EXPONENTIAL MIXING OF TORUS EXTENSIONS OVER EXPANDING MAPS

JIANYU CHEN AND HUYI HU

Abstract. We study the mixing property for the skew product $F : \mathbb{T}^d \times \mathbb{T}^\ell \to \mathbb{T}^d \times \mathbb{T}^\ell$ given by $F(x, y) = (Tx, y + \tau(x))$, where $T : \mathbb{T}^d \to \mathbb{T}^d$ is a C^∞ uniformly expanding endomorphism, and the fiber map $\tau : \mathbb{T}^d \to \mathbb{T}^\ell$ is a C^∞ map. We apply the semiclassical analytic approach to get the dichotomy: either F mixes exponentially fast or τ is an essential coboundary. In the former case, the Koopman operator $\widehat{F}$ of F has spectral gap in some Hilbert space that contains all $L^2(\mathbb{T}^d \times \mathbb{T}^\ell)$ functions, and in the latter case the system is semiconjugate to an expanding endomorphism crossing a torus rotation.

0. Introduction.

In this paper we study the mixing properties for torus extension of expanding maps. The systems F we consider are of the form of skew products with expanding $T : \mathbb{T}^d \to \mathbb{T}^d$ on the base and torus rotations with rotation vectors $\tau(x)$, $x \in \mathbb{T}^d$, on the fibers $\mathbb{T}^\ell$. (See (1.2) for the maps.) We obtain a dichotomy: either such a system has exponential decay of correlations with respect to the smooth invariant measure, or the rotation function $\tau(x)$ over $\mathbb{T}^d$ is an essential coboundary. The latter implies that the system is semiconjugate to an expanding endomorphism crossing a torus rotation, or simply semiconjugate to a circle rotation, and therefore cannot be weak mixing (Theorem 3 (iii)) or stably ergodic.

The methods we use to get exponential mixing is the semiclassical analytic approach. Instead of the Ruelle-Perron-Frobenius transfer operators acting on some Hölder function space, we study the dual operator, Koopman operator $\widehat{F}$, given by $\widehat{F}\phi = \phi \circ F$, acting on certain distribution space. By Fourier transform along $\mathbb{T}^\ell$, the fiber direction, the operator can be decompose to a family of operators $\{\widehat{F}_\nu\}_{\nu \in \mathbb{Z}^\ell}$, where ν is the frequency. Such operators can be regarded as Fourier integral operators. Using semiclassical analysis theory we can obtain that the spectral radius of $\widehat{F}_\nu$ is strictly less than 1 for all $\nu \neq 0$, and uniformly less than 1 for all ν with $|\nu|$ large whenever τ is not an essential coboundary, while 1 is the only eigenvalue of $\widehat{F}_0$ on the unit circle and it is simple. Hence the operator $\widehat{F}$ has a spectral gap, and the system has exponential decay of correlations.

Dolgopyat established exponential mixing property for compact group extensions of expanding maps under a generic condition called infinitesimally completely non-integrability (see [3]). Faure used semiclassical analysis in [5] to obtain exponential mixing for a simpler but intuitive model - a circle extension of an expanding map of T - under a so-called partially captive condition. For the low dimensional case, the dichotomy similar to that in Theorem 1 was obtained by Butterley and Eslami recently in [1] and discontinuities are allowed at a finite set for maps T and τ there.

Naud showed in [9] that such skew product cannot mix super-exponentially, not even for analytic observables.

Similar results were also obtained in the context of suspension semiflows over linear expanding maps. Pollicott [12] used Dolgopyat's estimates [2] to show that the generic suspension semiflows is exponentially mixing. Tsujii [15] constructed an anisotropic Sobolev space on which the transfer operator has spectral gap.

This paper is organized as the following. The setting and statements of results are given in Section 1. In Section 2 we introduce some notions and results from classical and semiclassical analysis, including Fourier transform, Sobolev spaces, Pseudo-differential operators, Fourier Integral Operators, Egorov's Theorem, and L^2-continuity theorems. This section is not necessary for the reader who is familar with the theory. We prove the theorems of the paper in Section 3 based on Proposition 3.1 and 3.2, which give the spectral radius of the Koopman operator, the dual operator of the transfer operator. The propositions are proved in Section 4, using classical and semiclassical analysis. A key estimates in the proof, stated in Lemma 5.1, is postponed in Section 5.

1. Statement of results.

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, and let $T : \mathbb{T}^d \to \mathbb{T}^d$ be a C^∞ uniformly expanding map such that

$$\gamma := \inf_{(x,v)\in S\mathbb{T}^d} |D_xT(v)| > 1, \tag{1.1}$$

where $S\mathbb{T}^d$ is the unit tangent bundle over $\mathbb{T}^d$. It is well known that T has a unique smooth invariant probability measure $d\mu(x) = h(x)dx$, where the density function $h \in C^\infty(\mathbb{T}^d, \mathbb{R}^+)$. Further, T is mixing with respect to μ.

Given a function $\tau \in C^\infty(\mathbb{T}^d, \mathbb{T}^\ell)$, we define the skew product $F : \mathbb{T}^d \times \mathbb{T}^\ell \to \mathbb{T}^d \times \mathbb{T}^\ell$ by

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} Tx \\ y + \tau(x) \pmod{\mathbb{Z}^\ell} \end{pmatrix}, \tag{1.2}$$

which preserves the product measure $dA = d\mu(x)dy$.

The mixing property of the system $(\mathbb{T}^{d+\ell}, F, dA)$ is quantified by the rates of decay of correlations. We say that the skew product F is *exponentially mixing* with respect to the smooth measure A for the observables $\phi \in L^\infty(\mathbb{T}^{d+\ell})$ and $\psi \in C^\alpha(\mathbb{T}^{d+\ell})$, $\alpha > 0$, if there exists $\rho \in [0,1)$ such that the correlation function

$$C_n(\phi, \psi; F, dA) = \left| \int \phi \circ F^n \cdot \psi dA - \int \phi dA \int \psi dA \right|.$$

satisfies $C_n(\phi, \psi; F, dA) \leqslant C_{\phi,\psi}\rho^n$ for all $n \geqslant 1$, where $C_{\phi,\psi} > 0$ is a constant depending on ϕ and ψ.

Certain cohomological conditions might give obstructions to the exponential mixing property.

Definition 1.1. *A function $\tau \in C^\infty(\mathbb{T}^d, \mathbb{T}^\ell)$ is called a* (directional) essential coboundary *if there exist $\nu \in \mathbb{Z}^\ell\backslash\{\mathbf{0}\}$, $c \in \mathbb{T}$ and a measurable function $u : \mathbb{T}^d \to \mathbb{T}$ such that*

$$\nu \cdot \tau(x) = c + u(x) - u(Tx), \quad \mu - a.e.\ x. \tag{1.3}$$

Remark 1.2. *By Livsic theory (see [7] for example), we actually have $u \in C^\infty(\mathbb{T}^d, \mathbb{T})$.*

Our main result is the following.

Theorem 1. *Let* $(\mathbb{T}^{d+\ell}, F, dA)$ *be the skew product as described above. We have the following dichotomy:*

(1) Either the system is exponentially mixing;
(2) Or $\tau(x)$ *is an essential coboundary.*

Remark 1.3. *The second case is very rare in the sense that the set consisting of all functions that are essential coboundaries is a countable union of finite and positive codimension subspaces in* $C^\infty(\mathbb{T}^d, \mathbb{T}^\ell)$. *It means that the first case that the system is exponentially mixing is generic.*

Parry and Pollicott [10] showed that $(\mathbb{T}^{d+\ell}, F, dA)$ is not mixing if and only if $\tau(x)$ is an essential coboundary. In other words, Theorem 1 asserts that F is exponentially mixing whence it is mixing. If $d = \ell = 1$, the result is proved by Butterley and Eslami in [1]. They allow the circle expansion T and the rotation τ have finite number of discontinuities.

We shall follow the semiclassical analytic approach in [5] to prove Theorem 1. Instead of the Ruelle-Perron-Frobenius transfer operators acting on some Hölder function space, we study the dual operator - Koopman operator - acting on certain distribution space.

More precisely, recall that the *Koopman operator* $\widehat{F} : L^2(\mathbb{T}^{d+\ell}, dA) \to L^2(\mathbb{T}^{d+\ell}, dA)$ defined by $\widehat{F}\phi = \phi \circ F$ is a unitary operator. Note that dA is equivalent to the Lebesgue measure $dxdy$, we instead study the action of $\widehat{F}$ on $L^2(\mathbb{T}^{d+\ell}) := L^2(\mathbb{T}^{d+\ell}, dxdy)$ as well as $\mathcal{D}'(\mathbb{T}^{d+\ell})$, the space of distributions on $\mathbb{T}^{d+\ell}$.

We say that the operator $\widehat{F}$ from a Banach space to itself has *spectral gap* if the spectrum

$$\operatorname{Spec}(\widehat{F}) = \{1\} \cup \mathcal{K}, \tag{1.4}$$

where 1 is a simple eigenvalue and $\mathcal{K}$ is a compact subset of the open disk $\{z \in \mathbb{C} : |z| < 1\}$.

Theorem 2. *If* $\tau(x)$ *is not an essential coboundary, there is an* $\widehat{F}$*-invariant Hilbert subspace* $L^2(\mathbb{T}^{d+\ell}) \subset \mathcal{W} \subset \mathcal{D}'(\mathbb{T}^{d+\ell})$ *such that* $\widehat{F}|\mathcal{W}$ *has spectral gap.*

We will specify the construction of the Hilbert space $\mathcal{W}$ in Subsection 2.3 (see (2.4)), prove this theorem in Subsection 3.3, and then show how Theorem 2 implies Theorem 1 in Subsection 3.4.

Remark 1.4. *It is well known that* $\widehat{F}|L^2(\mathbb{T}^{d+\ell})$ *does not have spectral gap. We get the result of the theorem since the norm in* $\mathcal{W}$ *is weaker than that in* $L^2(\mathbb{T}^{d+\ell})$.

Next we characterize the dynamical properties of F when $\tau(x)$ is an essential coboundary. The behaviors of the system in the $\mathbb{T}^\ell$ become very simple, as we see in part (iii) of the next theorem.

A foliation $\mathcal{L}$ of a smooth manifold M is of dimensional m if the leaves of $\mathcal{L}$ are m dimensional submanifolds. For a smooth dynamical system (F, M), a foliation $\mathcal{L}$ of M is F invariant if F preserves the leaves, that is, $F(\mathcal{L}(z)) = \mathcal{L}(F(z))$ for any $z \in M$, where $\mathcal{L}(z)$ is the leaf of $\mathcal{L}$ containing z.

Let $\Pi : \mathbb{T}^d \times \mathbb{T}^\ell \to \mathbb{T}^d$ denote the natural projection.

A smooth dynamical system (F, M) is semiconjugate to a smooth system (G, N) if there is a smooth map $\pi : M \to N$ such that $\pi \circ F = G \circ \pi$.

Theorem 3. *Let $F : \mathbb{T}^d \times \mathbb{T}^\ell \to \mathbb{T}^d \times \mathbb{T}^\ell$ be defined as in* (1.2). *The following conditions are equivalent.*

(i) *$\tau(x)$ is an essential coboundary.*
(ii) *There is an F invariant $d+\ell-1$ dimensional foliation $\mathcal{L}$ of $\mathbb{T}^d \times \mathbb{T}^\ell$ such that restricted to each leave $\mathcal{L}(z)$, $z \in \mathbb{T}^d \times \mathbb{T}^\ell$, $\Pi|_{\mathcal{L}(z)}$ is a covering map.*
(iii) *F is semiconjugate to a diffeomorphism $G = T \times R_c : \mathbb{T}^d \times \mathbb{T} \to \mathbb{T}^d \times \mathbb{T}$ through a map $\pi = \mathrm{id} \times \pi_1$, where R_c is a circle rotation with rotation number $c \in \mathbb{T}$, and $\pi_1 : \mathbb{T}^\ell \to \mathbb{T}$ is a continuous map. Further, F is semiconjugate to the circle rotation $R_c : \mathbb{T} \to \mathbb{T}$.*
(iv) *F is not weak mixing.*

Remark 1.5. *It is easy to see that if $\ell = 1$, then the leaves of the foliation $\mathcal{L}$ are unstable manifolds, and restricted to each leaf $\mathcal{L}(z)$, the natural projection $\Pi|_{\mathcal{L}(z)} : \mathcal{L}(z) \to \mathbb{T}^d$ is finite to one.*

2. Semiclassical Analysis: Preliminaries

In this section we introduce some notions and basic properties in semiclassical analysis which we are going to use. The distribution spaces and Sobolev spaces will be used in construction of the Hilbert space $\mathcal{W}$ in Theorem 2. The pseudo-differential operators (PDO) and Fourier integral operators (FIO) will be used to prove Proposition 3.1 and 3.2, where the Egorov's theorems and theorems for L^2-continuity are also used. For more information and details on this subject, one can see in standard references (e.g. [4, 14, 16]).

