0^1 J\int_{\mathrm{B}_\ell} \nabla H(x,t)\ \mathrm{d}x\ \mathrm{d}t\right| + \int_0^1 \int_{\mathrm{B}_{\ell+c_0}\setminus \mathrm{B}_{\ell-c_0}} |\nabla H(x,t)|\ \mathrm{d}x\ \mathrm{d}t \\
\leqslant\ &\left|\int_0^1 J\int_{\partial\mathrm{B}_\ell} H(x,t)\nu(x)\ \sigma(\mathrm{d}x)\ \mathrm{d}t\right| + c_0\big|\mathrm{B}_{\ell+c_0}\setminus \mathrm{B}_{\ell-c_0}\big| \\
\leqslant\ &c_0\sigma\big(\partial\mathrm{B}_\ell\big) + c_0\int_{\ell-c_0}^{\ell+c_0} \sigma(\partial\mathrm{B}_r)\ \mathrm{d}r \\
\leqslant\ &c_0\sigma\big(\partial\mathrm{B}_1\big)\big(\ell^{2d-1} + 2c_0(\ell+c_0)^{2d-1}\big) \\
\leqslant\ &c_1\ell^{2d-1}(1+c_0)^{2d+1},
\end{aligned}$$

where

$$\nu(x) = \frac{x}{|x|}$$

is the outward unit normal at $x \in \partial\mathrm{B}_\ell$, $\sigma(\mathrm{d}x)$ denotes the $2d-1$-surface measure on $\partial\mathrm{B}_\ell$, and c_1 is a constant that depends on d only. Hence, by stationarity of $\mathbb{P}$,

$$\begin{aligned}
\left|\int_{\mathcal{H}} \phi_1^H(0)\ \mathbb{P}(\mathrm{d}H)\right| &= \left|\int_{\mathcal{H}} \left[|\mathrm{B}_\ell|^{-1}\int_{\mathrm{B}_\ell} \big(\phi_1^H(x) - x\big)\ \mathrm{d}x\right]\ \mathbb{P}(\mathrm{d}H)\right| \\
&\leqslant c_2\ell^{-1}\int_{\mathcal{H}} \big(1 + \|H\|_{\mathrm{C}^1}\big)\big)^{2d+1}\ \mathbb{P}(\mathrm{d}H),
\end{aligned}$$

for a constant c_2. We now send $\ell \to \infty$ to complete the proof.

(iii) Evidently,

$$\big|\Phi^H(x) - x\big| \leqslant \left|\int_0^1 J\nabla H(\phi_s^H(x), s)\ \mathrm{d}s\right| \leqslant \ell.$$

On the other hand, if

$$V(x,t) = J\mathrm{D}^2 H\big(\phi_t^H(x), t\big), \qquad A(x,t) = \mathrm{D}\phi_t^H(x),$$

then

$$A_t(x,t) = V(x,t)A(x,t),$$

which leads to the identity

$$A(x,t) = I + \sum_{k=1}^{\infty}\int_{\Delta_n(t)} V(x,t_n)\dots V(x,t_1)\ \mathrm{d}t_1\dots\mathrm{d}t_n,$$

where

$$\Delta_n(t) = \big\{(t_1,\dots,t_n):\ 0 \leqslant t_1 \leqslant \dots \leqslant t_n \leqslant t\big\}.$$

From this we deduce (6.2) because $|V| \leqslant \ell$ by (6.1).

(iv) Note that if $H \in \mathcal{L}^2(\ell)$, then (6.1) implies that the Lipschitz constant of the vector field $J\nabla H$ is at most ℓ. Using this, we can write,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathrm{e}^{-\ell t}\left|\phi_t^{H'}(x)-\phi_t^H(x)\right|\right] \leqslant & \mathrm{e}^{-\ell t}\left|J\nabla H'(\phi_t^{H'}(x),t)-J\nabla H(\phi_t^H(x),t)\right| \\ & -\ell\mathrm{e}^{-\ell t}\left|\phi_t^{H'}(x)-\phi_t^H(x)\right| \\ \leqslant & \mathrm{e}^{-\ell t}\left|J\nabla H(\phi_t^{H'}(x),t)-J\nabla H(\phi_t^H(x),t)\right| \\ & +\mathrm{e}^{-\ell t}\|\nabla H'-\nabla H\|_{\mathrm{C}^0}-\ell\mathrm{e}^{-\ell t}\left|\phi_t^{H'}(x)-\phi_t^H(x)\right| \\ \leqslant & \mathrm{e}^{-\ell t}\|\nabla H'-\nabla H\|_{\mathrm{C}^0}.\end{aligned}$$

Integrating both sides with respect to t yields

$$\mathrm{e}^{-\ell}\left|\phi^{H'}(x)-\phi^H(x)\right| \leqslant \|\nabla H'-\nabla H\|_{\mathrm{C}^0},$$

which in (6.3).

(v) By stationarity, the left-hand side of (6.5) equals to

$$\begin{aligned}&\frac{1}{|\mathrm{B}_1(0)|}\int_{|a|\le 1}\int_{\mathcal{H}}\int_{\mathbb{R}^{2d}}\left|\big(\phi^H(x)-x\big)\cdot\big(\phi^H(a)-a\big)\right|\ |L(x-a)|\ \mathrm{d}x\mathrm{d}a\ \mathbb{P}(\mathrm{d}H) \\ &=\frac{1}{|\mathrm{B}_1(0)|}\int_{|a|\le 1}\int_{\mathcal{H}}\int_{\mathbb{R}^{2d}}\left|\left(\int_0^1 X\big(\phi_t^H(x),t\big)\ \mathrm{d}t\cdot\int_0^1 X(\phi_s^H(a),s)\ ds\right)\right|\ |L(x-a)|\ \mathrm{d}x\ \mathrm{d}a\ \mathbb{P}(\mathrm{dH}) \\ &\le\frac{1}{|\mathrm{B}_1(0)|}\int_{\mathcal{H}}\int_0^1\int_0^1\int_{|a|\le 1}\int_{\mathbb{R}^{2d}}\left|\nabla H\big(\phi_t^H(x),t\big)\cdot\nabla H(\phi_s^H(a),s)\right|\ |L(x-a)|\ \mathrm{d}x\ \mathrm{d}a\ \mathrm{d}t\ \mathrm{d}s\ \mathbb{P}(\mathrm{dH}),\end{aligned}$$

where $X(x,t)=J\nabla H(x,t)$. We now make a (volume preserving) change of variable

$$(y,z)=\big(\phi_t^H(x),\phi_s^H(a)\big),$$

for the d$x$da integration to rewrite the last expressions as

$$\frac{1}{|\mathrm{B}_1(0)|}\int_{\mathcal{H}}\int_0^1\int_0^1\int_{\phi_s^H(\mathrm{B}_1(0))}\int_{\mathbb{R}^{2d}}|\nabla H(y,t)\cdot\nabla H(z,s)|\ \left|L\big(\psi_t(y)-\psi_s(z)\big)\right|\ \mathrm{d}y\ \mathrm{d}z\ \mathrm{d}t\ \mathrm{d}s\ \mathbb{P}(\mathrm{dH}),$$

where ψ_t is the inverse of ϕ_t^H. This expression is bounded above by

$$\frac{1}{|\mathrm{B}_1(0)|}\int_{\mathcal{H}}\int_0^1\int_0^1\int_{\mathrm{B}_{\ell+1}(0)}\int_{\mathbb{R}^{2d}}|\nabla H(y,t)\cdot\nabla H(z,s)|\ \left|L\big(\psi_t(y)-\psi_s(z)\big)\right|\ \mathrm{d}y\ \mathrm{d}z\ \mathrm{d}t\ \mathrm{d}s\ \mathbb{P}(\mathrm{dH}),$$

because

$$\phi_s^H(\mathrm{B}_1(0))\subset\mathrm{B}_{\ell+1}(0),$$

$\mathbb{P}$-almost surely by (6.1). Observe that ψ_t is the flow of the Hamiltonian ODE associated with $J\nabla H$ with time reversed. So using (6.1),

$$|\psi_t(y)-y|\ ,\ |\psi_s(z)-z|\le\ell,$$

$\mathbb{P}$-almost surely, for $y,z\in\mathbb{R}^{2d}$, and $s,t\in[0,1]$. As a result,

$$|y-z|\geqslant 4\ell \quad\Longrightarrow\quad \left|\psi_t(y)-\psi_s(z)\right|\geqslant|y-z|-2\ell\geqslant\frac{1}{2}|x-y|. \tag{6.7}$$