2.1. **Distribution spaces.** Let $\mathcal{D}(\mathbb{T}^{d+\ell}) = C^\infty(\mathbb{T}^{d+\ell})$. Its dual space $\mathcal{D}'(\mathbb{T}^{d+\ell})$ is the space of distributions on $\mathbb{T}^{d+\ell}$. If $\phi \in \mathcal{D}(\mathbb{T}^{d+\ell}) \subset \mathcal{D}'(\mathbb{T}^{d+\ell})$ and $\psi \in \mathcal{D}(\mathbb{T}^{d+\ell})$, then the action of ϕ on ψ is given by standard L^2-paring, i.e.,

$$(\psi, \phi)_{\mathcal{D},\mathcal{D}'} = \langle \psi, \phi \rangle_{L^2} = \int_{\mathbb{T}^{d+\ell}} \psi \overline{\phi}\, dxdy.$$

Since $\widehat{F}\mathcal{D}(\mathbb{T}^{d+\ell}) = \mathcal{D}(\mathbb{T}^{d+\ell})$, we define $\widehat{F}^* : \mathcal{D}(\mathbb{T}^{d+\ell}) \to \mathcal{D}(\mathbb{T}^{d+\ell})$ by the duality $\langle \widehat{F}^*\psi, \phi \rangle_{L^2} = \langle \psi, \widehat{F}\phi \rangle_{L^2}$ for all $\phi, \psi \in \mathcal{D}(\mathbb{T}^{d+\ell})$. One can check that $\widehat{F}^*$ is exactly the RPF (Ruelle-Perron-Frobenius) transfer operator over $F : \mathbb{T}^{d+\ell} \to \mathbb{T}^{d+\ell}$, that is,

$$\widehat{F}^*\psi(x) = \sum_{F(y)=x} \frac{\psi(y)}{|\mathrm{Jac}(F)(y)|}, \qquad \psi \in \mathcal{D}(\mathbb{T}^{d+\ell}).$$

Then we extend $\widehat{F}$ on $\mathcal{D}'(\mathbb{T}^{d+\ell})$ via duality again by

$$(\psi, \widehat{F}\phi)_{\mathcal{D},\mathcal{D}'} = (\widehat{F}^*\psi, \phi)_{\mathcal{D},\mathcal{D}'}, \quad \text{for all } \psi \in \mathcal{D}(\mathbb{T}^{d+\ell}),\ \phi \in \mathcal{D}'(\mathbb{T}^{d+\ell}).$$

2.2. **Fourier transforms.** The *Fourier transform* of $\varphi \in \mathcal{D}(\mathbb{T}^d)$ is defined by

$$\widehat{\varphi}(\xi) = \int_{\mathbb{T}^d} \varphi(x) e^{-i2\pi x\cdot\xi} dx, \qquad \xi \in \mathbb{Z}^d. \tag{2.1}$$

The *inverse transform* is given by

$$\varphi(x) = \sum_{\xi \in \mathbb{Z}^d} \widehat{\varphi}(\xi) e^{i2\pi\xi\cdot x}, \qquad x \in \mathbb{T}^d.$$

Let ω be the counting measure over the lattice $\mathbb{Z}^d$ on $\mathbb{R}^d$, i.e., $\omega(\xi) = \sum_{\mathbf{n}\in\mathbb{Z}^d} \delta(\xi-\mathbf{n})$ for $\xi \in \mathbb{R}^d$. Then the above equation becomes

$$\varphi(x) = \int_{\mathbb{R}^d} \widehat{\varphi}(\xi) e^{i2\pi\xi\cdot x} d\omega(\xi), \qquad x \in \mathbb{T}^d. \tag{2.2}$$

2.3. **Sobolev spaces.** Denote $\langle\xi\rangle = \sqrt{1+|\xi|^2}$, and introduce s-inner product

$$\langle\varphi, \psi\rangle_s = \sum_{\xi\in\mathbb{Z}^d} \langle\xi\rangle^{2s} \widehat{\varphi}(\xi)\overline{\widehat{\psi}(\xi)}, \qquad \varphi, \psi \in \mathcal{D}(\mathbb{T}^d). \tag{2.3}$$

The Sobolev space $H^s(\mathbb{T}^d)$ is the completion of $\mathcal{D}(\mathbb{T}^d)$ under $\langle\cdot,\cdot\rangle_s$.

Proposition 2.1. *Sobolev spaces have the following properties:*

(i) $\mathcal{D}(\mathbb{T}^d) \subset H^s(\mathbb{T}^d) \subset \mathcal{D}'(\mathbb{T}^d)$ *for any* $s \in \mathbb{R}$;

(ii) $H^0(\mathbb{T}^d) = L^2(\mathbb{T}^d)$, *and* $H^s(\mathbb{T}^d) = \{\varphi : D_x^\beta\varphi \in L^2(\mathbb{T}^d), \text{for any } |\beta| \leqslant s\}$ *if* $s \in \mathbb{N}$, *where* $D_x^\beta\varphi$ *are weak derivatives of* φ;

(iii) $H^s(\mathbb{T}^d) \subset H^{s'}(\mathbb{T}^d)$ *if* $s > s'$;

(iv) $C^s(\mathbb{T}^d) \subset H^s(\mathbb{T}^d)$, *and if* $s > \frac{d}{2}$, *then* $H^s(\mathbb{T}^d) \subset C^{s-\frac{d}{2}-\varepsilon}(\mathbb{T}^d)$ *for any small* $\varepsilon > 0$;

(v) *the dual space of* $H^s(\mathbb{T}^d)$, $s > 0$, *is* $H^{-s}(\mathbb{T}^d)$, *and the dual action is exactly the standard* L^2*-paring.*

The Hilbert space $\mathcal{W}$ that we will use in Theorem 2 is of the form

$$\mathcal{W} = H^s(\mathbb{T}^d) \otimes L^2(\mathbb{T}^\ell), \quad s < 0. \tag{2.4}$$

Remark 2.2. *By Proposition 2.1(ii) and (iii), we have* $L^2(\mathbb{T}^d) \subset H^s(\mathbb{T}^d)$ *if* $s < 0$. *Since* $L^2(\mathbb{T}^d) \otimes L^2(\mathbb{T}^\ell) = L^2(\mathbb{T}^{d+\ell})$, *by part (i)* $L^2(\mathbb{T}^{d+\ell}) \subset \mathcal{W} \subset \mathcal{D}'(\mathbb{T}^{d+\ell})$.

For technical treatments, we will also use a different but equivalent inner product on $H^s(\mathbb{T}^d)$. (See Subsection 4.1.)

2.4. **Pseudo-differential operators.** The cotangent bundle over $\mathbb{T}^d$ can be identified as $T^*\mathbb{T}^d \cong \mathbb{T}^d \times \mathbb{R}^d$.

Choose a *Planck's constant* $\hbar \in (0, 1]$.

Let $\omega_\hbar$ be the counting measure over the lattice $(\hbar\mathbb{Z})^d$. Note that $\omega_1 = \omega$ is the same as introduced in Subsection 2.2.

Definition 2.3. *A (complex-valued) function* $a \in C^\infty(T^*\mathbb{T}^d)$ *is called a* symbol of order $m \in \mathbb{R}$ *if for any* $\alpha, \beta \in \mathbb{N}_0^d$, *there is a constant* $C_{\alpha\beta} > 0$ *such that*

$$|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)| \leqslant C_{\alpha\beta}\langle\xi\rangle^{m-|\beta|} \quad \text{for any} \quad (x,\xi) \in T^*\mathbb{T}^d,$$

where $\langle\xi\rangle = \sqrt{1+|\xi|^2}$. *We denote the space of symbols of order* m *by* S^m.

Definition 2.4. *Given a symbol* $a \in S^m$, *the linear operator* $\mathrm{Op}_\hbar(a) : \mathcal{D}(\mathbb{T}^d) \to \mathcal{D}(\mathbb{T}^d)$ *defined by*

$$\begin{aligned} \mathrm{Op}_\hbar(a)\varphi(x) &= \int_{T^*\mathbb{T}^d} a(x,\xi) e^{i2\pi\frac{\xi}{\hbar}\cdot(x-y)} \varphi(y) dy d\omega_\hbar(\xi) \\ &= \int_{T^*\mathbb{T}^d} a(x,\hbar\xi) e^{i2\pi\xi\cdot(x-y)} \varphi(y) dy d\omega(\xi) \end{aligned} \tag{2.5}$$

is called an $\hbar$*-scaled* pseudo-differential operator (PDO) *of order* m *corresponding to the symbol* $a \in S^m$. *We denote the space of* $\hbar$*-scaled PDOs of order* m *by* $\mathrm{OP}_\hbar S^m$.

The formula with $\hbar = 1$ in (2.5) gives the definition of classical pseudo-differential operator $\mathrm{Op}(a) = \mathrm{Op}_1(a)$. We denote $\mathrm{OP}S^m = \mathrm{OP}_1 S^m$. In this way, the $\hbar$-scaled PDO with symbol $a \in S^m$ can be regarded as the classical PDO with symbol $a_\hbar \in S^m$, that is, $\mathrm{Op}_\hbar(a) = \mathrm{Op}(a_\hbar)$, where $a_\hbar(x,\xi) = a(x, \hbar\xi)$.

We see by (2.1) and (2.2), that if $a(x,\xi) = 1$, then $\mathrm{Op}(a) = \mathrm{Id}$; and if $a(x,\xi) = i2\pi\xi$, then $\mathrm{Op}(a) = \dfrac{\partial}{\partial x}$.

By standard duality argument, we extend $\mathrm{Op}_\hbar(a) : \mathcal{D}'(\mathbb{T}^d) \to \mathcal{D}'(\mathbb{T}^d)$. Moreover, $\mathrm{Op}_\hbar(a) : H^s(\mathbb{T}^d) \to H^{s-m}(\mathbb{T}^d)$ is a bounded operator if $a \in S^m$. Some properties about symbols and PDOs are stated in Subsection 2.6-2.8.

2.5. **Fourier integral operators.**

Definition 2.5. *An $\hbar$-scaled* Fourier integral operator (FIO) $\Phi_\hbar : \mathcal{D}(\mathbb{T}^d) \to \mathcal{D}(\mathbb{T}^d)$ *with* amplitude $a \in S^m$ *and* phase $S \in C^\infty(\mathbb{T}^d \times (\mathbb{R}^d \backslash \{0\}))$ *is of the form*

$$\begin{aligned}\Phi_\hbar \varphi(x) &= \int_{T^*\mathbb{T}^d} a(x,\xi) e^{i2\pi \frac{1}{\hbar}[S(x,\xi) - y\cdot\xi]} \varphi(y) dy d\omega_\hbar(\xi) \\ &= \int_{T^*\mathbb{T}^d} a(x,\hbar\xi) e^{i2\pi [S(x,\xi) - y\cdot\xi]} \varphi(y) dy d\omega(\xi),\end{aligned}$$

where the phase function $S(x,\xi)$ *satisfies the following conditions:*

(1) $S(x,\xi)$ is homogeneous of degree 1 in ξ for all $|\xi| \neq 0$;

(2) $S(x,\xi)$ is non-degenerate, that is, $\det\left(\dfrac{\partial^2 S}{\partial x \partial \xi}\right) \neq 0$.

The classical Fourier integral operator $\Phi = \Phi_\hbar : \mathcal{D}(\mathbb{T}^d) \to \mathcal{D}(\mathbb{T}^d)$ *is the one with* $\hbar = 1$.

Remark 2.6. (i) *If we take $S(x,\xi) = x \cdot \xi$, then $\Phi_\hbar$ becomes an $\hbar$-scaled pseudo-differential operator.*

(ii) *If $a(x,\xi) = a(x)$, a function independent of ξ, and $S(x,\xi) = R(x)\cdot\xi$ for any map $R : \mathbb{T}^d \to \mathbb{T}^d$, then $\Phi_\hbar\varphi(x) = a(x)\varphi(R(x))$ by* (2.1) *and* (2.2).

By standard duality argument, we can extend $\Phi_\hbar : \mathcal{D}'(\mathbb{T}^d) \to \mathcal{D}'(\mathbb{T}^d)$. Further, $\Phi_\hbar : H^s(\mathbb{T}^d) \to H^{s-m}(\mathbb{T}^d)$ is a bounded operator if its amplitude $a \in S^m$.

Definition 2.7. *The canonical transformation associated to an $\hbar$-scaled FIO with phase S is the transformation $(x,\xi) \mapsto (y,\eta)$ given by*

$$y = \frac{\partial S(x,\eta)}{\partial \eta}, \qquad \xi = \frac{\partial S(x,\eta)}{\partial x}. \tag{2.6}$$

In other words, the phase function S serves as the generating function of the canonical transformation.

2.6. **Principal symbol and symbol calculus.**

If $m < m'$, then $S^m \subset S^{m'}$ and $\mathrm{OP}_\hbar S^m \subset \mathrm{OP}_\hbar S^{m'}$. Set $S^{-\infty} = \bigcap_{m\in\mathbb{R}} S^m$. If $a \in S^{-\infty}$, then $\mathrm{Op}_\hbar(a)$ is a smoothing (and hence compact) operator.

Definition 2.8. *A symbol $a_0 \in S^m$ is called a* classical principal symbol *of* $\mathrm{Op}(a) \in \mathrm{OP}S^m$ *if $a - a_0 \in S^{m-1}$ and $a_0(x,\xi)$ is homogeneous in ξ of degree m for all $(x,\xi) \in T^*\mathbb{T}^d$ with $|\xi|$ sufficiently large.*

A symbol $a_0 \in S^m$ is called a semiclassical principal symbol *of* $\mathrm{Op}_\hbar(a) \in \mathrm{OP}_\hbar S^m$ *if $a_\hbar - (a_0)_\hbar \in \hbar S^m$ as $\hbar \to 0$, that is, for any $\alpha, \beta \in \mathbb{N}_0^d$, there is $C_{\alpha\beta} > 0$ and $\hbar_{\alpha\beta} > 0$ such that for all $(x,\xi) \in T^*\mathbb{T}^d$ and $\hbar \in (0, \hbar_{\alpha\beta}]$,*

$$|\partial_x^\alpha \partial_\xi^\beta (a(x,\hbar\xi) - a_0(x,\hbar\xi))| \leqslant C_{\alpha\beta}\hbar\langle\xi\rangle^{m-|\beta|}.$$

Note that usually the classical principal symbol for $\mathrm{Op}(a)$ and the semiclassical principal symbol for $\mathrm{Op}_\hbar(a)$ do not coincide. Also, principal symbol is not unique in S^m but unique in the quotient class S^m/S^{m-1}.

Theorem 2.9. *For classical PDOs, we have the following.*

(1) Adjoint: If $A \in \mathrm{OP}S^m$ has a principal symbol a_0, then the adjoint operator $A^* \in \mathrm{OP}S^m$ has a principal symbol $\overline{a_0}$.
(2) Composition: If $A \in \mathrm{OP}S^m$ has a principal symbol a_0 and $B \in \mathrm{OP}S^{m'}$ has a principal symbol b_0, then the compositions $A \circ B, B \circ A \in \mathrm{OP}S^{m+m'}$ both have a principal symbol $a_0 b_0$.
(3) Inverse: If $A \in \mathrm{OP}S^m$ has a principal symbol a_0 and is invertible, then $A^{-1} \in \mathrm{OP}S^{-m}$ has a principal symbol a_0^{-1}.

The above rules are also true for semiclassical PDOs when $A \in \mathrm{OP}_\hbar S^m$ with a semiclassical principal symbol a_0 and $B \in \mathrm{OP}_\hbar S^{m'}$ with a semiclassical principal symbol b_0.

2.7. **Egorov's Theorem.** We first state the original version of Egorov's theorem in [4] for the invertible case.

Theorem 2.10. *Let $A \in \mathrm{OP}S^m$ with principal symbol a_0, and Φ be a classical FIO with amplitude $b \in S^0$ and phase S. Let $\mathcal{F}(x,\xi) = (y,\eta)$ be the canonical transformation associated to Φ, and assume that there is a domain $\Omega \subset T^*\mathbb{T}^d$ such that $\mathcal{F} : \Omega \to \mathcal{F}(\Omega)$ is bijective. Then the operator $\Phi^* A \Phi|_{\mathcal{F}(\Omega)} \in \mathrm{OP}S^m$ has a principal symbol $\overline{a}_0$ such that*

$$\overline{a}_0(y,\eta) = \overline{a}_0(\mathcal{F}(x,\xi)) = a_0(x,\xi)|b(x,\xi)|^2 \left|\det\left(\frac{\partial^2 S}{\partial x \partial \xi}\right)\right|^{-1}.$$

For our purpose, we need the following version of Egorov's theorem.

Theorem 2.11. *Let $A \in \mathrm{OP}S^m$ with principal symbol a_0, and Φ be a classical FIO with amplitude $b \in S^0$ and phase S. Let $\mathcal{F}(x,\xi) = (y,\eta)$ be the canonical transformation associated to Φ, and assume that $\mathcal{F}$ is a surjective local diffeomorphism with finite inverse branches. Then the operator $\Phi^* A \Phi \in \mathrm{OP}S^m$ has a principal symbol $\overline{a}_0$ such that*

$$\overline{a}_0(y,\eta) = \sum_{\mathcal{F}(x,\xi)=(y,\eta)} a_0(x,\xi)|b(x,\xi)|^2 \left|\det\left(\frac{\partial^2 S}{\partial x \partial \xi}\right)\right|^{-1}. \tag{2.7}$$

Proof. Given an inverse branch $\mathcal{G}$ of $\mathcal{F}$, denote $\Omega_\mathcal{G} = \mathcal{G}(T^*\mathbb{T}^d)$. By Theorem 2.10,

$$\Phi^* A \Phi = \sum_{\mathcal{G}} \Phi^* A \Phi|_{\Omega_\mathcal{G}} \in \mathrm{OP}S^m$$

has a principal symbol of the form in (2.7). ☐

Remark 2.12. *We can easily adapt the proof of Theorem 2.10 and 2.11 in the semiclassical situation and show that the semiclassical principal symbol of $\Phi_\hbar^* A \Phi_\hbar$*

is still given by (2.7), where $A \in \mathrm{OP}_\hbar S^m$ *has a semiclassical principal symbol* a_0 *and* $\Phi_\hbar$ *is the* $\hbar$*-scaled FIO with amplitude* $b \in S^0$ *and phase* S*.*

2.8. **L^2-Continuity.** We first state a version of L^2-continuity for a classical PDO of order 0 established in [6].

Theorem 2.13. *If* $a(x,\xi) \in S^0$*, then* $\mathrm{Op}(a) : L^2(\mathbb{T}^d) \to L^2(\mathbb{T}^d)$ *is a bounded operator. Moreover, for any* $\varepsilon > 0$*, there is a decomposition*

$$\mathrm{Op}(a) = K + R$$

such that $K : L^2(\mathbb{T}^d) \to L^2(\mathbb{T}^d)$ *is a compact operator and*

$$\|R\|_{L^2\to L^2} \leqslant \sup_x \limsup_{|\xi|\to\infty} |a(x,\xi)| + \varepsilon = \sup_x \limsup_{|\xi|\to\infty} |a_0(x,\xi)| + \varepsilon,$$

where a_0 *is a principal symbol of* $\mathrm{Op}(a)$*.*

For a semiclassical PDO of order 0, we need a version of Carderon-Vaillancourt theorem established in [8].

Theorem 2.14. *If* $a(x,\xi) \in S^0$*, then* $\mathrm{Op}_\hbar(a) : L^2(\mathbb{T}^d) \to L^2(\mathbb{T}^d)$ *is a bounded operator. Moreover, let* $a_0 \in S^0$ *be a semiclassical principal symbol of* $\mathrm{Op}_\hbar(a)$*, then as* $\hbar \to 0$*,*

$$\|\mathrm{Op}_\hbar(a)\|_{L^2\to L^2} \leqslant \sup_{(x,\xi)\in T^*\mathbb{T}^d} |a_0(x,\xi)| + \hbar C(a,d),$$

where the constant $C(a,d)$ *only depends on the* C^k*-norms of the symbol* a*,* $0 \leqslant k \leqslant 2d$*.*

3. Spectral Gap and Coboundary: Proof of the Theorems

3.1. **Decomposition of Koopman operator.** In this subsection we decompose the Koopman operator $\widehat{F} : \mathcal{W} \to \mathcal{W}$ according to fiberwise Fourier expansion, where $\mathcal{W} = H^s(\mathbb{T}^d) \otimes L^2(\mathbb{T}^\ell)$ is given in (2.4).

Given $\phi \in \mathcal{W}$, we write the Fourier series expansion along $\mathbb{T}^\ell$-direction as

$$\phi(x,y) = \sum_{\nu\in\mathbb{Z}^\ell} \phi_\nu(x) e^{i2\pi\nu\cdot y},$$

where the Fourier coefficients are defined by

$$\phi_\nu(x) = \int_{\mathbb{T}^\ell} \phi(x,y) e^{-i2\pi\nu\cdot y} dy \in H^s(\mathbb{T}^d), \qquad \nu \in \mathbb{Z}^\ell.$$

Note that for each Fourier mode $\nu \in \mathbb{Z}^\ell$,

$$\widehat{F}(\phi_\nu(x) e^{i2\pi\nu\cdot y}) = [\phi_\nu(Tx) e^{i2\pi\nu\cdot\tau(x)}] e^{i2\pi\nu\cdot y},$$

and it can be shown that $\phi_\nu(Tx)e^{i2\pi\nu\cdot\tau(x)} \in H^s(\mathbb{T}^d)$.[1] This observation suggests an $\widehat{F}$-invariant decomposition

$$\mathcal{W} = H^s(\mathbb{T}^d) \otimes L^2(\mathbb{T}^\ell) = \bigoplus_{\nu\in\mathbb{Z}^\ell} H^s_\nu,$$

[1]This fact is easy to show for $s \in \mathbb{N} \cup \{0\}$, and hence is also true when s is a negative integer by duality. For the general case, treat H^s as the interpolation between $H^{\lfloor s\rfloor}$ and $H^{\lfloor s\rfloor+1}$. See Section 4.2 in [14] for details.

where $H^s_\nu := \{\varphi(x)e^{i2\pi\nu\cdot y} : \varphi \in H^s(\mathbb{T}^d)\} \cong H^s(\mathbb{T}^d)$. Correspondingly, we decompose $\widehat{F} = \bigoplus_{\nu\in\mathbb{Z}^\ell} \widehat{F}_\nu$, where each $\widehat{F}_\nu = \widehat{F}|H^s_\nu \cong \widehat{F}|H^s(\mathbb{T}^d)$ acts by

$$\widehat{F}_\nu\varphi(x) = \varphi(Tx)e^{i2\pi\nu\cdot\tau(x)}, \qquad \varphi \in H^s(\mathbb{T}^d). \tag{3.1}$$

In the case when $s < 0$, using the fact that $H^s(\mathbb{T}^d)$ is the dual space of $H^{-s}(\mathbb{T}^d)$ via L^2-paring, we get

$$\widehat{F}^*_\nu\psi(x) = \sum_{Ty=x} \frac{e^{-i2\pi\nu\cdot\tau(y)}}{|\mathrm{Jac}(T)(y)|}\psi(y), \qquad \psi \in H^{-s}(\mathbb{T}^d). \tag{3.2}$$

In other words, $\widehat{F}^*|H^{-s}(\mathbb{T}^d)$ is the RPF transfer operator over $T : \mathbb{T}^d \to \mathbb{T}^d$ for the complex potential function $-\log|\mathrm{Jac}(T)| - i2\pi\nu\cdot\tau$. In the case when $\nu = \mathbf{0}$, we have that $\widehat{F}^*_{\mathbf{0}}h = h$, that is, the density function $h(x)$ of μ w.r.t. dx is provided by the eigenvector corresponding to the leading simple eigenvalue 1 of $\widehat{F}^*_{\mathbf{0}}$. See [13] for more details.