We now assume that $d > 1$ so that $L(x)$ is given by a constant multiple of $|x|^{2-2d}$. From of this, and (6.7), we learn that the left-hand side of (6.5) is bounded above by

$$\Lambda_1 + \Lambda_2,$$

where

$$\begin{aligned}\Lambda_1 =& \frac{1}{|\mathrm{B}_1(0)|}\int_{\mathcal{H}}\int_0^1\int_0^1\iint_{E_\ell}|\nabla H(y,t)\cdot\nabla H(z,s)|\ \big|L\big(\psi_t(y)-\psi_s(z)\big)\big|\ \mathrm{d}y\ \mathrm{d}z\ \mathrm{d}t\ \mathrm{d}s\ \mathbb{P}(\mathrm{dH}),\\ \Lambda_2 =& c_0\ \frac{1}{|\mathrm{B}_1(0)|}\int_{\mathcal{H}}\int_0^1\int_0^1\int_{\mathrm{B}_{\ell+1}(0)}\int_{\mathbb{R}^{2d}}|\nabla H(y,t)\cdot\nabla H(z,s)|\ \big|L(y-z)\big|\ \mathrm{d}y\ \mathrm{d}z\ \mathrm{d}t\ \mathrm{d}s\ \mathbb{P}(\mathrm{dH}),\end{aligned}$$

for a constant c_0, and

$$E_\ell = \big\{(y,z)\in\mathbb{R}^{2d}:\ |y|\le \ell+1,\ |y-z|\le 2\ell\big\}.$$

By stationarity of $\mathbb{P}$,

$$\begin{aligned}\Lambda_2 =& c_0\ \frac{1}{|\mathrm{B}_1(0)|}\int_{\mathcal{H}}\int_0^1\int_0^1\int_{\mathrm{B}_{\ell+1}(0)}\int_{\mathbb{R}^{2d}}|\nabla H(y-z,t)\cdot\nabla H(0,s)|\ \big|L(y-z)\big|\ \mathrm{d}y\ \mathrm{d}z\ \mathrm{d}t\ \mathrm{d}s\ \mathbb{P}(\mathrm{dH})\\ =& c_0\ \frac{|\mathrm{B}_{\ell+1}(0)|}{|\mathrm{B}_1(0)|}\int_{\mathcal{H}}\int_0^1\int_0^1\int_{\mathbb{R}^{2d}}|\nabla H(x,t)\cdot\nabla H(0,s)|\ \big|L(x)\big|\ \mathrm{d}x\ \mathrm{d}t\ \mathrm{d}s\ \mathbb{P}(\mathrm{dH}),\end{aligned}$$

and this is finite by our assumption (6.4). It remains to show $\Lambda_1 < \infty$. Indeed from our assumption, (6.1) holds $\mathbb{P}$-almost surely, which yields the bound

$$\Lambda_1 \le \frac{\ell^2}{|\mathrm{B}_1(0)|}\int_{\mathcal{H}}\iint_{E_\ell}\big|L\big(\psi_t(y)-\psi_s(z)\big)\big|\ \mathrm{d}y\ \mathrm{d}z\ \mathbb{P}(\mathrm{dH}).$$

We make the change of variable

$$(x,a) = \Psi(y,z) := \big(\psi_t(y),\psi_s(z)\big),$$

to rewrite the integral as

$$\int_{\mathcal{H}}\iint_{\Psi(E_\ell)}|L(x-a)|\ \mathrm{d}x\ \mathrm{d}a\ \mathbb{P}(\mathrm{dH}).$$

Using (6.1), we have

$$\Psi(E_\ell)\subset \mathrm{B}_{2\ell+1}(0)\times\mathrm{B}_{4\ell+1}(0).$$

From this and local integrability of $L(x)$ we deduce that $\Lambda_1 < \infty$.

The case $d = 1$ can be treated in a similar fashion.

(vi) Let us write

$$\begin{aligned}O(H) &:= \big\{\theta_a H:\ a\in\mathbb{R}^n\big\},\\ \widehat{O}(H) &:= \big\{\theta'_a\phi^H:\ a\in\mathbb{R}^n\big\}.\end{aligned}$$

By part (i), we know

$$\mathcal{G}\big(O(H)\big) = \widehat{O}(H). \tag{6.8}$$

If $H\in\mathcal{C}^2(\ell)$ and ∇H is almost periodic, then

$$O(H)\subset\mathcal{C}^2(\ell),$$

and $O(\nabla H)$ is precompact with respect to C^0-topology. From this and (6.8) we deduce the precompactness of $\widehat{O}(H)$. Since

$$\theta'_a \phi^H - \theta'_b \phi^H = \theta_a \mathcal{F}(\phi^H) - \theta_b \mathcal{F}(\phi^H),$$

we deduce the precompactness of the set $\{\theta_a \mathcal{F}(\phi^H)\}$. As a consequence, the map $\mathcal{F}(\phi^H)$ is almost periodic.