We shall use the fact $\widehat{F}^*_{\mathbf{0}}h = h$ in the following particular form:

$$\sum_{Ty=x} \mathcal{A}(y) = 1, \quad \text{for all } x \in \mathbb{T}^d,$$

where

$$\mathcal{A}(y) = \frac{1}{|\mathrm{Jac}(T)(y)|}\frac{h(y)}{h(Ty)}. \tag{3.3}$$

Similarly, we have for all $n \in \mathbb{N}$,

$$\sum_{T^ny=x} \mathcal{A}_n(y) = 1, \quad \text{for all } x \in \mathbb{T}^d,$$

where

$$\mathcal{A}_n(y) = \frac{1}{|\mathrm{Jac}(T^n)(y)|}\frac{h(y)}{h(T^ny)}. \tag{3.4}$$

3.2. **Spectral gap.** Recall that the notion of spectral gap is given right before Theorem 2 is stated (see (1.4)).

According to the decomposition of $\widehat{F} : \mathcal{W} \to \mathcal{W}$, the spectral gap property follows from the following propositions. The proof of the propositions will be given in the next section, using the classical and semiclassical analysis theory.

Proposition 3.1. *Let $s < 0$ and $\nu \in \mathbb{Z}^\ell$. There are $C_1 > 0$ and $\rho_1 \in (0,1)$ such that $\widehat{F}_\nu : H^s(\mathbb{T}^d) \to H^s(\mathbb{T}^d)$ can be written as*

$$\widehat{F}_\nu = K_\nu + R_\nu,$$

where K_ν is a compact operator and

$$\|R^n_\nu|H^s(\mathbb{T}^d)\| \leqslant C_1\rho_1^n, \quad n \in \mathbb{N}. \tag{3.5}$$

Proposition 3.2. *Let $s < 0$ and assume that τ is not an essential coboundary. There are $\rho_2 \in (0,1)$ and $\nu_1 > 0$ such that for any $\nu \in \mathbb{Z}^\ell$ with $|\nu| \geqslant \nu_1$, the spectral radius*

$$\mathrm{Sp}(\widehat{F}_\nu|H^s(\mathbb{T}^d)) \leqslant \rho_2. \tag{3.6}$$

Remark 3.3. (i) *The quasi-compactness property is well known for Ruelle-Perron-Frobenius transfer operator on Hölder function spaces over expanding maps. Proposition 3.1 can be regarded as its dual version. The estimate in (3.5) shows that the essential spectral radius of $\widehat{F}_\nu$ is no more than ρ_1;*

(ii) *Proposition 3.2 shows that the operator $\widehat{F}_\nu|H^s(\mathbb{T}^d)$ is essentially a contraction when the Fourier mode ν is very large, since the spectral radius of $\widehat{F}_\nu$ is no more than ρ_2.*

3.3. **Proof of Theorem 2.** Recall that the space $\mathcal{W} = H^s(\mathbb{T}^d)\otimes L^2(\mathbb{T}^\ell)$ is defined in (2.4), where $s < 0$.

Lemma 3.4. *The spectral radius* $\mathrm{Sp}(\widehat{F}_\nu|H^s(\mathbb{T}^d)) \leqslant 1$ *for* $\nu \in \mathbb{Z}^\ell$.

Proof. The proof is similar as in [6], as we sketch here.

Choose $\rho_3 \in (0,1)$ such that $\max\{\rho_1,\rho_2\} < \rho_3 < 1$, where ρ_1 and ρ_2 are given in Proposition 3.1 and 3.2 respectively. By Proposition 3.1, we can rewrite

$$\widehat{F}_\nu = K_\nu + R_\nu = (K_\nu^1 + K_\nu^2) + R_\nu = K_\nu^1 + (K_\nu^2 + R_\nu) = K_\nu^1 + R'_\nu$$

such that the spectral radius of R'_ν is less than ρ_3. This can be done by defining K_ν^1 and K_ν^2 to be the spectral projection of K_ν outside and inside the circle $\{z : |z| = \rho_3\}$ respectively. K_ν^1 has finite rank since K_ν is compact. The general Jordan decomposition of K_ν^1 can be written

$$K_\nu^1 = \sum_{i=1}^k \left(\lambda_i \sum_{j=1}^{d_i} v_{ij}\otimes w_{ij} + \sum_{j=1}^{d_i-1} v_{ij}\otimes w_{i(j+1)}\right)$$

where d_i is the dimension of the Jordan block associated with the eigenvalue λ_i, with $v_{ij} \in H^s(\mathbb{T}^d)$ and $w_{ij} \in H^{-s}(\mathbb{T}^d)$. We arrange eigenvalues such that $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_k|$.

Now if $|\lambda_1| > 1$, we can choose $\varphi,\psi \in \mathcal{D}(\mathbb{T}^d)$ such that $v_{11}(\overline{\psi}) \neq 0$ and $w_{11}(\varphi) \neq 0$ since $\mathcal{D}(\mathbb{T}^d)$ is dense in both $H^s(\mathbb{T}^d)$ and $H^{-s}(\mathbb{T}^d)$. On one hand,

$$|(\psi, \widehat{F}_\nu^n\varphi)_{H^{-s},H^s}| = \left|\int_{\mathbb{T}^d} \psi \widehat{F}_\nu^n \varphi dx\right| \leqslant |\psi|_{C^0}|\varphi|_{C^0}.$$

On the other hand,

$$|(\psi, \widehat{F}_\nu^n\varphi)_{H^{-s},H^s}| \geqslant |(\psi, (K_\nu^1)^n\varphi)_{H^{-s},H^s}| - |(\psi, (R'_\nu)^n\varphi)_{H^{-s},H^s}|.$$

The second term converges to 0 since $\|(R'_\nu)^n|H^s(\mathbb{T}^d)\| = O(\rho_3^n)$, while the first term

$$|(\psi, (K_\nu^1)^n\varphi)_{H^{-s},H^s}| = \left|\sum_{i=1}^k \sum_{r=0}^{\min(n,d_i-1)} \binom{n}{r}\lambda_i^{n-r}\sum_{j=1}^{d_i-r} v_{ij}(\overline{\psi})w_{i(j+r)}(\varphi)\right| \tag{3.7}$$

has a leading growth $|\lambda_1|^n|v_{11}(\overline{\psi})||w_{11}(\varphi)| \to \infty$ as $n \to \infty$, which is a contradiction. Therefore, all eigenvalues of K_ν^1 are of modulus no more than 1, and so are eigenvalues of $\widehat{F}_\nu$. □

Lemma 3.5. *If τ is not an essential coboundary, then the spectral radius* $\mathrm{Sp}(\widehat{F}_\nu|H^s(\mathbb{T}^d)) < 1$ *for $\nu \in \mathbb{Z}^\ell\backslash\{\mathbf{0}\}$. Moreover, 1 is the only eigenvalue of $\widehat{F}_{\mathbf{0}}$ on the unit circle, which is simple with eigenspace of constant functions.*

Proof. The proof is essentially due to Pollicott [11]. Let λ be an eigenvalue of $\widehat{F}_\nu$ with modulus 1, and $\varphi \in H^s(\mathbb{T}^d) \subset H^{s-\frac{d}{2}-1}(\mathbb{T}^d)$ be a corresponding eigenvector such that $\widehat{F}_\nu \varphi = \lambda\varphi$. It is sufficient to show that $\nu = \mathbf{0}$, $\lambda = 1$, and φ is a constant function.

By duality, there is $\psi \in H^{-s+\frac{d}{2}+1} \subset C^1(\mathbb{T}^d)$ such that $\widehat{F}_\nu^* \psi = \overline{\lambda}\psi$. Let $\lambda = e^{i2\pi c}$ for some $c \in \mathbb{T}$, and set $\overline{\psi} = \frac{\psi}{h}$, where $h(x)$ is the density function of μ w.r.t. dx. By the definition of $\widehat{F}_\nu^*$ in (3.2) and $\mathcal{A}(y)$ in (3.3), we have

$$\sum_{Ty=x} \mathcal{A}(y) e^{i2\pi[c-\nu\cdot\tau(y)]}\overline{\psi}(y) = \overline{\psi}(x). \tag{3.8}$$

Now choose z such that $|\overline{\psi}(z)|$ obtains maximum. Since $\sum_{Ty=z} \mathcal{A}(y) = 1$, we must have $|\overline{\psi}(y)| = |\overline{\psi}(z)|$ for all $y \in T^{-1}(z)$. By induction, we get that $|\overline{\psi}(y)| = |\overline{\psi}(z)|$ for all $y \in \bigcup_{n=1}^\infty T^{-n}(z)$. Since T is mixing, the set $\bigcup_{n=1}^\infty T^{-n}(z)$ is dense in $\mathbb{T}^d$, and hence $|\overline{\psi}(x)| = |\overline{\psi}(z)|$ is constant. Thus (3.8) is a convex combination of points of a circle which lies on the circle. From this we deduce that

$$e^{i2\pi[c-\nu\cdot\tau(y)]}\overline{\psi}(y) = \overline{\psi}(Ty)$$

for all $y \in \mathbb{T}^d$, and hence

$$\nu \cdot \tau(y) = c + \frac{1}{2\pi} \arg \overline{\psi}(y) - \frac{1}{2\pi} \arg \overline{\psi}(Ty).$$

Since τ is not an essential coboundary, we must have $\nu = \mathbf{0}$. By integrating w.r.t. $d\mu$, we also have $c = 0$ and $\arg \overline{\psi} \equiv$ constant. Thus, $\lambda = 1$ and $\overline{\psi} \equiv$ constant. □

Now we are ready to prove the spectral gap property for $\widehat{F} : \mathcal{W} \to \mathcal{W}$.

Proof of Theorem 2. To sum up, by Proposition 3.1, 3.2 and Lemma 3.4, 3.5, we have $\mathrm{Spec}(\widehat{F}_0) = \{1\} \cup \mathcal{K}_0$, where 1 is an simple eigenvalue of $\widehat{F}_0$ and $\mathcal{K}_0$ is a compact subset of the open unit disk. Thus, there is $\rho_4 \in (0, 1)$ such that $\mathcal{K}_0 \subset \{z \in \mathbb{C} : |z| < \rho_4\}$. Also,

$$\mathrm{Sp}(\widehat{F}_\nu) \begin{cases} \leqslant \rho_2, & \text{for all } |\nu| \geqslant \nu_1, \\ < 1, & \text{for all } 0 < |\nu| \leqslant \nu_1, \end{cases}$$

where ν_1 is given in Proposition 3.2. Hence we can assume that $\bigcup_{\nu\neq\mathbf{0}} \mathrm{Spec}(\widehat{F}_\nu)$ is inside $\{z \in \mathbb{C} : |z| < \rho_4\}$ (by enlarging ρ_4 if necessary). Adding the spectrums of $\{\widehat{F}_\nu\}_{\nu\in\mathbb{Z}^\ell}$ together, we get that $\widehat{F} = \oplus \widehat{F}_\nu$ has spectrum

$$\mathrm{Spec}(\widehat{F}) = \{1\} \cup \mathcal{K},$$

where $\mathcal{K} = \mathcal{K}_0 \cup \bigcup_{\nu\neq\mathbf{0}} \mathrm{Spec}(\widehat{F}_\nu) \subset \{z \in \mathbb{C} : |z| < \rho_4\}$. □

3.4. **Proof of Theorem 1.** Now we use Theorem 2 to prove Theorem 1. What we need to do is to show that if $\widehat{F} : \mathcal{W} \to \mathcal{W}$ has a spectral gap, then it is exponentially mixing. In the proof we regard the observables ϕ and ψ as elements in $\mathcal{W}$ and $\mathcal{W}'$ respectivly, and use the fact that the dual action is standard L^2-paring.

Proof of Theorem 1. Since $\widehat{F} : \mathcal{W} \to \mathcal{W}$ has a spectral gap, we can write $\widehat{F} = \mathcal{P} + \mathcal{N}$ such that

(1) $\mathcal{P}$ is a 1-dimensional projection, i.e., $\mathcal{P}^2 = \mathcal{P}$;
(2) $\mathcal{N}$ is a bounded operator with spectral radius $\mathrm{Sp}(\mathcal{N}) < 1$;
(3) $\mathcal{P}\mathcal{N} = \mathcal{N}\mathcal{P} = 0$.

From the proof of Theorem 2, we know that 1 is the greatest simple eigenvalue for $\widehat{F}_0$ with eigenvector 1 as well as for $\widehat{F}_0^*$ with eigenvector h, and therefore, $\mathcal{P} = 1 \otimes h$. Further, $\mathrm{Sp}(\mathcal{N}) \leqslant \rho_4 < 1$, then there is $C_2 > 0$ such that $\|\mathcal{N}^n\| \leqslant C_2 \rho_4^n$ for all $n \in \mathbb{N}$.

Suppose $\phi \in L^\infty(\mathbb{T}^{d+\ell})$ and $\psi \in C^\alpha(\mathbb{T}^{d+\ell})$ are given. Pick $s \in [-\alpha, 0)$ and let $\mathcal{W} = H^{-s}(\mathbb{T}^d) \otimes L^2(\mathbb{T}^\ell)$. Then the dual space of $\mathcal{W}$ is $\mathcal{W}' = H^{-s}(\mathbb{T}^d) \otimes L^2(\mathbb{T}^\ell)$. Note that $C^\alpha(\mathbb{T}^{d+\ell}) \subset \mathcal{W}'$ and $L^\infty(\mathbb{T}^{d+\ell}) \subset L^2(\mathbb{T}^{d+\ell}) \subset \mathcal{W}$. We have $\phi \in \mathcal{W}$ and $\psi \in \mathcal{W}'$. Hence,

$$\begin{aligned}\int (\phi \circ F^n)\psi dA =&(\psi h, \widehat{F}^n(\overline{\phi}))_{\mathcal{W}',\mathcal{W}}\\ =&(\psi h, \mathcal{P}(\overline{\phi}))_{\mathcal{W}',\mathcal{W}} + (\psi h, \mathcal{N}^n(\overline{\phi}))_{\mathcal{W}',\mathcal{W}}\\ =&1 \otimes h(\psi h, \overline{\phi}) + (\psi h, \mathcal{N}^n(\overline{\phi}))_{\mathcal{W}',\mathcal{W}}\\ =&\int \psi dA \int \phi dA + (\psi h, \mathcal{N}^n(\overline{\phi}))_{\mathcal{W}',\mathcal{W}}.\end{aligned}$$

That is, the correlation function

$$C_n(\phi, \psi; F, dA) = |(\psi h, \mathcal{N}^n(\overline{\phi}))_{\mathcal{W}',\mathcal{W}}| \leqslant \|\mathcal{N}^n\| \|\psi h\|_{\mathcal{W}'} \|\phi\|_{\mathcal{W}} \leqslant C_{\phi,\psi} \rho_4^n$$

where $C_{\phi,\psi} = C_2 \|\psi h\|_{\mathcal{W}'} \|\phi\|_{\mathcal{W}}$. □

Remark 3.6. *Using some Sobolev inequalities, it is not hard to get that* $\|\psi h\|_{\mathcal{W}'} \leqslant C_3 \|\psi\|_{C^\alpha} \|h\|_{C^1}$ *and* $\|\phi\|_{\mathcal{W}} \leqslant C_4 \|\phi\|_{L^\infty}$*, and hence* $C_{\phi,\psi} \leqslant C_5 \|\phi\|_{C^\alpha} \|\psi\|_{L^\infty}$*.*

3.5. **Proof of Theorem 3.** Now we show the characters of the system in the case that $\tau(s)$ is an essential coboundary.

Proof of Theorem 3. (i)⇒(ii). Suppose $\tau(x)$ is an essential coboundary. For any $(x, y) \in \mathbb{T}^d \times \mathbb{T}^\ell$, we define

$$\mathcal{L}(x, y) = \{(x', y') \in \mathbb{T}^d \times \mathbb{T}^\ell : \nu \cdot y' + u(x') = \nu \cdot y + u(x)\},$$

where $\nu \in \mathbb{Z}^\ell \backslash \{\mathbf{0}\}$ and $u : \mathbb{T}^d \to \mathbb{T}$ is given by Definition 1.1. Since u is a smooth map, $\mathcal{L}(x, y)$ is a smooth $d + \ell - 1$ dimensional manifold, and $\{\mathcal{L}(x, y) : (x, y) \in \mathbb{T}^d \times \mathbb{T}^\ell\}$ form a foliation of $\mathbb{T}^d \times \mathbb{T}^\ell$.

For $(x', y') \in \mathcal{L}(x, y)$, the definition of F gives

$$F(x, y) = (Tx, y + \tau(x)) \quad \text{and} \quad F(x', y') = (Tx', y' + \tau(x')).$$

By (1.3) we get

$$\nu \cdot (y + \tau(x)) + u(Tx) = \nu \cdot y + \nu \cdot \tau(x) + u(Tx) = \nu \cdot y + c + u(x)$$

and similarly $\nu \cdot (y' + \tau(x')) + u(Tx') = \nu \cdot y' + c + u(x')$. Hence we obtain

$$\nu \cdot (y' + \tau(x')) + u(Tx') = \nu \cdot (y + \tau(x)) + u(Tx).$$

By definition of $\mathcal{L}$, we get $F(x', y') \in \mathcal{L}(F(x, y))$, that is, the foliation is F invariant.

(i)⇒(iii). Define a continuous map $\pi : \mathbb{T}^d \times \mathbb{T}^\ell \to \mathbb{T}^d \times \mathbb{T}$ by $\pi(x, y) = (x, \nu \cdot y + u(x))$ and a diffeomorphism $G : \mathbb{T}^d \times \mathbb{T} \to \mathbb{T}^d \times \mathbb{T}$ by $G(x, y) = (Tx, y + c)$, where both ν and c are given by Definition 1.1. So by (1.3) we have

$$\begin{aligned}\pi\big(F(x, y)\big) =&\pi\big(Tx,\ y + \tau(x)\big) = \big(Tx,\ \nu \cdot (y + \tau(x)) + u(Tx)\big)\\ =&\big(Tx,\ \nu \cdot y + c + u(x)\big) = G\big(x,\ \nu \cdot y + u(x)\big) = G\big(\pi(x, y)\big).\end{aligned}$$

It is obvious that the map $G = T \times R_c : \mathbb{T}^d \times \mathbb{T} \to \mathbb{T}^d \times \mathbb{T}$ is semiconjugate to the circle rotation $R_c : \mathbb{T} \to \mathbb{T}$.

(ii)⇒(iii). Let $p \in \mathbb{T}^d$ be a fixed point of T. Restricted to $\{p\} \times \mathbb{T}^\ell$, the leaves of the foliation $\mathcal{L}$ become $(\ell - 1)$ dimensional tori, and $F : \{p\} \times \mathbb{T}^\ell \to \{p\} \times \mathbb{T}^\ell$ preserve the leaves. Hence the quotient space $\mathbb{T}^\ell/\sim$ is a circle $\mathbb{T}$, where $y \sim y'$ if y and y' are in the same leave of $\mathcal{L}|_{\{p\}\times\mathbb{T}^\ell}$. Let $\pi|_{\{p\}\times\mathbb{T}^\ell} : \{p\} \times \mathbb{T}^\ell \to \{p\} \times \mathbb{T}$ be the quotient map, and $G|_{\{p\}\times\mathbb{T}} : \{p\} \times \mathbb{T} \to \{p\} \times \mathbb{T}$ be the map induced by F. Clearly $\pi|_{\{p\}\times\mathbb{T}}$ is continuous and $G|_{\{p\}\times\mathbb{T}}$ is a rotation, denoted by R_c, where $c \in \mathbb{T}$.

$\pi|_{\{p\}\times\mathbb{T}}$ and $G|_{\{p\}\times\mathbb{T}}$ can be extended to map $\pi : \mathbb{T}^d \times \mathbb{T}^\ell \to \mathbb{T}^d \times \mathbb{T}$ and $G : \mathbb{T}^d \times \mathbb{T} \to \mathbb{T}^d \times \mathbb{T}$ in a natural way such that $\pi = \mathrm{id}_{\mathbb{T}^d} \times \pi|_{\{p\}\times\mathbb{T}}$ and $G = T \times R_c$, where π_1 is a map from $\mathbb{T}^\ell$ to $\mathbb{T}$. That is, for any $(x, y) \in \mathbb{T}^d \times \mathbb{T}^\ell$, let $y' \in \mathcal{L}(x, y) \cap (\{p\} \times \mathbb{T}^d)$, and define $\pi(x, y) = (x, \pi|_{\{p\}\times\mathbb{T}^\ell}(y'))$; and for any $(x, \bar{y}) \in \mathbb{T}^d \times \mathbb{T}$ define $G(x, \bar{y}) = (Tx, G_{\{p\}\times\mathbb{T}}(\bar{y})) = (Tx, R_c(\bar{y}))$. It is easy to check that $\pi \circ G = F \circ \pi$.

(iii)⇒(iv). This is because the circle rotation is not weak mixing.

(iv)⇒(i). If the map is not mixing, it cannot be exponentially mixing. Then we use the dichotomy in Theorem 1. □

4. Spectrums of $\widehat{F}_\nu$: Proof of Proposition 3.1 and 3.2

We shall use the classical formulation to prove Proposition 3.1, and the semi-classical one to prove Proposition 3.2.

4.1. The spaces and operators. We first construct a particular symbol on $T^*\mathbb{T}^d$ as follows. Choose

$$R > \frac{\max\{1,\ 2\|D\tau\|\}}{\gamma - 1}, \tag{4.1}$$

where γ is given in (1.1), and let $g_0 \in C^\infty(\mathbb{R}^+)$ be such that

$$g_0(t) = \begin{cases} 1, & t \leqslant R, \\ t, & t \geqslant R + 1, \end{cases} \tag{4.2}$$

and $g_0(t)$ is strictly increasing for $t \in [R, \infty)$. Set $g(\xi) = g_0(|\xi|)$ for $\xi \in \mathbb{R}^d$. Given $s < 0$, define a symbol

$$\lambda_s(x, \xi) = h(x)^{\frac{1}{2}} g(\xi)^s \in S^s, \tag{4.3}$$

where $h(x)$ is the density function of μ w.r.t. dx.

Denote $\Lambda_{s,\hbar} = \mathrm{Op}_\hbar(\lambda_s) \in \mathrm{OP}_\hbar S^s$ and $\Lambda_s = \Lambda_{s,1}$. We define a parameter family of inner products on $H^s(\mathbb{T}^d)$ by

$$\langle \varphi, \psi \rangle_{\Lambda_{s,\hbar}} = \langle \Lambda_{s,\hbar}\varphi, \Lambda_{s,\hbar}\psi \rangle_{L^2}, \qquad \varphi, \psi \in H^s(\mathbb{T}^d),\ \hbar \in (0, 1].$$

It is easy to check that every inner product in this family is equivalent to the standard one $\langle \cdot, \cdot \rangle_s$ defined in (2.3) since $\lambda_s(x, \xi) \asymp \langle \xi \rangle^s$ for $|\xi|$ large. When equipped with the inner product $\langle \cdot, \cdot \rangle_{\Lambda_{s,\hbar}}$, $H^s(\mathbb{T}^d)$ is denoted by $H_{\Lambda_{s,\hbar}}(\mathbb{T}^d)$ instead. The Sobolev space $H_{\Lambda_{s,\hbar}}$ is unitarily equivalent to the L^2 space, that is,

$$\Lambda_{s,\hbar} H_{\Lambda_{s,\hbar}}(\mathbb{T}^d) \cong L^2(\mathbb{T}^d), \quad \text{or} \quad H_{\Lambda_{s,\hbar}}(\mathbb{T}^d) \cong \Lambda_{s,\hbar}^{-1} L^2(\mathbb{T}^d).$$

4.2. **Proof of Proposition 3.1.** Recall that $\widehat{F}_\nu$ is given in (3.1).

Proof of Proposition 3.1. In Definition 2.5, we choose $\hbar = 1$, and choose amplitude $a(x,\xi) = e^{i2\pi\nu\cdot\tau(x)} \in S^0$ and phase $S(x,\xi) = Tx\cdot\xi$. Then by (3.1) and Remark 2.6(ii), we have

$$\widehat{F}_\nu\varphi(x) = a(x,\xi)\varphi(Tx) = \int_{T^*\mathbb{T}^d} e^{i2\pi\nu\cdot\tau(x)} e^{i2\pi[Tx\cdot\xi - y\cdot\xi]}\varphi(y)dyd\omega(\xi)$$

for any $\nu\in\mathbb{Z}^\ell$. The canonical transformation $(x,\xi)\mapsto(y,\eta)$ associated to $\widehat{F}_\nu$ is given by

$$y = Tx, \qquad \eta = [(D_xT)^t]^{-1}\xi.$$

With $\Lambda_s = \Lambda_{s,1}$ constructed in the previous subsection, we get the following commutative diagram

$$\begin{array}{ccc} H_{\Lambda_s}(\mathbb{T}^d) & \xrightarrow{\widehat{F}_\nu} & H_{\Lambda_s}(\mathbb{T}^d) \\ {\scriptstyle\Lambda_s}\downarrow & & \downarrow{\scriptstyle\Lambda_s} \\ L^2(\mathbb{T}^d) & \xrightarrow{Q_\nu} & L^2(\mathbb{T}^d), \end{array}$$

where $Q_\nu = \Lambda_s\widehat{F}_\nu\Lambda_s^{-1}$. Denote $P_\nu = Q_\nu^*Q_\nu$. Then we have

$$P_\nu = Q_\nu^*Q_\nu = (\Lambda_s^{-1})^*\left[\widehat{F}_\nu^*(\Lambda_s^*\Lambda_s)\widehat{F}_\nu\right]\Lambda_s^{-1} : L^2(\mathbb{T}^d)\to L^2(\mathbb{T}^d).$$

Note that $\sqrt{P_\nu}$ and $\widehat{F}_\nu$ share the same spectrum in modulus.