(vii) Set $\zeta_t(\omega) = \phi_t^{H(\cdot,\omega)}(0)$. We claim that ζ is periodic. Observe that if $x(t)$ solves the ODE

$$\dot{x}(t) = J\nabla H(x(t), t) = JA^*\nabla K(\omega + Ax(t), t), \qquad x(0) = 0,$$

then $\omega(t) := \omega + Ax(t)$ satisfies

$$\dot{\omega}(t) = AJA^*\nabla K(\omega(t), t).$$

From this we learn that if ψ_t is the flow of the vector field

$$\hat{X} := AJA^*\nabla K,$$

then

$$\psi_t(\omega) = \omega + A\ \phi_t^{H(\cdot,\omega)}(0) = \omega + A\zeta_t(\omega).$$

Since $\hat{K}$ is periodic, we learn that $\mathcal{F}(\psi_t) = \psi_t - \mathrm{id}$ is periodic by part (i). Hence $A\zeta_t(\omega)$ is periodic. Since A has a trivial null space, we deduce that ζ_t is periodic. On the other hand,

$$\begin{aligned} \phi^{H(\cdot,\omega)}(x) - x =& \big(\theta'_x \phi^{H(\cdot,\omega)}\big)(0) = \big(\phi^{\theta_x H(\cdot,\omega)}\big)(0) = \big(\phi^{H(\cdot,\theta_x\omega)}\big)(0) \\ =& \zeta_1(\theta_x\omega) = \zeta_1(\omega + Ax). \end{aligned}$$

From this, and the periodicity of ζ_1, we deduce the quasiperiodicity of the left-hand side. □

6.1. **Proof of Theorem 2.6.** We are now ready to offer a more precise statement of Theorem 2.6 and give a proof.

Theorem 6.2. *Assume that $H(x,t,\omega) = K(\Theta_{Ax}\omega, t)$, for a C^1 function $K : \mathbb{T}^N \times \mathbb{T} \to \mathbb{R}$. Let $\mathbb{P}$ denotes the Lebesgue measure on $\mathbb{T}^N$, and assume $K \in \mathcal{C}^2(\ell)$, for some $\ell \in (0, \log 2)$, and that* (6.4) *holds. Then the set*

$$\{x \in \mathbb{R}^{2d} :\ \phi^{H(\cdot,\omega)}(x) = x\},$$

is of positive (possibly infinite) density, $\mathbb{P}$-almost surely.

Proof. It suffices to show that the conditions of Theorem 5.1 hold true for

$$\Phi(x,\omega) = \phi^{H(\cdot,\omega)}(x).$$

These conditions have been verified in Propositions 6.1, parts (i)–(iii), (v), (vii). □

Remark 6.3 As we discussed in Example 3.6, a Diophantine-type condition on A, and the existence of certain number of derivatives of the function K would guarantee the validity of (6.4) for the corresponding Hamiltonian function H in the quasiperodic setting of Theorem 6.2.

7. Appendix. The 2 dimensional case: Random Poincare-Birkhoff theorem

To put in context the stochastic Conley-Zehnder theory we develop in this paper, we very briefly review the simpler case of stochastic symplectic maps in dimension 2. In order to describe our results, let us write, following [PR], $\mathcal{T}$ for the space of area preserving twist maps. Let $\overline{\mathcal{T}}$ be the space of maps

$$\bar{F} : (\mathcal{A} = \mathbb{R} \times [-1,\, 1]) \to \mathcal{A}$$

such that if

$$\ell(\bar{F})(q,p) := (q,0) + \bar{F}(q,p),$$

then $\ell(\bar{F}) \in \mathcal{T}$. Consider the operator $\ell : \overline{\mathcal{T}} \to \mathcal{T}$ which send $\bar{F}$ to $F = \ell(\bar{F})$. So we have a family of shifts

$$\big\{\tau_a : \overline{\mathcal{T}} \to \overline{\mathcal{T}} :\ a \in \mathbb{R}\big\},$$

defined by $\tau_a \bar{F}(q,p) = \bar{F}(q+a,p)$. For any $F \in \mathcal{T}$ we write

$$\mathrm{Fix}(F) = \big\{x \in \mathcal{A} :\ F(x) = x\big\}.$$

Also, with a slight abuse of notation we write τ_a instead of

$$\tau_a A = \{x :\ x + a \in A\}.$$

We then have the trivial commutative relationship

$$\tau_a \mathrm{Fix}(\ell(\bar{F})) = \mathrm{Fix}(\ell(\tau_a \bar{F})). \tag{7.1}$$