Denote by $p(x,\xi)$ the principal symbol of P_ν (see Subsection 2.6 for definition). By Theorem 2.9 and L^2-continuity theorem (Theorem 2.13), the operator $P_\nu : L^2(\mathbb{T}^d)\to L^2(\mathbb{T}^d)$ is of order 0 and hence bounded such that $P_\nu = K + R$, where K is compact. By Lemma 4.1 below, the definition of g in the previous subsection, and the definition γ in (1.1), we get

$$\begin{aligned}\|R\|_{L^2\to L^2} &\leqslant \sup_x \limsup_{|\xi|\to\infty} |p(x,\xi)| + \varepsilon \\ &= \sup_x \limsup_{|\xi|\to\infty} \sum_{x=Ty}\mathcal{A}(y)\left(\frac{g((D_yT)^t\xi)}{g(\xi)}\right)^{2s} + \varepsilon \\ &\leqslant \sup_x \sum_{x=Ty}\mathcal{A}(y)\limsup_{|\xi|\to\infty}\left(\frac{|(D_yT)^t\xi|}{|\xi|}\right)^{2s} + \varepsilon \\ &\leqslant \sup_x \sum_{x=Ty}\mathcal{A}(y)\gamma^{2s} + \varepsilon = \gamma^{2s}+\varepsilon.\end{aligned}$$

Choose ε small enough such that $\rho_1 := \sqrt{\gamma^{2s}+\varepsilon} < 1$. By polar decomposition, $Q_\nu = \sqrt{P_\nu}U_\nu$ for some unitary operator $U_\nu : L^2(\mathbb{T}^d)\to L^2(\mathbb{T}^d)$. By spectral theorem, Q_ν also has a similar decomposition $Q_\nu = K_1 + R_1$ such that K_1 is compact and $\|R_1\|\leqslant\rho_1$. By unitary equivalence between $Q_\nu : L^2(\mathbb{T}^d)\to L^2(\mathbb{T}^d)$ and $\widehat{F}_\nu : H_{\Lambda_s}(\mathbb{T}^d)\to H_{\Lambda_s}(\mathbb{T}^d)$, there is a similar decomposition $\widehat{F}_\nu = K_\nu + R_\nu$ such that K_ν is compact, and $\|R_\nu|H_{\Lambda_s}(\mathbb{T}^d)\|\leqslant\rho_1$. Choose $C_1>0$ such that

$$\frac{1}{\sqrt{C_1}} \leqslant \frac{\|\varphi\|_{\Lambda_s}}{\|\varphi\|_s} \leqslant \sqrt{C_1} \quad \text{for any} \quad \varphi\in H^s(\mathbb{T}^d),$$

then we have

$$\|R_\nu^n|H^s(\mathbb{T}^d)\| \leqslant C_1\|R_\nu^n|H_{\Lambda_s}(\mathbb{T}^d)\| \leqslant C_1\|R_\nu|H_{\Lambda_s}(\mathbb{T}^d)\|^n \leqslant C_1\rho_1^n.$$

This completes the proof of Proposition 3.1. ◻

Lemma 4.1. *$P_\nu \in \mathrm{OP}S^0$ has a principal symbol*

$$p(x,\xi) = \sum_{x=Ty} \mathcal{A}(y) \left(\frac{g((D_yT)^t\xi)}{g(\xi)} \right)^{2s},$$

where $\mathcal{A}(y)$ is defined in (3.3).

Proof. Note that Λ_s has a principal symbol λ_s. By Theorem 2.9, $\Lambda_s^* \in \mathrm{OP}S^s$ has a principal symbol λ_s, and $\Lambda_s^{-1}, (\Lambda_s^{-1})^* \in \mathrm{OP}S^{-s}$ both have a principal symbol λ_s^{-1}. Further, $\Lambda_s^*\Lambda_s \in \mathrm{OP}S^{2s}$ has a principal symbol λ_s^2. Then by Egorov's theorem 2.11, $\widehat{F}_\nu^*(\Lambda_s^*\Lambda_s)\widehat{F}_\nu \in \mathrm{OP}S^{2s}$ has a principal symbol

$$\begin{aligned} \overline{a}_0(y,\eta) &= \sum_{\substack{y=Tx,\\ \eta=[(D_xT)^t]^{-1}\xi}} \lambda_s^2(x,\xi) \left| e^{i2\pi\nu\cdot\tau(x)} \right|^2 |\det(D_xT)^t|^{-1} \\ &= \sum_{y=Tx} \frac{\lambda_s^2(x,(D_xT)^t\eta)}{|\mathrm{Jac}(T)(x)|}. \end{aligned}$$

Use the composition rule again and recall the definition of λ_s in (4.3), we have $P_\nu \in \mathrm{OP}S^0$ with a principal symbol

$$\begin{aligned} p(y,\eta) &= \sum_{y=Tx} \frac{\lambda_s^2(x,(D_xT)^t\eta)}{|\mathrm{Jac}(T)(x)|} \frac{1}{\lambda_s^2(y,\eta)} \\ &= \sum_{y=Tx} \frac{1}{|\mathrm{Jac}(T)(x)|} \frac{h(x)}{h(y)} \frac{g((D_xT)^t\eta)^{2s}}{g(\eta)^{2s}} \\ &= \sum_{y=Tx} \mathcal{A}(x) \left(\frac{g((D_xT)^t\eta)}{g(\eta)} \right)^{2s}. \end{aligned}$$

This is what we need. ◻

4.3. **Proof of Proposition 3.2.** While we use the classical formulation to prove Proposition 3.1, we use the semiclassical one to show Proposition 3.2. The key step of the proof is the estimate stated and proved in Lemma 5.1 in the next section.

Proof of Proposition 3.2. Let $\nu \in \mathbb{Z}^\ell \backslash \{\mathbf{0}\}$. Take $\hbar = \dfrac{1}{|\nu|}$.

In Definition 2.5, we choose amplitude $a(x,\xi) = 1 \in S^0$ and phase function

$$S_\nu(x,\xi) = Tx\cdot\xi + \frac{\nu}{|\nu|}\cdot\tau(x). \tag{4.4}$$

Then by Remark 2.6(ii), $\widehat{F}_\nu$ is an $\hbar$-scaled FIO of the form

$$\widehat{F}_\nu\varphi(x) = \varphi(Tx)e^{i2\pi\nu\cdot\tau(x)} = \int_{T^*\mathbb{T}^d} e^{i2\pi\frac{1}{\hbar}\{[Tx\cdot\xi + \frac{\nu}{|\nu|}\cdot\tau(x)] - y\cdot\xi\}}\varphi(y)dyd\omega_\hbar(\xi).$$

Inductively, we have

$$\begin{aligned} \widehat{F}_\nu^n\varphi(x) =& \varphi(T^nx)e^{i2\pi\nu\cdot\sum_{k=0}^{n-1}\tau(T^kx)} \\ =& \int e^{i2\pi\frac{1}{\hbar}\{[T^nx\cdot\xi + \frac{\nu}{|\nu|}\cdot\sum_{k=0}^{n-1}\tau(T^kx)] - y\cdot\xi\}}\varphi(y)dyd\omega_\hbar(\xi), \end{aligned}$$

where the phase function now is

$$S_{\nu,n}(x,\xi) = T^n x \cdot \xi + \frac{\nu}{|\nu|} \cdot \sum_{k=0}^{n-1} \tau(T^k x).$$

The canonical transformation $(x,\xi) \mapsto (y,\eta)$ associated to $\widehat{F}_\nu^n$ is given by

$$y = T^n x, \qquad \eta = [(D_x T^n)^t]^{-1}[\xi - W_{\nu,n}(x)],$$

where

$$W_{\nu,n}(x) = W_n(x) \cdot \frac{\nu}{|\nu|} \quad \text{and} \quad W_n(x) = \sum_{k=0}^{n-1} (D_x T^k)^t [D_{T^k x}\tau]^t. \tag{4.5}$$

The following commutative diagram

$$\begin{array}{ccc} H_{\Lambda_{s,\hbar}}(\mathbb{T}^d) & \xrightarrow{\widehat{F}_\nu^n} & H_{\Lambda_{s,\hbar}}(\mathbb{T}^d) \\ {\scriptstyle \Lambda_{s,\hbar}}\downarrow & & \downarrow{\scriptstyle \Lambda_{s,\hbar}} \\ L^2(\mathbb{T}^d) & \xrightarrow{\widetilde{Q}_\nu^n} & L^2(\mathbb{T}^d) \end{array}$$

suggests that we study the operator

$$\widetilde{P}_{\nu,n} = (\widetilde{Q}_\nu^n)^* \widetilde{Q}_\nu^n = (\Lambda_{s,\hbar}^{-1})^* \left[(\widehat{F}_\nu^n)^* (\Lambda_{s,\hbar}^* \Lambda_{s,\hbar}) \widehat{F}_\nu^n \right] \Lambda_{s,\hbar}^{-1} : L^2(\mathbb{T}^d) \to L^2(\mathbb{T}^d)$$

since the spectral radius

$$\mathrm{Sp}(\widehat{F}_\nu | H_{\Lambda_{s,\hbar}}(\mathbb{T}^d)) = \mathrm{Sp}(\widetilde{Q}_\nu | L^2(\mathbb{T}^d)) = \lim_{n\to\infty} \sqrt[2n]{\|\widetilde{P}_{\nu,n} | L^2(\mathbb{T}^d)\|}. \tag{4.6}$$

By Lemma 4.2 below, for all $n \in \mathbb{N}$ and $\nu \in \mathbb{Z}^\ell \backslash \{\mathbf{0}\}$, $\widetilde{P}_{\nu,n} \in \mathrm{OP}_\hbar S^0$ has a principal symbol $\widetilde{p}_{\nu,n}(x,\xi)$ given by (4.7). By Theorem 2.14 and Lemma 5.1 in the next section, we know that there exist $n_0 \in \mathbb{N}$ and $\widetilde{p}_0 < 1$ such that

$$\|\widetilde{P}_{\nu,n_0} | L^2(\mathbb{T}^d)\| \leqslant \sup_{(x,\xi)\in T^*\mathbb{T}^d} \widetilde{p}_{\nu,n_0}(x,\xi) + \hbar C_d \leqslant \widetilde{p}_0 + \frac{C_d}{|\nu|}$$

for some constant $C_d > 0$. Choose $\nu_1 > 0$ such that

$$\rho_2 := \sqrt[2n_0]{\widetilde{p}_0 + \frac{C_d}{\nu_1}} < 1,$$

then $\sup_{|\nu|\geqslant\nu_1} \|\widetilde{P}_{\nu,n_0} | L^2(\mathbb{T}^d)\| \leqslant \rho_2^{2n_0}$. By (4.6) and the subadditivity of the sequence $\{\|\widetilde{P}_{\nu,n} | L^2(\mathbb{T}^d)\|\}_{n\in\mathbb{N}}$, we have the spectral radius

$$\mathrm{Sp}(\widehat{F}_\nu | H^s(\mathbb{T}^d)) = \mathrm{Sp}(\widehat{F}_\nu | H_{\Lambda_{s,\hbar}}(\mathbb{T}^d)) = \lim_{n\to\infty} \sqrt[2n]{\|\widetilde{P}_{\nu,n} | L^2(\mathbb{T}^d)\|} \leqslant \rho_2.$$

This completes the proof of Proposition 3.2. □

Lemma 4.2. *For each $n \in \mathbb{N}$ and $\nu \in \mathbb{Z}^\ell \backslash \{\mathbf{0}\}$, $\widetilde{P}_{\nu,n} \in \mathrm{OP}_\hbar S^0$ has a principal symbol*

$$\widetilde{p}_{\nu,n}(x,\xi) = \sum_{x=T^n y} \mathcal{A}_n(y) \left(\frac{g((D_y T^n)^t \xi + W_{\nu,n}(y))}{g(\xi)} \right)^{2s}, \tag{4.7}$$

where $\mathcal{A}_n(y)$ is given in (3.4) and $W_{\nu,n}(y) = \frac{\nu}{|\nu|} W_n(y)$ is given in (4.5).

Proof. Note that $\Lambda_{s,\hbar} \in \mathrm{OP}_\hbar S^s$ has a principal symbol λ_s. By Theorem 2.9, $\Lambda^*_{s,\hbar} \in \mathrm{OP}_\hbar S^s$ has a principal symbol λ_s, and $\Lambda^{-1}_{s,\hbar}$, $(\Lambda^{-1}_{s,\hbar})^* \in \mathrm{OP}_\hbar S^{-s}$ both have a principal symbol λ_s^{-1}. Further, $\Lambda^*_{s,\hbar}\Lambda_{s,\hbar} \in \mathrm{OP}_\hbar S^{2s}$ has a principal symbol λ_s^2. By Remark 2.12 for Egorov's theorem, $(\widehat{F}^n_\nu)^*(\Lambda^*_{s,\hbar}\Lambda_{s,\hbar})\widehat{F}^n_\nu \in \mathrm{OP}_\hbar S^{2s}$ has a principal symbol

$$\begin{aligned}\overline{a}^n_0(y,\eta) &= \sum_{\substack{y=T^n x,\\ \eta=[(D_xT^n)^t]^{-1}[\xi-W_{\nu,n}(x)]}} \lambda_s^2(x,\xi)\cdot 1^2\cdot |\det(D_xT^n)^t|^{-1}\\ &= \sum_{y=T^n x} \frac{\lambda_s^2(x,(D_xT^n)^t\eta+W_{\nu,n}(x))}{|\mathrm{Jac}(T^n)(x)|}.\end{aligned}$$

By composition rule again, $\widetilde{P}_{\nu,n} \in \mathrm{OP}_\hbar S^0$ has a principal symbol

$$\begin{aligned}\widetilde{p}_{\nu,n}(y,\eta) &= \sum_{y=T^n x} \frac{\lambda_s^2(x,(D_xT^n)^t\eta+W_{\nu,n}(x))}{|\mathrm{Jac}(T^n)(x)|}\frac{1}{\lambda_s^2(y,\eta)}\\ &= \sum_{y=T^n x} \frac{1}{|\mathrm{Jac}(T^n)(x)|}\frac{h(x)}{h(y)}\frac{g((D_xT^n)^t\eta+W_{\nu,n}(x))^{2s}}{g(\eta)^{2s}}\\ &= \sum_{y=T^n x} \mathcal{A}_n(x)\left(\frac{g((D_xT^n)^t\eta+W_{\nu,n}(x))}{g(\eta)}\right)^{2s}.\end{aligned}$$

This is what we need. $\square$

5. The Principal Symbol: Proof of Lemma 5.1

The estimates given in Lemma 5.1 in the section is the most important step to obtain the radius of the spectrum of $\widehat{F}_\nu|H^s(\mathbb{T}^d)$ strictly less than 1.

Lemma 5.1. *If $\tau(x)$ is not an essential coboundary, then there exists $n_0 \in \mathbb{N}$ such that*

$$\widetilde{p}_0 := \sup_{\nu\in\mathbb{Z}^\ell\backslash\{\mathbf{0}\}}\ \sup_{(x,\xi)\in T^*\mathbb{T}^d} \widetilde{p}_{\nu,n_0}(x,\xi) < 1. \tag{5.1}$$

Proof. Given $\nu \in \mathbb{Z}^\ell\backslash\{\mathbf{0}\}$ and $x \in \mathbb{T}^d$, we consider the affine map $\mathcal{F}_{\nu,x} : \mathbb{R}^d \to \mathbb{R}^d$ given by

$$\mathcal{F}_{\nu,x}(\xi) = (D_xT)^t\xi + [D_x\tau]^t\frac{\nu}{|\nu|}, \tag{5.2}$$

and the n-th iterates

$$\mathcal{F}^n_{\nu,x}(\xi) = \prod_{k=0}^{n-1}\mathcal{F}_{\nu,T^kx}(\xi) = (D_xT^n)^t\xi + W_{\nu,n}(x). \tag{5.3}$$

Therefore, we can rewrite (4.7) as

$$\widetilde{p}_{\nu,n}(x,\xi) = \sum_{x=T^n y}\mathcal{A}_n(y)\left[\frac{g(\mathcal{F}^n_{\nu,y}(\xi))}{g(\xi)}\right]^{2s}.$$

By Sublemma 5.3 below, for any $\nu \in \mathbb{Z}^\ell\backslash\{\mathbf{0}\}$, there is $n_0(\nu) > 0$ such that

$$\sup_{(x,\xi)\in T^*\mathbb{T}^d}\widetilde{p}_{\nu,n_0(\nu)}(x,\xi) < 1.$$

We can strength it as follows. By Sublemma 5.2 below, $\widetilde{p}_{\nu,n_0(\nu)}(x,\xi) < 1$ implies that there is $y \in T^{-n}(x)$ such that $|\mathcal{F}_{\nu,y}^{n_0(\nu)}(\xi)| > R$. Since $\mathcal{F}_{\nu,y}^{n_0(\nu)}(\xi)$ depends on $\dfrac{\nu}{|\nu|}$ continuously, there is $\varepsilon(\nu) > 0$ such that $|\mathcal{F}_{\nu',y}^{n_0(\nu)}(\xi)| > R$ whenever $\left|\dfrac{\nu'}{|\nu'|} - \dfrac{\nu}{|\nu|}\right| < \varepsilon(\nu)$. Moreover, by proof of Sublemma 5.2, we know that

$$|\mathcal{F}_{\nu',y}^{n}(\xi)| \geqslant |\mathcal{F}_{\nu',y}^{n_0(\nu)}(\xi)| > R,$$

for any $n \geqslant n_0(\nu)$. By Sublemma 5.2 again, we have

$$\sup_{(x,\xi)\in T^*\mathbb{T}^d} \widetilde{p}_{\nu',n}(x,\xi) < 1,$$

for any $|\nu'/|\nu'| - \nu/|\nu|| < \varepsilon(\nu)$ and $n \geqslant n_0(\nu)$.

Since all $\{\nu/|\nu|\}_{\nu\in\mathbb{Z}^\ell}$ lies on the unit $(\ell-1)$-sphere $\mathbb{S}^{\ell-1}$, which is compact, there are $\nu_1, \nu_2, \dots, \nu_k \in \mathbb{Z}^\ell$ such that the finite collection of open balls $\{B(\nu_j/|\nu_j|, \varepsilon(\nu_j)\}_{1\leqslant j\leqslant k}$ covers $\mathbb{S}^{\ell-1}$. Therefore, we obtain (5.1) if we set

$$n_0 = \max\{n_0(\nu_1), \dots, n_0(\nu_k)\}.$$

Then we can obtain what we need. $\square$

Sublemma 5.2. *Let $\nu \in \mathbb{Z}^\ell\backslash\{\mathbf{0}\}$.*

(1) $\widetilde{p}_{\nu,n}(x,\xi) \leqslant 1$ for all $n \in \mathbb{N}$ and $(x,\xi) \in T^*\mathbb{T}^d$;

(2) $\widetilde{p}_{\nu,n}(x,\xi) < 1$ if and only if there is $y \in T^{-n}x$ such that $|\mathcal{F}_{\nu,y}^{n}(\xi)| > R$. In other words, $\widetilde{p}_{\nu,n}(x,\xi) = 1$ if and only if $|\mathcal{F}_{\nu,y}^{k}(\xi)| \leqslant R$ for any $y \in T^{-n}x$ and $0 \leqslant k \leqslant n$.

Proof. First recall that $\sum_{x=T^ny} \mathcal{A}_n(y) = 1$, and every $\mathcal{A}_n(y)$ is positive.

Given $\xi \in \mathbb{R}^d$ with $|\xi| > R$, we observe that all iterates $\mathcal{F}_{\nu,y}^{k}(\xi)$ will be outside the interval $[-R, R]$. Indeed, by the choice of R in (4.1), we have $|\xi| > \dfrac{2\|D\tau\|}{\gamma-1}$. So by (5.2)

$$|\mathcal{F}_{\nu,y}(\xi)| \geqslant |(D_yT)^t\xi| - \left|[D_y\tau(y)]^t \frac{\nu}{|\nu|}\right| \geqslant \gamma|\xi| - \|D\tau\| \geqslant \gamma|\xi| - \frac{\gamma-1}{2}|\xi| = \frac{\gamma+1}{2}|\xi|,$$

and hence for all $k \geqslant 0$,

$$|\mathcal{F}_{\nu,y}^{k}(\xi)| \geqslant \left(\frac{\gamma+1}{2}\right)^k |\xi| \geqslant |\xi| > R.$$

By the definition of $g(\xi)$ defined in (4.2), the quotient

$$\left[\frac{g(\mathcal{F}_{\nu,y}^{n}(\xi))}{g(\xi)}\right]^{2s} \begin{cases} = 1 & \text{if } |\mathcal{F}_{\nu,y}^{k}(\xi)| \leqslant R \text{ for all } 0 \leqslant k \leqslant n; \\ \\ < 1 & \text{otherwise.} \end{cases}$$

In either case, we always get $\left[\dfrac{g(\mathcal{F}_{\nu,y}^{n}(\xi))}{g(\xi)}\right]^{2s} \leqslant 1$ and hence

$$\widetilde{p}_{\nu,n}(x,\xi) = \sum_{x=T^ny} \mathcal{A}_n(y) \left[\frac{g(\mathcal{F}_{\nu,y}^{n}(\xi))}{g(\xi)}\right]^{2s} \leqslant \sum_{x=T^ny} \mathcal{A}_n(y) = 1.$$

Now clearly, $\widetilde{p}_{\nu,n}(x,\xi) = 1$ if and only if we are in the first case, that is, $|\mathcal{F}^k_{\nu,y}(\xi)| \leqslant R$ for any $y \in T^{-n}x$ and $0 \leqslant k \leqslant n$. $\square$

Next we show that

Sublemma 5.3. *Suppose $\tau(x)$ is not an essential coboundary. Then for any $\nu \in \mathbb{Z}^\ell \backslash \{\mathbf{0}\}$, there is $n_0(\nu) \in \mathbb{N}$ such that*

$$\sup_{(x,\xi)\in T^*\mathbb{T}^d} \widetilde{p}_{\nu,n_0(\nu)}(x,\xi) < 1.$$

Proof. Let us argue by contradiction. If

$$\sup_{(x,\xi)\in T^*\mathbb{T}^d} \widetilde{p}_{\nu,n}(x,\xi) = 1$$

for all $n \in \mathbb{N}$, then according to the proof of Sublemma 5.2, we actually have

$$\sup_{(x,\xi)\in T^*\mathbb{T}^d} \widetilde{p}_{\nu,n}(x,\xi) = \max_{(x,\xi)\in\mathbb{T}^d\times[-R,R]} \widetilde{p}_{\nu,n}(x,\xi) = 1.$$

So for each $n \in \mathbb{N}$, there is $(x_n, \xi_n) \in \mathbb{T}^d \times [-R, R]$ such that $\widetilde{p}_{\nu,n}(x_n, \xi_n) = 1$.

Using (5.3) and then Sublemma 5.2 (2), we get that

$$|\mathcal{F}^k_{\nu,y}(\xi_n)| = |(D_yT^k)^t\xi_n + W_{\nu,k}(y)| = \left|(D_yT^k)^t\left(\xi_n + \widetilde{W}_{\nu,k}(y)\right)\right| \leqslant R,$$

for any $y \in T^{-n}(x_n)$ and $1 \leqslant k \leqslant n$, where

$$\widetilde{W}_{\nu,k}(y) = [(D_yT^k)^t]^{-1}W_{\nu,k}(y) = \sum_{j=0}^{k-1}[(D_{T^jy}T^{k-j})^t]^{-1}[D_{T^jy}\tau]^t\frac{\nu}{|\nu|}. \tag{5.4}$$

By (1.1), we have

$$\left|\xi_n + \widetilde{W}_{\nu,k}(y)\right| \leqslant \frac{R}{\gamma^k}, \qquad \forall y \in T^{-n}(x_n), \quad 1 \leqslant k \leqslant n. \tag{5.5}$$

We would like to rewrite $\widetilde{W}_{\nu,k}(y)$ in terms of x_n as follows. Suppose the degree of the expanding endomorphism $T : \mathbb{T}^d \to \mathbb{T}^d$ is N. We denote

$$\Sigma^n_N = \{\mathbf{i} = (i_1, i_2, \ldots, i_n) : \ i_j = 0, 1, \ldots, N-1\}, \quad 1 \leqslant n \leqslant \infty. \tag{5.6}$$

Let $T_0^{-1}, T_1^{-1}, \ldots, T_{N-1}^{-1}$ be the inverse branches of T. Given $x \in \mathbb{T}^d$ and $\mathbf{i} \in \Sigma^n_N$, for any $1 \leqslant k \leqslant n$, we denote $T_{\mathbf{i}}^{-k}x = T_{i_k}^{-1}\ldots T_{i_1}^{-1}x$. So $T_{\mathbf{i}}^{-k}x$ is well defined whenever $k \leqslant n \leqslant \infty$.