Furthermore, we adopt the following notations and terminologies: we denote by $\overline{\mathcal{MT}_+}$ the space of

$$\bar{F} = (\bar{Q}, \bar{P}) \in \overline{\mathcal{T}}$$

with the property that for every q we have that the function $f\colon [-1,\, 1] \to \mathbb{R}$ defined by $f(p) := \bar{Q}(q,\, p)$ is increasing; we denote by $\mathcal{MT}_+$ the space of $\ell(\bar{F})$, with $\bar{F} \in \overline{\mathcal{MT}_+}$; we denote by $\mathcal{MT}_-$ the space of F such that $F^{-1} \in \mathcal{MT}_+$; the elements of $\mathcal{MT}_+$ are the positive monotone twist maps; the elements of $\mathcal{MT}_-$ are the negative monotone twist maps; a fixed point $x = (q,\, p)$ of $F(\cdot,\, \cdot)\colon \mathcal{S} \to \mathcal{S}$ is of $+$ (respectively $-$) type if the eigenvalues of $\mathrm{d}F(q,\, p)$ are positive (respectively negative); finally we write

$$\mathrm{Fix}_{\pm}(F) = \big\{x \in \mathrm{Fix}(F) :\ x \text{ is of } \pm \text{ type}\big\}.$$

For any $F = \ell(\bar{F}) \in \mathcal{MT}_+$ there is a scalar-valued function

$$\mathcal{G}(q,Q) = \mathcal{G}(q,Q;\bar{F})$$

such that

$$F(q, -\mathcal{G}_q(q,Q)) = (Q, \mathcal{G}_Q(q,Q)). \tag{7.2}$$

Due to the existence of boundary conditions for

$$F(q,p) = (P(q,p), Q(q,p)),$$

we only need to define $\mathcal{G}(q,Q)$ for (q,Q) such that $Q(q,-1) \leqslant Q \leqslant Q(q,+1)$. This means in particular that $\psi(q;\bar{F}) := \mathcal{G}(q,q;\bar{F})$ is well defined. By (7.1),

$$\psi(\cdot;\tau_a\bar{F}) = \tau_a \psi(\cdot;\bar{F}).$$

In [PR, Theorem B] we saw that if $\mathbb{Q}$ is a translation invariant ergodic probability measure on $\big(\overline{\mathcal{T}},\mathcal{F}\big)$ satisfying that

$$\mathbb{Q}\big(\overline{\mathcal{M}\mathcal{T}_+}\big)=1$$

then all the sets $\mathrm{Fix}_\pm\big(\ell(\bar F)\big)$ are nonempty with probability one with respect to $\mathbb{Q}$. Moreover, also with probability one, if the random pair

$$\left(\frac{\mathrm{d}}{\mathrm{d}q}\psi(q;\bar F),\frac{\mathrm{d}^2}{\mathrm{d}q^2}\psi(q;\bar F)\right),$$

has a probability density $\rho(a,b;\bar F)$ (which is independent of q due to the translation invariance), then the sets $\mathrm{Fix}_\pm\big(\ell(\bar F)\big)$ have positive density $\lambda_\pm$ given by

$$\lambda_\pm=\int\left[\int_{-\infty}^{\infty}b^\pm\rho(0,b;\bar F)\ \mathrm{d}b\right]\ \mathbb{Q}(\mathrm{d}\bar F).$$

Now denote by Ω_0 the space of functions $\omega\colon\mathbb{R}^2\to\mathbb{R}$ such that $\omega(q,a)>0$ for $a>0$, $\omega(q,0)=0$, and,

$$\eta(q;\omega)=\inf\{a\mid\omega(q,a)=2\}<\infty,$$

for every q. Define

$$\begin{aligned}Q^-(q;\omega)&=\frac{1}{2}\int_0^{\eta(q;\omega)}\omega(q,a)\ \mathrm{d}a-\eta(q;\omega);\\ G(q,Q;\omega)&=\omega(q,Q-q-Q^-(q;\omega)).\end{aligned}$$