For all $x \in \mathbb{T}^d$, $\mathbf{i} \in \Sigma^n_N$ and $1 \leqslant k \leqslant n < \infty$, we define

$$V_{\nu,k}(\mathbf{i}, x) := \sum_{j=1}^{k} D_x\left[\frac{\nu}{|\nu|}\cdot\tau(T_{\mathbf{i}}^{-j}(x))\right] = \sum_{j=1}^{k}[(D_{T_{\mathbf{i}}^{-j}x}T^j)^t]^{-1}[D_{T_{\mathbf{i}}^{-j}x}\tau]^t\frac{\nu}{|\nu|}. \tag{5.7}$$

Since for any $x \in \mathbb{T}^d$,

$$\sum_{j=k}^{\infty}\left|[(D_{T_{\mathbf{i}}^{-j}x}T^j)^t]^{-1}[D_{T_{\mathbf{i}}^{-j}x}\tau]^t\frac{\nu}{|\nu|}\right| \leqslant |D\tau|_\infty\sum_{j=k}^{\infty}\gamma^{-j} \to 0 \quad \text{as } k \to \infty, \tag{5.8}$$

and the convergence is uniform, $V_{\nu,k}(\mathbf{i}, x)$ can be defined for $\mathbf{i} \in \Sigma^\infty_N$ and $k = \infty$. We denote $V_\nu(\mathbf{i}, x) = V_{\nu,\infty}(\mathbf{i}, x)$ in this case.

Comparing (5.4) and (5.7), we see $\widetilde{W}_{\nu,k}(y) = V_{\nu,k}(\mathbf{i}, x)$ whenever $y = T_{\mathbf{i}}^{-n}(x)$. Therefore, we can rewrite (5.5) as

$$|\xi_n + V_{\nu,k}(\mathbf{i}, x_n)| \leqslant \frac{R}{\gamma^k}$$

for any $\mathbf{i} \in \Sigma_N^n$ and $1 \leqslant k \leqslant n$. Since $\{(x_n, \xi_n)\}$ lies in the compact region $\mathbb{T}^d \times [-R, R]$, there is an accumulation point (x^*, ξ^*). By taking $n \to \infty$ and $k \to \infty$ in the above inequality, we get $V_\nu(\mathbf{i}, x^*) = -\xi^*$, regardless of the choice for $\mathbf{i} \in \Sigma_N^\infty$.

For $x \in \mathbb{T}^d$, take $w \in \{0, 1, \dots, N-1\}$ such that such that $x = T_w^{-1}(Tx)$. For any $\mathbf{i} \in \Sigma_N^\infty$, we can check directly using (5.7), to get

$$(D_{T_w^{-1}x}T)^t V_\nu(w\mathbf{i}, x) = V_\nu(\mathbf{i}, T_w^{-1}x) + [D_{T_w^{-1}x}\tau]^t \frac{\nu}{|\nu|}. \tag{5.9}$$

By Claim 1 below we know that $V_\nu(\mathbf{i}, x)$ is independent of $\mathbf{i}$. Hence, we can define a function $V_\nu : \mathbb{T}^d \to \mathbb{R}^\ell$ by

$$V_\nu(x) = V_\nu(\mathbf{i}, x), \quad \text{for any } \mathbf{i} \in \Sigma_N^\infty, \tag{5.10}$$

if $x \in \mathbb{T}^d$. Replacing x by Tx in (5.9), we get

$$(D_x T)^t V_\nu(T(x)) = V_\nu(x) + [D_x\tau]^t \frac{\nu}{|\nu|}. \tag{5.11}$$

By Claim 2, the function $u : \mathbb{T}^d \to \mathbb{T}$ given by

$$u(x) = \int_0^1 V_\nu(tx_1, tx_2, \dots, tx_d)\ dt$$

is well-defined. Integrating (5.11), we get

$$u(Tx) = u(x) + \frac{\nu}{|\nu|} \cdot \tau(x) + c$$

for some constant $c \in \mathbb{T}$, which contradicts to the fact that $\tau(x)$ is not a directional essential coboundary. □

Claim 1. *For any $x \in \mathbb{T}^d$, $V_\nu(\mathbf{i}, x)$ is independent of $\mathbf{i}$, that is, $V_\nu(\mathbf{i}, x) = V_\nu(\mathbf{i}', x)$ for all $\mathbf{i}, \mathbf{i}' \in \Sigma_N^\infty$.*

Proof. Take $x = x^*$ and using the fact $V_\nu(\mathbf{i}, x^*) = -\xi^*$ in (5.9), we can get

$$V_\nu(\mathbf{i}, T_w^{-1}x^*) = -(D_{T_w^{-1}x^*}T)^t \xi^* - [D_{T_w^{-1}x^*}\tau]^t \frac{\nu}{|\nu|}.$$

The right hand side is independent of $\mathbf{i}$, and hence, $V_\nu(\mathbf{i}, T_w^{-1}x^*) = V_\nu(\mathbf{0}, T_w^{-1}x^*)$, where $\mathbf{0} = (0, 0, \dots) \in \Sigma_N^\infty$.

Inductively, one can show that $V_\nu(\mathbf{i}, x) = V_\nu(\mathbf{0}, x)$ for all $x \in \bigcup_{n=1}^\infty T^{-n}(x^*)$ and thus for all $x \in \mathbb{T}^d$, since the set $\bigcup_{n=1}^\infty T^{-n}x^*$ is dense in $\mathbb{T}^d$. □

Claim 2. *For any $x = (x_1, x_2, \dots, x_d) \in \mathbb{T}^d = (\mathbb{R}/\mathbb{Z})^d$ and $1 \leqslant k \leqslant d$,*

$$\int_0^1 V_\nu(x_1, \dots, x_{k-1}, t, x_{k+1}, \dots, x_d)\ dt = 0.$$

Proof. By (5.10) and (5.7), we rewrite for arbitrary $M \in \mathbb{N}$,

$$
\begin{aligned}
&V_\nu(x) = \frac{1}{N^M} \sum_{\mathbf{i} \in \Sigma_N^M} V_\nu(\mathbf{i0}, x) \\
=& \frac{1}{N^M} \sum_{\mathbf{i} \in \Sigma_N^M} V_{\nu,M}(\mathbf{i0}, x) + \frac{1}{N^M} \sum_{\mathbf{i} \in \Sigma_N^M} [V_\nu(\mathbf{i0}, x) - V_{\nu,M}(\mathbf{i0}, x)] \\
=& \frac{1}{N^M} \sum_{j=1}^{M} \sum_{\mathbf{i} \in \Sigma_N^j} D_x \left[\frac{\nu}{|\nu|} \cdot \tau(T_{i_j}^{-1} \ldots T_{i_1}^{-1}(x)) \right] + \frac{1}{N^M} \sum_{\mathbf{i} \in \Sigma_N^M} [V_\nu(\mathbf{i0}, x) - V_{\nu,M}(\mathbf{i0}, x)].
\end{aligned}
$$

For the first term, we have

$$
\begin{aligned}
&\int_0^1 \sum_{\mathbf{i} \in \Sigma_N^j} D_x \left[\frac{\nu}{|\nu|} \cdot \tau(T_{i_j}^{-1} \ldots T_{i_1}^{-1}(x_1, \ldots, x_{k-1}, t, x_{k+1}, \ldots, x_d)) \right] dt \\
=& \frac{\nu}{|\nu|} \cdot \sum_{\mathbf{i} \in \Sigma_N^j} \Big[\tau(T_{i_j}^{-1} \ldots T_{i_1}^{-1}(x_1, \ldots, x_{k-1}, 1, x_{k+1}, \ldots, x_d)) \\
&\qquad - \tau(T_{i_j}^{-1} \ldots T_{i_1}^{-1}(x_1, \ldots, x_{k-1}, 0, x_{k+1}, \ldots, x_d)) \Big]
\end{aligned}
$$

Since $(x_1, \ldots, x_{k-1}, 0, x_{k+1}, \ldots, x_d)$ and $(x_1, \ldots, x_{k-1}, 1, x_{k+1}, \ldots, x_d)$ are the same point in $\mathbb{T}^d$, and both $\{T_{i_j}^{-1} \ldots T_{i_1}^{-1}(x_1, \ldots, x_{k-1}, 0, x_{k+1}, \ldots, x_d) : \mathbf{i} \in \Sigma_N^j\}$ and $\{T_{i_j}^{-1} \ldots T_{i_1}^{-1}(x_1, \ldots, x_{k-1}, 1, x_{k+1}, \ldots, x_d) : \mathbf{i} \in \Sigma_N^j\}$ are the sets of all jth preimages of the point. Hence they are the same on $\mathbb{T}^d$. Since τ is a function on $\mathbb{T}^d$, the right hand side of the equality must be 0.

On the other hand, by (5.8) the convergence $V_{\nu,M}(\mathbf{i0}, x) \to V_\nu(\mathbf{i0}, x)$ is uniform in $\mathbf{i}$ as $M \to \infty$. By choosing M large enough, the integral of $V_\nu(x)$ is arbitrary small and hence 0. $\square$

References

[1] Oliver Butterley and Peyman Eslami. Exponential mixing for skew products with discontinuities. *Preprint*, 2014.

[2] Dmitry Dolgopyat. On decay of correlations in anosov flows. *Ann. of Math. (2)*, 147:357–390, 1998.

[3] Dmitry Dolgopyat. On mixing properties of compact group extensions of hyperbolic systems. *Israel J. Math.*, 130:157–205, 2002.

[4] Yu. V. Egorov. *Linear differential equations of principal type*. Contemporary Soviet Mathematics. Consultants Bureau, New York, 1986. Translated from the Russian by Dang Prem Kumar.

[5] Frédéric Faure. Semiclassical origin of the spectral gap for transfer operators of a partially expanding map. *Nonlinearity*, 24(5):1473–1498, 2011.

[6] Frédéric Faure, Nicolas Roy, and Johannes Sjöstrand. Semi-classical approach for Anosov diffeomorphisms and Ruelle resonances. *Open Math. J.*, 1:35–81, 2008.

[7] Oliver Jenkinson. Smooth cocycle rigidity for expanding maps, and an application to Mostow rigidity. *Math. Proc. Cambridge Philos. Soc.*, 132(3):439–452, 2002.

[8] André Martinez. *An introduction to semiclassical and microlocal analysis*. Universitext. Springer-Verlag, New York, 2002.

[9] Frédéric Naud. Heat kernel and rates of mixing for compact extensions of expanding maps. *Preprint*, 2011.

[10] W. Parry and M. Pollicott. Stability of mixing for toral extensions of hyperbolic systems. *Tr. Mat. Inst. Steklova*, 216(Din. Sist. i Smezhnye Vopr.):354–363, 1997.

[11] Mark Pollicott. A complex Ruelle-Perron-Frobenius theorem and two counterexamples. *Ergodic Theory Dynam. Systems*, 4(1):135–146, 1984.

[12] Mark Pollicott. On the mixing of Axiom A attracting flows and a conjecture of Ruelle. *Ergodic Theory Dynam. Systems*, 19(2):535–548, 1999.

[13] David Ruelle. *Thermodynamic formalism*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 2004. The mathematical structures of equilibrium statistical mechanics.

[14] Michael E. Taylor. *Partial differential equations I. Basic theory*, volume 115 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2011.

[15] Masato Tsujii. Decay of correlations in suspension semi-flows of angle-multiplying maps. *Ergodic Theory Dynam. Systems*, 28(1):291–317, 2008.

[16] Maciej Zworski. *Semiclassical analysis*, volume 138 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA

E-mail address: hu@math.msu.edu, jychen@math.msu.edu