Denote by Ω_1 the space of all $\omega\in\Omega_0$ satisfying that $G_q(q,Q;\omega)<0$ for every (q,Q). In [PR, Theorem C] we proved that for each $\omega\in\Omega_1$ there is a unique function $\bar F(\cdot,\cdot;\omega)$ such that if

$$F(\cdot,\cdot;\omega)=\ell\big(\bar F(\cdot,\cdot;\omega)\big)$$

then (7.2) holds for $\mathcal{G}(q,Q)=\mathcal{G}(q,Q;\omega)$, that is given by

$$\mathcal{G}(q,Q;\omega)=\int_{q+Q^-(q;\omega)}^{Q}\omega(q,a)\ \mathrm{d}a-(Q-q).$$

Moreover, we proved that if

$$\tau_a\omega(q,v)=\omega(q+a,v)$$

then we have that

$$\bar F(\cdot,\cdot;\tau_a\omega)=\tau_a\bar F(\cdot,\cdot;\omega).$$

To continue the discussion of results on the 2-dimensional case let Ω_2 denote the set C^2 Hamiltonian functions $\omega(q,p,t)$ with the property that they have uniformly bounded second derivatives and such that $\pm\omega_p(q,\pm1,t)>0$ and $\omega_q(q,\pm1,t)=0$. If $\omega\in\Omega_2$ let

$$\tau_a\omega(q,p,t)=\omega(q+a,p,t)$$

as before, and also let

$$\phi_t^\omega(q,p)$$

be the flow of the corresponding Hamiltonian system

$$\dot q=\omega_p(q,p,t),\dot p=-\omega_q(q,p,t).$$

One can show that if

$$F^t(q,p;\omega)=\phi_t^\omega(q,p)$$

and

$$\bar{F}^t(q,p;\omega) = \phi_t^\omega(q,p) - (q,0)$$

then

$$F^t(\cdot,\cdot;\omega) \in \mathcal{T}$$

and

$$\bar{F}^t\big(q,p;\tau_a\omega\big) = \tau_a \bar{F}^t\big(q,p;\omega\big).$$

With this in mind, we proved in [PR, Theorem D] that if $\mathbb{P}$ is a τ-invariant ergodic probability measure $\mathbb{P}$ on Ω_2, then for every $t \geqslant 0$ we have that

$$\mathbb{P}(\#\mathrm{Fix}(F^t(\cdot,\ \cdot,\ ;\omega)) = \infty) = 1.$$

Denote by $\mathcal{C}([0,1];\overline{\mathcal{T}})$ the space of C^1 maps $\gamma : [0,1] \to \overline{\mathcal{T}}$ for which $\ell(\gamma(0))$ is the identity. The operator τ on $\overline{\mathcal{T}}$ induces a new operator (denoted in the same way) on $\mathcal{C}([0,1];\overline{\mathcal{T}})$ defined by

$$(\tau_a\gamma)(t) = \tau_a(\gamma(t)).$$

We know that if $\mathbb{P}$ is a stationary ergodic measure on Ω_2 then

$$\omega \mapsto \big(\bar{F}^t(q,p;\omega) = \phi_t^\omega(q,p) - (q,0) :\ t \in [0,1]\big),$$

pushes forward $\mathbb{P}$ onto a stationary ergodic probability measure $\mathcal{Q}$ on $\mathcal{C}([0,1];\overline{\mathcal{T}})$. The converse also holds; a stationary ergodic probability measure $\mathcal{Q}$ on $\mathcal{C}([0,1];\overline{\mathcal{T}})$ always comes from a unique a stationary ergodic measure $\mathbb{P}$ on Ω_2.

Let $\mathbb{Q}$ be a stationary ergodic measure on $\overline{\mathcal{T}}$. A natural question is whether we can find a stationary ergodic measure $\mathcal{Q}$ on $\mathcal{C}([0,1];\overline{\mathcal{T}})$ such that $\mathbb{Q}$ is the push forward of $\mathcal{Q}$ under the time-1 map $\pi_1 : \mathcal{C}([0,1];\overline{\mathcal{T}}) \to \overline{\mathcal{T}}$ (by time-1 map we mean $\pi_1\gamma := \gamma(1)$). In order to discuss this question let $\mathcal{D}$ denote the space of diffeomorphisms $F : \mathcal{A} \to \mathcal{A}$. We also denote by $\overline{\mathcal{D}}$ the space of functions $\bar{F}$ with the property that $\ell(\bar{F}) \in \mathcal{D}$.

Let $\mathcal{Q}$ be a stationary ergodic measure on $\mathcal{C}([0,1];\overline{\mathcal{D}})$. In [PR] we called $\mathcal{Q}$ regular if

$$\int \sup_{t\in[0,1]} \left[\|\dot{\gamma}(t)\|_\infty + \|\mathrm{d}\gamma(t)\|_\infty + \left\|\mathrm{d}\gamma(t)^{-1}\right\|_\infty\right]\ \mathcal{Q}(\mathrm{d}\gamma) < \infty.$$

Here $\|\cdot\|_\infty$ denotes the L^∞ norm and $\dot{\gamma}(t)$ and $\mathrm{d}\gamma(t)$ denote the derivatives of $\gamma(t)$ with respect to t and $x = (q,p)$, and

$$\frac{1}{2}\int \left[\int_{-1}^{1} \det(\mathrm{d}\gamma(t)(q,p))\,\mathrm{d}p\right]\ \mathcal{Q}(\mathrm{d}\gamma) = 1;$$

for all $t \in [0,1]$.

If we start with a stationary ergodic measure $\mathbb{Q}$ on $\overline{\mathcal{T}}$, in [PR, Theorem E] we proved that if there exists a regular stationary ergodic measure $\mathcal{Q}$ on $\mathcal{C}([0,1];\overline{\mathcal{D}})$, such that $\mathbb{Q}$ is the push forward of $\mathcal{Q}$ under the time-1 map $\pi_1\gamma := \gamma(1)$, then there is another stationary ergodic measure $\mathcal{Q}'$ on $\mathcal{C}([0,1];\overline{\mathcal{T}})$ with the property that $\mathbb{Q}$ is the push forward of $\mathcal{Q}$ under π_1.

Finally, our paper [PR] concluded by showing that if $\mathbb{P}$ is a τ-invariant ergodic probability measure $\mathbb{P}$ on Ω_2 and $F = F^1$ is as in the result we described earlier, then there exists [PR, Theorem E] a deterministic integer $N \geqslant 0$ and area-preserving random twists F_j, $0 \leqslant j \leqslant N$, such that for $\mathbb{P}$ almost all $\omega \in \Omega_2$, we have a decomposition:

$$F(\cdot,\cdot;\omega) = F_N(\cdot,\cdot;\omega) \circ \ldots \circ F_2(\cdot,\cdot;\omega) \circ F_1(\cdot,\cdot;\omega) \circ F_0(\cdot,\cdot;\omega),$$

where the map F_j is positive monotone if j is an odd integer, F_j is negative monotone if j is an even integer, and

$$\bar{F}_j(q,p;\tau_a\omega) = \tau_a \bar{F}_j(q,p;\omega)$$

for every j.

References

[AT07] R. Adler and J.E. Taylor: *Random Fields and Geometry.* Springer Monographs in Mathematics. Springer, New York, 2007.

[AT09] R. Adler and J.E. Taylor: *Topological Complexity of Smooth Random Functions.* Ecole d'été de probabilités de Saint-Flour XXXIX-2009. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 2011.

[AW09] J. Azais and M. Wschebor: Level Sets and Extrema of Random Processes and Fields. John Wiley and Sons, Inc., Hoboken, NJ, 2009.

[Ar78] V.I. Arnol'd: *Mathematical Methods of Classical Mechanics*, Springer-Verlag, Berlin and NY, 1978.

[Bi13] G. D. Birkhoff: Proof of Poincaré's last geometric theorem, *Trans. AMS* **14** (1913) 14-22.

[Bi26] G. D. Birkhoff: An extension of Poincaré's last geometric theorem, *Acta. Math.* **47** (1926) 297-311.

[Ch84] M. Chaperon: *Une idée du type "géodésiques brisées" pour les systémes hamiltoniens.* (French) [A "broken geodesic" method for Hamiltonian systems] C. R. Acad. Sci. Paris Sér. I Math. 298 (1984), no. 13, 293-296.

[Ch84b] M. Chaperon: *An elementary proof of the Conley-Zehnder theorem in symplectic geometry.* Dynamical systems and bifurcations (Groningen, 1984), 1-8, Lecture Notes in Math., 1125, Springer, Berlin, 1985.

[Ch89] M. Chaperon: *Recent results in symplectic geometry.* Dynamical systems and ergodic theory (Warsaw, 1986), 143-159, Banach Center Publ., 23, PWN, Warsaw, 1989.

[CZ83] C.C. Conley and E. Zehnder: The Birkhoff-Lewis fixed point theorem and a conjecture of V. I. Arnol'd. *Invent. Math.* **73** (1983) 33-49.

[Fl88] A. Floer: Morse theory for Lagrangian intersections. *J. Diff. Geom.* **28** (1988) 513-547.

[Fl89] A. Floer: Witten's complex and infinite-dimensional Morse theory. *J. Diff. Geom.* **30** (1989) 207-221.

[Fl89b] A. Floer: Symplectic fixed points and holomorphic spheres. *Comm. Math. Phys.* **120** (1989) 575-611.

[Fl91] A. Floer: *Elliptic methods in variational problems.* A plenary address presented at the International Congress of Mathematicians held in Kyoto, August 1990. ICM-90. Mathematical Society of Japan, Tokyo; distributed outside Asia by the American Mathematical Society, Providence, RI, 1990.

[Go01] C. Golé: *Symplectic Twist Maps. Global Variational Techniques.* Advanced Series in Nonlinear Dynamics, 18. World Scient. Publis. Co., Inc., River Edge, NJ, 2001. xviii+305 pp.

[Ho85] H. Hofer: Lagrangian embeddings and critical point theory. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **2** (1985) 407-462.

[HS95] H. Hofer and D. Salamon: Floer homology and Novikov rings. *The Floer Memorial Volume*, 483-524, Progr. Math., 133, Birkhäuser, Basel, 1995.

[HZ94] H. Hofer and E. Zehnder: *Symplectic Invariants and Hamiltonian Dynamics*, Birkhäuser, Basel (1994).

[La02] P. D. Lax: *Functional Analysis*, John Wiley and Sons, Inc. 2002.

[LT98] G. Liu and G. Tian: Floer homology and Arnol'd conjecture. *J. Differential Geom.* **49** (1998) 1-74.

[On95] K. Ono: On the Arnol'd conjecture for weakly monotone symplectic manifolds, *Invent. Math.* **119** (1995) 519-537.

[PR] Á. Pelayo, F. Rezakhanlou: Poincare-Birkhoff theorems in random dynamics, *Trans. Amer. Math. Soc.* 370 (2018) 601-639.

[Po93] H. Poincaré: Les Méthodes Nouvelles de la Mécanique Céleste, Tome I, Paris, Gauthier-Viltars, 1892. Republished by Blanchard, Paris, 1987.

[Po12] H. Poincaré: Sur un théorème de géométrie, *Rend. Circ. Mat. Palermo* **33** (1912) 375-407.

[Pol01] L. Polterovich: *The Geometry of the Group of Symplectic Diffeomorphisms.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2001. xii+132 pp.

[R] F. Rezakhanlou: Lectures on Dynamical Systems. https://math.berkeley.edu/∼rezakhan/dyn-partII.pdf

[Vi11] C. Viterbo: Symplectic topology as the geometry of generating functions, *Math. Ann.* **292** (1992) 685-710.

[We86] A. Weinstein: On extending the Conley-Zehnder fixed point theorem to other manifolds. *Nonlinear functional analysis and its applications, Part 2 (Berkeley, Calif., 1983), 541-544, Proc. Sympos. Pure Math., 45, Part 2, Amer. Math. Soc., Providence, RI,* 1986,

[Ze86] E. Zehnder: The Arnol'd conjecture for fixed points of symplectic mappings and periodic solutions of Hamiltonian systems. *Proceedings of the International Congress of Mathematicians, Vol. 1, 2* (Berkeley 1986), 1237-1246, Amer. Math. Soc., Providence, RI, 1987.

Álvaro Pelayo
alvpel01@ucm.es
Facultad de Ciencias Matemáticas
Universidad Complutense de Madrid
28040 Madrid, Spain

Fraydoun Rezakhanlou
rezakhan@math.berkeley.edu
Department of Mathematics
University of California, Berkeley
Berkeley, CA 94720-3840 USA